$MIP^{co} = co RE$

Junqiao Lin

$$\mathsf{MIP}^{\mathrm{co}} = \mathsf{coRE}$$

This research was supported by ERC STG grant 101040624-ASC-Q.

Printed and bound by Ipskamp Printing.

ISBN: 978-94-6536-201-4.

$\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 8 september 2026, te 16.00 uur

door

Jun Qiao Lin

geboren te Shenzhen

Promotiecommissie

Promotor:	prof. dr. S.M. Jeffery	Universiteit van Amsterdam
Copromotor:	prof. dr. J. Briët	Universiteit Leiden
Overige leden:	prof. dr. T. Vidick	EPFL
	prof. dr. R.M. de Wolf	Universiteit van Amsterdam
	prof. dr. M. Walter	Ludwig-Maximilians-Universität München
	dr. J. Zuiddam	Universiteit van Amsterdam
	prof. dr. H. Peters	Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

List of publications

This dissertation is based on the following papers (in chronological order).

[Lin24] Jun Qiao Lin, *Tracially embeddable strategies: Lifting* MIP^* *tricks to* $\mathsf{MIP}^{\mathsf{co}}$. arXiv preprint (2024). arXiv:2304.01940

Contributed talk at TQC 2024.

[Lin26] Jun Qiao Lin, $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$. Proceedings of the 58th Annual ACM Symposium on Theory of Computing. STOC'26. New York, NY, USA:Association for Computing Machinery, 2026, pp.314–322. arXiv:2510.07162

Contributed talk at QIP 2026, STOC 2026 (selected as Danny Lewin best student paper award).

To all my friends and colleagues who supported me throughout this wonderful journey.

Acknowledgement

Wow, it is still a bit hard to believe I made it this far into my academic career. From almost failing courses (funny enough, on introduction to analysis and a coding course) back in the middle of second year to getting the STOC best student paper, this has been one hell of a ride! There are many peoplewho have accompanied me on this journey who I wish to extend my sincerest gratitude to in this section.

First and foremost, I want to thank Stacey Jeffery, who has been my supervisor for the past four years. I know that I am a pain to deal with at times, but I truly appreciate that you are always there for me regardless, and I want to thank you for your wisdom and your support for the past four years. You are the best advisor I could have asked for for during this journey.

Next, I want to thank my committee, Jop, Thomas, Ronald, Micheal, Jeroen and Hans, Thank you all for taking the time to read my thesis and for providing valuable feedback.

I would like to thank Richard Cleve, Henry Yuen and William Slofstra. Richard, I truly appreciate that you took a chance on me during my undergrad and for offering me an undergraduate research position, as well as becoming my masters supervisor. Without you, I don't think I would even have become a researcher. Henry, I want to thank you for letting me sit in on your course on quantum complexity, and dragging me into the project regarding the parallel repetition for the commuting operator model. This is what began the five year project which led to this thesis and an amazing journey. William, thank you for the early guidance in paper writing and academic advice in general, and for pointing me in the right direction several times when everything seemed to be crashing down. I would have never made it this far without your help!

I want to thank Thomas Vidick for many helpful exchanges about the $\mathsf{MIP}^* = \mathsf{RE}$ result, and answering some of my stupid emails here and there. Additionally, I want to thank Anand Natarajan, John Wright and Ji Zhengfeng for producing such a hardcore result for me to ~~suffer through~~ follow up on. Jokes aside, I really do appreciate you guys taking time to have helpful discussions with me. I want to thank Andrew Marks for the insightful discussions about the compression theorem, which ended up being the favourite part of my thesis. I also want to thank Zhao Yuming for always being there to bounce ideas off and being a great, supportive friend. I want to thank Michael Chapman, Mikael de la Salle, Vern Paulsen, Laura Mančinska and many others for helpful description about non-local games. Additionally, I want to thank Conner Paul Paddock, Kieran Mastel, Eric Culf, Aareyan Manzoor, Hamoon Mousavi, Honghao Fu, Seyed Sajjad Nezhadi, Chen Ranyiliu, Taro Spirig, Sigurd Storgaard, Alexander Kulpe and Valentine Blanpain for being apart of this amazing academic journey.

I was very fortunate to have a lot of travel opportunities as a Ph.D. student. I want to thank François Le Gall, Laura Mančinska, Matthias Christandl, William Slofstra, Graeme Smith, Henry Yuen, Micheal Chapman, Arthur Mehta, Alex Grilo, Anand Natarajan, Alexander Müller-Hermes, Thomas Vidick, Matt Coudron, David Pérez-García, Ludovico Lami, Ulysse Charbad, Marc-Olivier Renou, Scott Aaronson, Conner Paul Paddock, Adam Bene Watts, Xavier Coiteux-Roy, András Gilyén, Clark Lyons, Zoltan Vidnyanszky, Tom Gur and Marco Fanizza for taking the time to host me during my visits, organizing seminars and having tons of interesting research discussions! which shaped a lot of my current academic interest.

I want to thank John Bostanci for lots of interesting discussions, and for always convincing me to go climbing during conferences. I want to thank Francisco and Antonio Anna Mele for all the amazing adventures in conferences. I want to thank Álvaro Yángüez for showing me both Paris and Madrid during my travels. I want to thank Elias Theil for all the board game events and fun times we have whenever I visited Copenhagen. I want to thank Lane Gundermen for the continuous support. Additionally, I want to thank Liu Yupan, Wu Peixue, Lia Yah, Jiang Jiaqing, Vishnu, Uma, Zoe, Nunzia and many others for making my research visit memorable.

I was really fortunate to spend the past four years in Qusoft. I want to thank Ronald de Wolf for always keeping his office door open (despite the intimidating sign), and making his wisdom available. I want to thank Florian Speelman for being available to chat whenever I have some random ideas (both research and non-research related). I want to thank Jeroen Zuiddam for a lot of conversation about research outside of quantum computing and all the random coffee conversations. I want to thank Jonas Helsen for a lot of dad jokes and wisdom and advice. Additionally, I want to thank Harry Buhrman, Christian Schaffner, Kareljan Schoutens, Jop Briët, Māris Ozols, Krystal Guo, John van de Wetering, Stefano Polla, Ludovico Lami, and Freek Witteveen for fostering such a wonderful research community.

Not trying to brag, but I was fortunate enough to be given the best office mate one can ask for during a Ph.D. degree. I want to start with Francisco Escudero Gutiérrez and Llorenç Escolà-Farràs, who also happen to be the two paranymph for my Ph.D. defence. Fran, I want to thank you for always being there for me, even during some of the lowest moments during my Ph.D. Your jokes and encouragement really did help a lot in this journey. Llorenç, your smile is the first thing I see when I walk into my office every day (fine, every week at least :P), and thank you for always bringing in positivity to our office. Additionally, I want to thank my other office mates throughout the last four years. Sebastian Zur, I really appreciate all the advice you gave when I first moved to Europe, and helping me settle down with all the support. Garazi, thank you for the schrodinger's cat plushy and organizing the crypto reading group. Luca, thank you for the chocolate, and indoctrinating me into the Italian mafia gang. Filippo, thank you for making some of the best pasta I've ever had. Lorenzo, thank you for the adventure we had in Lithuania, and the political discussions (when we are both sober and intoxicated). Philippe van Dordrecht, thank you for the amazing hikes that we did in Salt Lake city, and giving me the most questionable book in my collection. Bao Zongbo, thank you for all the fun banters in the office, 还有一起在办公室摸鱼.

I also want to thank many of my colleague within Qusoft and CWI. First I want to thank everyone involved with the foosball table (you know who you are :P) for bullying me into becoming a mediocre defender on the foosball table. I want to thank Alexander Blomenhofer for coming up with the brilliant idea of hooking up Mario kart onto a movie theatre, and the continuous support throughout the year. Marten, for the fun banters around the coffee table. Dyon, for the board game sessions (especially for agreeing to play the EU4 boardgame with me). Jordi, for carrying me up to the peak of mount Fuji. Subha, for showing off and complaining about Lego. Harold, for the fun and positive energy. Davi, for the chaotic energy around the foosball table and showing me around Cambridge. Léo, for the drinks and the interesting cryptography conversations we had. Phillip Verduyn Lunel, for the collaborations, problem solving sessions, and helping with the different Dutch translations. Yaroslav, for the mentoring advice. Dimity, for always be willing to bounce interesting research ideas off. Huang Yu-Hsuan for allowing me to join your weekly routine (of going to Jen's bing and Imi Bubble tea). Michael Yonli and Michelle Sweering for trying different restaurants around Amsterdam. Nick, for all the basketball conversations we've had. Furthermore, I want to thank Adam, Ailsa, Alicja, Ake, Aldo, Amira, Anna, Arie, Arjan, Artimis, Aryan, Carla, Carli, Chris Cade, Daan Planken, Daan Schoneveld, Doutzen, Galina, Giada, Gina, Hemant, Ido, Jelena, Joppe, Koean Groenland, Koen Leijnse, Lisa, Lynn, Mani, Maxim, Mert, Neils, Nikhil, Poojith, Quinten, Rene, Salvatore, Sarah, Sebastian Verschoor, Seenivasan, Susanne, Tony, Vania, Vladyslav, Yanlin, Zarin for making my

time as a Qusofter such a unique experience. I want to also thank Emil, Guillermo, Henrik, Ila, Irene, Joan-Anne, Lisa, Lynda, Ludo, Rik, Shane, Simona and the rest of CWI for a memorable experience for the four and half years that I was here.

Outside of academia, Amsterdam gave me countless wonderful memories as well. I want to thank Julia Kastner for indoctrinating me into the bouldering cult. I want to thank George, Ruben, David, Sina, Oussama, Sina, Lieke, Jeannine, Tim, Simon, Justin and Parsa for the weekly bouldering sessions. Additionally, I want to thank Jamy, Vincent, Eline, Jip-jan, Maxim, and Daniele for going on two amazing trips to Fontainebleau during my Ph.D. I want to thank Amanda for running the best tea shop in Amsterdam (Formocha), I really enjoy the cozy and friendly environment the shop offers.

Back in NA, I want to thank Wei Zichuan, Hanzhen Lin, Robert Huang, Amy Tan, and Caleb Suan for the continuous friendship and support since undergraduate.

Finally, I want to thank Delphine Martres, for encouraging me to stay creative, both academically and in life.

Abstract

In 2020, a landmark result by Ji, Natarajan, Vidick, Wright, and Yuen showed that MIP^*, the class of languages that can be decided by a classical verifier interacting with multiple computationally unbounded provers sharing entanglement in the tensor product model, is equal to RE. We show that the class $\mathsf{MIP}^{\mathsf{co}}$, a complexity class defined similarly to MIP^*, except that the provers share entanglement in the commuting operator model, is equal to the class coRE.[1] This shows that giving the provers two different models of entanglement leads to two completely different computational powers for interactive proof systems. Our main result, in addition to providing another disproof of the Connes embedding problem, Kirchberg's problem, and Tsirelson's problem, also yields several corollaries in continuous model theory and the theory of noncommutative polynomials that do not follow immediately from the $\mathsf{MIP}^* = \mathsf{RE}$ theorem. Along the way, we introduce several innovations to compatibility theory and the commuting operator model of entanglement which we list below:

We introduce a new equivalence condition for $\mathsf{RE}/\mathsf{coRE}$-complete problems, which we call the weakly compressible condition. In addition to non-local games, this new condition could also potentially be applicable to other classes of promise problems.

Additionally, we prove that any two-party correlation in the commuting operator model can be approximated using a *tracially embeddable strategy*, a class of strategies defined on a finite tracial von Neumann algebra, which we define in this thesis. Using this characterization, we show that many known theorems in the tensor product model, such as the so-called "rounding" theorem by Vidick [JMP 2022], as well as the quantum soundness of the quantum tensor code test from Ji et al. [FOCS 2022] extend naturally to the commuting operator model.

We also give the first information-theoretic parallel repetition proof for non-local games in the commuting operator model of entanglement. This is done by combining the anchored parallel repetition proof by Bavarian et al. [SIAM J. Comput. 2022] along with the relative entropy framework given by Araki [PRIMS 1977]. We also define several analogues of information-theoretic tools for the tracial von Neumann algebra setting, such as mutual information and relative entropy, which may be of independent interest.

By using these innovations, we prove $\mathsf{MIP}^* = \mathsf{RE}$ and $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ simultaneously by showing that the compression theorem from the original $\mathsf{MIP}^* = \mathsf{RE}$ proof satisfies the same completeness and soundness guarantees in the commuting operator model. This shows that both MIP^* and $\mathsf{MIP}^{\mathsf{co}}$ satisfy this condition through the compression theorem, and thereby establish that the uncomputability of MIP^* and $\mathsf{MIP}^{\mathsf{co}}$ can be proved under a unified framework (despite these two complexity classes being different). Notably, this approach also gives an alternative proof of the $\mathsf{MIP}^* = \mathsf{RE}$ theorem, which does not rely on the preservation of the entanglement bound. By using this compression theorem, we show that deciding whether a game has its tensor product value equal to its commuting operator value is equivalent to the halting problem. This shows that there are infinitely many examples of "Bell test" which distinguish between the tensor product model and the commuting operator model (although this also means that one cannot use a computer to perform a brute-force search for such a Bell test.). We also give a more streamlined proof of the compression theorem for non-local games by incorporating the synchronous framework used by Mousavi et al. [STOC 2022], as well as the improved Pauli basis test introduced by de la Salle [ArXiv:2204.07084].

[1] We remark that the *co* modifiers on both sides of $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ refer to different things!

Samenvatting

In 2020 toonden Ji, Natarajan, Vidick, Wright en Yuen in een baanbrekend resultaat aan dat MIP^*, de klasse van talen die kan worden beslist door een klassieke verifier die communiceert met meerdere computationeel onbegrensde provers die entanglement delen in het tensorproductmodel, gelijk is aan RE. Wij tonen aan dat de klasse $\mathsf{MIP}^{\mathsf{co}}$, een complexiteitsklasse die op vergelijkbare wijze is gedefinieerd als MIP^*, met als enige verschil dat de provers entanglement delen in het model van commuterende operatoren, gelijk is aan de klasse coRE.[2] Dit laat zien dat het toestaan van twee verschillende modellen van entanglement voor de provers leidt tot twee volledig verschillende computationele vermogens voor interactieve bewijssystemen. Ons hoofdresultaat heeft, naast een nieuw tegenbewijs voor het Connes-inbeddingsprobleem, het probleem van Kirchberg en het probleem van Tsirelson, ook verschillende gevolgen voor de continue modeltheorie en de theorie van niet-commutatieve polynomen die niet onmiddellijk volgen uit de stelling $\mathsf{MIP}^* = \mathsf{RE}$. In dit proefschrift introduceren we verschillende vernieuwingen in de compatibiliteitstheorie en in het model van entanglement met commuterende operatoren, die we hieronder opsommen.

We introduceren een nieuwe equivalentievoorwaarde voor RE-/coRE-volledige problemen, die we de zwak comprimeerbare voorwaarde noemen. Naast niet-lokale spellen zou deze nieuwe voorwaarde mogelijk ook kunnen worden toegepast op andere klassen van belofteproblemen.

Daarnaast bewijzen we dat elke tweepartijencorrelatie in het model van commuterende operatoren kan worden benaderd met behulp van een *traciëel inbedbare strategie*, een klasse van strategieën die we in dit proefschrift definiëren op eindige traciële von Neumann-algebra's. Met behulp van deze karakterisering tonen we aan dat veel bekende stellingen uit het tensorproductmodel, zoals de zogenoemde "afrondingsstelling" van Vidick [JMP 2022] en de quantumcorrectheid van de quantumtensorcodetest van Ji et al. [FOCS 2022], op natuurlijke wijze kunnen worden uitgebreid naar het model van commuterende operatoren.

We geven tevens het eerste informatietheoretische bewijs van parallelle herhaling voor niet-lokale spellen in het model van entanglement met commuterende operatoren. Dit doen we door het bewijs van verankerde parallelle herhaling van Bavarian et al. [SIAM J. Comput. 2022] te combineren met het door Araki [PRIMS 1977] ontwikkelde raamwerk voor relatieve entropie. Daarnaast definiëren we verschillende informatietheoretische begrippen voor de setting van traciële von Neumann-algebra's, waaronder wederzijdse informatie en relatieve entropie. Deze definities kunnen ook op zichzelf van belang zijn.

Met behulp van deze vernieuwingen bewijzen we $\mathsf{MIP}^* = \mathsf{RE}$ en $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ gelijktijdig door aan te tonen dat de compressiestelling uit het oorspronkelijke bewijs van $\mathsf{MIP}^* = \mathsf{RE}$ in het model van commuterende operatoren dezelfde volledigheids- en correctheidsgaranties biedt. Dit laat zien dat zowel MIP^* als $\mathsf{MIP}^{\mathsf{co}}$ via de compressiestelling aan deze voorwaarde voldoen, waardoor we de onberekenbaarheid van MIP^* en $\mathsf{MIP}^{\mathsf{co}}$ binnen één uniform raamwerk kunnen bewijzen, ondanks het feit dat deze twee complexiteitsklassen verschillend zijn. Deze benadering levert bovendien een alternatief bewijs van de stelling $\mathsf{MIP}^* = \mathsf{RE}$ op dat niet afhankelijk is van het behoud van een bovengrens op de hoeveelheid entanglement. Met behulp van deze compressiestelling bewijzen we bovendien dat het beslissen of de waarde van een spel in het tensorproductmodel gelijk is aan zijn waarde in het model van commuterende operatoren, equivalent is aan het stopprobleem. Hieruit volgt dat er oneindig veel voorbeelden van "Bell-tests" bestaan die onderscheid maken tussen het tensorproductmodel en het model van commuterende operatoren. Dit betekent echter ook dat men niet met behulp van een computer via brute force naar een dergelijke Bell-test kan zoeken. Ten slotte geven we een gestroomlijnder bewijs van de compressiestelling voor niet-lokale spellen door gebruik te maken van het synchrone raamwerk van Mousavi et al. [STOC 2022] en van de verbeterde Pauli-basistest die werd geïntroduceerd door de la Salle [ArXiv:2204.07084].

[2] We merken op dat de aanduidingen *co* aan beide zijden van $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ naar verschillende zaken verwijzen!

Contents

III Interactive proof systems and the compression theorem

IV Implications and open problems

CHAPTER 1

Introduction

This thesis proves a simply stated theorem, $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$. Using slightly more technical jargon, this thesis shows that the complexity class defined by an interactive proof system with the commuting operator model of entanglement is equivalent to the non-halting problem. However, this weird title[1] or the long, cryptic statement doesn't give proper credit to the main result of this thesis and the story behind it. This theorem makes an interesting connection between computational complexity theory and operator algebra through quantum information, and also has various interesting implications for these communities. Before diving into this theorem and the rigorous mathematics that proves this statement, it would be interesting to first go into the history behind these connections.

1.1 Background information

The history behind the complexity class $\mathsf{MIP}^{\mathsf{co}}$ is also an interesting story in itself, as it has been considered by three different research communities independently, with very different goals in mind! Much like the proverb of "the blind men touching an elephant", different viewpoints reveal different facets about the same problem, yet none alone captures the full story. One might tell a similar story about a dragon: in Chinese culture, the dragon is not a single creature but a composite being, assembled from the features of many animals. Seen from one angle, one notices the horns; from another, the scales; from yet another, the claws. Only by bringing these perspectives together does one see the complete dragon emerge from the sky. I want to start this thesis by giving a (relatively) basic, intuitive summary of all three research communities in this section, and hopefully give you all the different perspectives required to see the dragon emerging from the night sky.

1.1.1 Modelling entanglement: the physics perspective

Our story begins with the study of quantum mechanics back in the 1920s, when the notion of entanglement, where subatomic particles can be correlated over a long distances, was first proposed. Of course, there was a lot of skepticism back then because of how counter-intuitive this sounded in comparison to classical mechanics. The debate between Bohr and Einstein on this topic marks an important page in the development of modern physics and quantum mechanics. However, these debates naturally stayed in the theoretical landscape, and a very natural question back then was whether one could come up with an experimental setup which would give us concrete evidence for particles being correlated over distances in the real world.

This question was answered by John Bell in the 1960s with the introduction of the Bell test [Bel64]. There are many formulations of the Bell test, but the most famous one given is

[1] which my office mate jokes that it sounds like a complexity theory paper title generated by ChatGPT

x	y	$x \wedge y$
0	0	0
1	0	0
0	1	0
1	1	1

a	b	$a \oplus b$
0	0	0
1	0	1
0	1	1
1	1	0

Figure 1.1: The logic table for $x \wedge y$ and $a \oplus b$.

the formulation due to Clauser, Horne, Shimony and Holt back in the 1960s [CHS+69]. In this formulation, the experiment can be described using two cooperating players, Alice and Bob, and a referee whom I named Richard[2]. In this setting, Alice and Bob are trying to win against Richard, and the experiment goes as follows: Richard will first pick two uniformly random bits $x, y \in \{0,1\}$, and send x to Alice and y to Bob. Then Alice (resp. Bob), depending on x (resp. y), will send a bit $a \in \{0,1\}$ (resp. $b \in \{0,1\}$) back to Richard. Alice and Bob win this theoretical game if

$$x \wedge y = a \oplus b,$$

where $\wedge$ and $\oplus$ are described in Figure 1.1.

In this experiment, Alice and Bob can strategize before the game begins, but cannot communicate with each other once they have received their bit from Richard. Since Alice's answer a can only depend on x and Bob's answer b can only depend on y, by looking at Figure 1.1 carefully, one can deduce that Alice and Bob can win this scenario with probability at most 0.75. However, by the laws of quantum mechanics, if Alice and Bob prepare an entangled pair between them, then they could potentially sample from a richer set of probability distributions, or "correlations", between them. Given this larger correlation set, it can be shown that Alice and Bob can win with probability $\cos^2(\frac{\pi}{8})$, which is around 0.85 (and in fact, under the models of quantum mechanics, they cannot perform better than $\cos^2(\frac{\pi}{8})$). Hence, the experiment goes as follows: The referee first asks Alice and Bob to prepare many different entangled pairs ahead of time, then asks Alice and Bob (in this case, it could be two different experimental labs thousands of kilometres apart) to repeatedly play the game. If their win rate on average hovers around 0.75, then this is evidence against quantum entanglement, and consequently, if their win rate hovers around 0.85, then this is evidence that particles indeed can be "correlated" over long distances.

The scenario described above is known as the CHSH game in the literature, and this experiment, along with subsequent work by Alain Aspect and Anton Zeilinger, forms the basis of the Nobel Prize in Physics for 2022 [Nob22]. This experiment has been implemented many times after its discovery (see, for example [AGR82]), and the fact that this thesis is funded and published provides evidence that particles can be correlated over long distances.

Given the CHSH game, a natural follow-up question is whether there are more scenarios which can distinguish between the "classical" model (or models where entanglement does not exist) and the "quantum" models of entanglement. This is where the notion of non-local games arises in the literature, where intuitively, a non-local game is a generalization of the CHSH game. To be more precise, a non-local game $\mathcal{G}$ is specified by two finite sets, a question set $\mathcal{X}$, and an answer set $\mathcal{A}$, a question distribution $\mu \sim \mathcal{X}^2$ and a verification function $V : \mathcal{X}^2 \times \mathcal{A}^2 \to \{\text{win}, \text{lose}\}$. The game plays out similarly to the CHSH game, except Richard samples the question distribution $(x, y) \sim \mu$ (instead of $(x, y) \sim \{0,1\}^2$ in the CHSH game). After receiving the answer pair $(a, b) \in \mathcal{A}^2$ (instead of the answer pair $(a, b) \in \{0,1\}^2$), the referee decides whether Alice and Bob win the given game based on the outcome of $V(x, y, a, b)$.

It turns out that there are many known examples of non-local games in which the players can achieve a strictly higher success rate when given entanglement in comparison to players

[2] In honour of my master's supervisor, Richard Cleve.

without entanglement, and one such example is the Mermin-Peres magic square game [Mer90; Per90]. Non-local games will be one of the key objects studied in this thesis.

Tsirelson's problem

In addition to finding a concrete experimental setup to verify quantum entanglement, another natural question about entanglement was considered by the physics community: What is the best mathematical framework for describing quantum entanglement? The "conventional" physics axiom[3] used to model entanglement is called the *quantum tensor product model of entanglement*. In this model, a pair of entangled systems is described by a vector state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_A$ and $\mathcal{H}_B$ are two (potentially infinite-dimensional) Hilbert spaces. Intuitively, $\mathcal{H}_A$ and $\mathcal{H}_B$ describe the space for each system, and any measurement operator made on one system is defined locally in their respective Hilbert spaces (i.e. the measurement operator on system A is defined as $A_a \otimes \mathbb{I}_B \in \mathcal{B}(\mathcal{H}_A) \otimes \mathbb{I}_B$). This model is commonly known as the *quantum tensor product* model in the literature. For a non-local game $\mathcal{G}$, the *tensor product value* of the game is defined to be the optimal success rate for players sharing the tensor product model of entanglement, and I use $\omega^*(\mathcal{G})$ to denote this quantity in this thesis. As explained in the next subsection, the tensor product is a more "analytic" way of defining entanglement.

The *commuting operator model* is another natural model used to model entanglement, which is used in, e.g., the Haag-Kastler axioms of quantum mechanics [HK64]. In this model, a pair of entangled systems is described by a vector state $|\psi\rangle \in \mathcal{H}$ defined within a single Hilbert space $\mathcal{H}$, and the only restriction is that the measurement operators associated with the two systems must pairwise commute. To be more precise, for any measurement operator $A \in \mathcal{B}(\mathcal{H})$ for the first system and $B \in \mathcal{B}(\mathcal{H})$ for the second system, it must be the case that $[A,B] = AB - BA = 0$. Similarly, the *commuting operator value* of the game is defined to be the optimal success rate for players sharing the commuting operator model of entanglement, and I use $\omega^{co}(\mathcal{G})$ to denote this quantity in this thesis. In some sense, this is a more "algebraic" way of defining entanglement.

If the measurement operator is defined in two separate tensor factors, then they automatically commute (i.e. $[A \otimes \mathbb{I}_B, \mathbb{I}_A \otimes B] = 0$). This implies that the commuting operator model is a more general notion of entanglement. In fact, these two models of entanglement are equivalent to each other mathematically when the underlying Hilbert space is finite-dimensional. Boris Tsirelson asked, in 1993, whether these two models of entanglement are equivalent [Tsi93][4]. This question was partially answered by William Slofstra in 2016, as he showed that the set of probability distributions generated by the tensor product model, or the set of *tensor product correlations*, is not closed [Slo19b], while the set of *commuting operator correlations* is closed. However, a more interesting question is whether there exists a non-local game such that the tensor product value of the game is strictly smaller than its commuting operator value. This is one formulation of the famous *Tsirelson's problem* in the literature, and a positive resolution of this problem would imply that there are **no** non-local games which can be used to distinguish these two models of entanglement (hence, roughly speaking, implying equivalence between the two models).

1.1.2 Classifying von Neumann algebra: the operator algebra perspective

The story from the mathematics side comes from the classification of von Neumann algebra. For readers without a background in von Neumann algebras, it is helpful to think of a von Neumann algebra as a generalization of an algebra of infinite-dimensional matrices. This set of

[3] There are many debates about what is considered conventional, but we will not discuss this further in this thesis since this is not the main focus.

[4] He technically did not ask this question, but rather assumed, without proof, that they are equivalent in this work. This was later pointed out to be nontrivial.

algebras contains many nice theoretical properties which make them ideal for modelling quantum mechanics, along with many other applications.

In 1976, Alain Connes showed that a von Neumann algebra can be classified by the different type of traces that they admit. Roughly speaking, a trace function τ for a von Neumann algebra is a linear functional which satisfies the cyclic property (i.e. $\tau(AB) = \tau(BA)$ for all A and B in the algebra). As the name suggests, a trace function is designed to mimic the structure of a trace for finite dimensional matrix algebras. In the seminal work by Connes [Con76], he showed that any von Neumann algebra can be decomposed as the direct product of the following 5 types of von Neumann algebras, where a von Neumann algebra is ...

1. ... I_n if it admits a trace which maps projectors to $\{0, \cdots, n\}$ for some integer n. This intuitively corresponds to a matrix algebra $\mathscr{M}_n(\mathbb{C})$ in the finite dimension setting.

2. ... I_∞ if it admits a trace which maps projectors to $\{0, \cdots, \infty\}$. This intuitively corresponds to a matrix algebra over an infinite-dimensional Hilbert space.

3. ... II_1 if it admits a trace which maps projectors to $[0,1]$ (i.e. a continuous mass between 0 and 1).

4. ... II_∞ if it admits a trace which maps projectors to $[0,\infty]$.

5. ... III if it admits a trace which maps projectors to 0 or ∞.

In the same work, he proposed a question which relates type I_n von Neumann algebra with type II_1 von Neumann algebras. To understand this question, we have to first define a *hyperfinite* II_1 *factor*. The definition of the hyperfinite II_1 factor starts with the matrix algebra $\mathscr{R}_1 = \mathscr{M}_2(\mathbb{C})$, or the algebra generated by the set of two by two matrices. $\mathscr{R}_1$ can be embedded into $\mathscr{R}_2 = \mathscr{M}_{2^2}(\mathbb{C})$ by mapping

$$A \to I_2 \otimes A = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix},$$

and the (normalized) trace can be extended accordingly. One can further embed this into $\mathscr{R}_3 = \mathscr{M}_{2^3}(\mathbb{C})$ using a similar embedding and define $\{\mathscr{R}_n\}_{n\in\mathbb{N}}$ accordingly. The hyperfinite II_1 factor $\mathscr{R}$ is defined as the weak $*$-closure (a key topology used to define a von Neumann algebra) of the infinite union of $\mathscr{R}_n$. $\mathscr{R}$ naturally admits a trace τ, since by definition each $\mathscr{R}_n$ admits a trace which extends to $\mathscr{R}_m$ for any $n \leq m$. Since each projector $P \in \mathscr{R}$ originates from one of the $\mathscr{R}_n$ (and hence $\tau(P) = \frac{l}{2^n}$ for some l between 0 and $2^n - 1$), this implies that τ maps any projector to $[0,1]$, thus showing that $\mathscr{R}$ is a type II_1 von Neumann algebra. In his seminal work, Connes asked the following famous question:

Can every II_1 factor on a separable Hilbert space be embedded into the free ultrapower of $\mathscr{R}$, the hyperfinite II_1 factor?

This is known as the *Connes embedding problem* in the literature. In the statement above, a factor is any von Neumann algebra whose centre (i.e. the intersection between the von Neumann algebra and its commutant) consists of scalar multiples of the identity[5]. A Hilbert space is separable if it contains a countable basis. For the sake of intuition, one should interpret a type II_1 factor on a separable Hilbert space as a "natural" type II_1 von Neumann algebra. An ultrapower, in essence, can be thought of as a generalization of a limit, and the free ultrapower of $\mathscr{R}$ can be intuitively interpreted as "infinite copies" of $\mathscr{R}$. Since by definition, the hyperfinite II_1 factor can be approximated by finite-dimensional matrices, or type I_n von Neumann algebras, a less scary/hand-wavy way to think about the Connes embedding problem is the following:

[5]This class of von Neumann algebras is actually the class for which Connes proved the famous classification theorem mentioned earlier in this subsection.

Can every type II_1 von Neumann algebra be embedded into another algebra defined by the limits of finite-dimensional matrices?

This problem is a famous open problem in operator algebra, and it has also been shown to be related to different problems in non-commutative probability and non-commutative real algebraic geometry. However, the most interesting connection[6] is to the different tensor product models of C^*-algebras, which I summarize below.

Roughly speaking, a C^*-algebra is an abstract algebra equipped with a norm $\|\cdot\|$ and an involution operation $*$ that satisfies $\|a^* a\| = \|a\|^2$ for all a in the algebra (in this case, one should intuitively think of the involution operation as an abstraction of the complex conjugation for finite-dimensional matrices). A C^*-algebra is *concrete* if it is represented as a subalgebra of some $\mathscr{B}(\mathscr{H})$. The well-known GNS construction [Bla06, p. II.6.4] states that every C^*-algebra can be represented on some $\mathscr{B}(\mathscr{H})$.

Given two C^*-algebras $\mathscr{A}$ and $\mathscr{B}$, one can define the tensor product between $\mathscr{A}$ and $\mathscr{B}$ in one of two ways. The first way, known as the *minimal tensor product* in the literature (labelled as $\otimes_{\min}$), is by first mapping $\mathscr{A}$ and $\mathscr{B}$ into the concrete representations which have underlying Hilbert spaces $\mathscr{B}(\mathscr{H}_A)$ and $\mathscr{B}(\mathscr{H}_B)$ respectively, using the GNS construction. Then one can take advantage of the underlying concrete representation $\mathscr{B}(\mathscr{H}_A) \otimes \mathscr{B}(\mathscr{H}_B)$ in order to define $\mathscr{A} \otimes_{\min} \mathscr{B}$. This is the more analytic approach, and the resulting norm is the smallest possible norm which can arise from any tensor product construction.

The second way, known as the *maximal tensor product* (labelled as $\otimes_{\max}$), instead considers the algebra $\mathscr{A} \odot \mathscr{B}$ defined as follows: $\mathscr{A} \odot \mathscr{B}$ is defined in terms of linear combinations of elements of the form $a \odot b$ for $a \in \mathscr{A}$ and $b \in \mathscr{B}$. Furthermore, the algebra satisfies the following rules:

- $(a_1 \odot b_1) \cdot (a_2 \odot b_2) = (a_1 \cdot a_2) \odot (b_1 \cdot b_2)$
- $(a \odot b)^* = (a^* \odot b^*)$

Under these multiplication rules, $\mathscr{A} \odot \mathscr{B}$ forms an $*$-algebra, and in order to make it a C^*-algebra, one takes the maximal norm $\|\cdot\|_{\max}$ to be

$$\|(a \odot b)\|_{\max} = \sup_{\pi} \|\pi((a \odot b))\|$$

where the supremum is taken over all concrete representations of $\mathscr{A} \odot \mathscr{B}$. The resulting norm constructed above is the largest possible norm arising from any tensor product construction.

An interesting question is: given a C^*-algebra $\mathscr{A}$, in what scenario is the algebra $\mathscr{A} \otimes_{\max} \mathscr{A}$ equivalent to $\mathscr{A} \otimes_{\min} \mathscr{A}$. A C^*-algebra is called *nuclear* iff the norm defined by $\mathscr{A} \otimes_{\max} \mathscr{A}$ coincides with $\mathscr{A} \otimes_{\min} \mathscr{A}$ [Bla06, Definition II.9.4.1]. $\mathscr{A}$ is known to be nuclear if $\mathscr{A}$ is abelian (i.e. the multiplicative structure is commutative) or represented within a finite-dimensional Hilbert space. The interesting scenario is when $\mathscr{A}$ is taken to be $C^*(\mathbb{F}_2)$ or the group C^*-algebra over the free group over two generators. In this case, the free group $\mathbb{F}_2$ is generated by 2 generators and their inverses, or $\{s_1, s_1^{-1}, s_2, s_2^{-1}\}$ where $s_i s_i^{-1} = I$ for all $i \in \{1,2\}$ (with no commutation relationships defined). One can transform $\mathbb{F}_2$ into a C^*-algebra by considering the unitary representation over some $\mathscr{B}(\mathscr{H})$. To this end, we have the following question:

Is the C^*-algebra $C^*(\mathbb{F}_2)$ nuclear?

This is one of the formulations of the well-known Kirchberg's problem in the literature [Kir93]. This problem is shown to be equivalent to the Connes embedding problem by Eberhard Kirchberg in the same paper. To be more precise, if $C^*(\mathbb{F}_\infty)$ is nuclear, then every II_1 factor on a separable Hilbert space can be embedded into the free ultrapower of R.

[6] with some bias.

It turns out, Kirchberg's problem is also quite relevant to quantum information. Let $\mathbb{F}_m^n$ denote the free group with m generators and order n (i.e. $s_i^n = I$ for all $i \in \{1,\dots,m\}$). Kirchberg's problem can be formulated as the following problem:

$$\text{Is the } C^*\text{-algebra } C^*(\mathbb{F}_m^n) \text{ nuclear for all } n, m \in \mathbb{N}?$$

To relate the problem back to quantum information, consider n sets of PVMs each with m outcomes, i.e. $\{P_a^x\}_{a\in\{1,\dots,n\}}$ for POVMs indexed by $x \in \{1,\dots,m\}$. By considering the unitary observable

$$U^x = \sum_{k=1}^{n} e^{\frac{2\pi i}{n}} P_k^x,$$

for all $x \in \{1,\cdots,m\}$, U^x is precisely a representation for $C^*(\mathbb{F}_m^n)$. In this case, the minimal tensor product between two copies of $C^*(\mathbb{F}_m^n)$ corresponds precisely to the tensor product model of entanglement between two quantum systems with n different measurements, each with m measurement outcomes. Similarly, the maximal tensor product between two copies of $C^*(\mathbb{F}_m^n)$ corresponds to the commuting operator model of entanglement. Hence, if Tsirelson's problem, introduced in the previous section, has a positive answer, this immediately implies that both Kirchberg's problem and Connes embedding problem are true!

Since the focus of this thesis is not on von Neumann algebra theory, I refer the reader to the excellent survey paper by Isaac Goldbring [Gol21] and an excellent paper by Tobias Fritz [Fri12] for more details on the implications of Connes embedding problem and how Tsirelson's problem is related to it.

1.1.3 Interactive proof system: the computer science perspective

Finally, I will bring the story back to the theoretical computer science community, with one of the (still open) Millennium Prize Problems, P vs NP. P and NP are perhaps the most famous complexity classes in the literature, where P is the set of decision problems which are solvable by a deterministic polynomial-time Turing machine, and NP is the set of decision problems solvable by a nondeterministic polynomial-time Turing machine. There are many definitions for these two complexity classes, but the definition that I use in this introduction will be in the languages of proofs. There are many notions of proofs for mathematical statements[7].

In complexity theory, a language L is a subset of $\{0,1\}^*$ and, from the proof perspective, corresponds to the set of "correct" statements (or the binary encoding of correct statements). From this point of view, a decision problem in complexity theory can be phrased as follows: given a statement $x \in \{0,1\}^*$, decide whether x is "correct", i.e. whether $x \in \mathsf{L}$ or $x \notin \mathsf{L}$. In this case, P corresponds to the set of languages in which one can deduce whether the statement is correct given a polynomial-time algorithm, and NP corresponds to the set of languages for which there exists a polynomial-time algorithm that can verify a statement when given a proof string y. Clearly, if one can deduce the statement correctly, one can also verify (even without using the proof) the statement correctly. So fundamentally, P vs NP is about whether all problems which can be verified efficiently can also be proven efficiently.

It turns out, proving whether P = NP is true is very hard. So people began to define a more "relaxed" notion of NP in hopes of making progress on this problem. One such simplification is the *interactive proof system* model, introduced by Shafi Goldwasser, Silvio Micali and Charles Rackoff in 1985 [GMR85]. Roughly speaking, this model is defined in terms of two parties, a verifier known as Arthur and a prover known as Merlin. Given a language $\mathsf{L} \subseteq \{0,1\}^*$, Arthur, when given an instance $x \in \{0,1\}^*$, wants to decide whether the instance corresponds to a correct statement using a polynomial-time algorithm. However, Arthur can also query Merlin probabilistically, and through multiple rounds of interaction (i.e. Arthur can make the next query

[7] As some statements which are 10 characters long would require a 50 page paper followed by a 167 page paper, along with a few more previous results in order to show that the statement is correct.

based on the answer he gets in the previous round). Merlin is assumed to be computationally unbounded, but he is always trying to convince Arthur that $x \in \mathsf{L}$ (even in the case where $x \notin \mathsf{L}$). This becomes an interesting scenario, since Arthur needs to find an intelligent way to query Merlin to make sure that Merlin is not deceiving him. A language L is in the complexity class IP if there exists an interaction between Arthur and Merlin such that

- (Completeness) If $x \in \mathsf{L}$, then Arthur always "accepts" the given instance (i.e. Arthur concludes that $x \in \mathsf{L}$).
- (Soundness) If $x \notin \mathsf{L}$, then no matter how Merlin responds during the interaction, Arthur will accept the given instance with probability at most 0.5.

I just wanted to quickly remark that this is not the standard completeness/soundness gap used in the literature describing IP[8]. Intuitively, one should interpret P, NP and IP in the following manner:

- P intuitively corresponds to you, as the reader, trying to prove a statement is correct without any assistance.
- NP intuitively corresponds to you, as the reader, trying to verify the statement when given a paper that claims to prove the statement.
- IP intuitively corresponds to you, as the reader, trying to verify the statement in a refereeing process for a paper. You can query the author freely, but the author of the paper is determined to convince you that the paper is correct (even in the case where the paper might contain an error).

Given the above intuition, it is unsurprising that IP is known to be more powerful than NP. A seminal result due to Adi Shamir [Sha90] shows that IP is equal to PSPACE, where, recall, PSPACE corresponds to the set of problems which are solvable by a polynomial space algorithm.

In a follow-up work, László Babai, Lance Fortnow and Carsten Lund introduce the notion of a *multiprover interactive proof system* [BFL91]. In this model, the verifier can additionally interact with multiple non-interactive provers, or provers that cannot interact with each other during the verification procedure. The provers can coordinate their answers beforehand, but they do not know the question given to the other provers during the verification process. At first glance, it might seem that the verifier cannot gain any additional advantage by interacting with multiple provers. However, the additional provers give the verifier the power of "cross-referencing", meaning that the verifier can ask two different provers the **same** question, and expect the same answer back from both provers. In this case, the provers will look extremely suspicious if they gave different answers to the same question. In the same paper, Babai et al. show that MIP, the set of languages decidable by a multiplayer interactive proof system, is NEXP-complete, where NEXP corresponds to NP but with exponential validation time and proof size. Interestingly, it can be shown that the verifier only needs to make one round of communication simultaneously with two provers in order to verify a NEXP instance. The proof of $\mathsf{MIP} = \mathsf{NEXP}$ has led to many foundational results in modern computer science. This includes some of the early development of the probabilistic checkable proofs (PCP) [AS98; ALM+98], and the unique games conjecture [Kho02] for the hardness of approximation of NP-hard problems.

Moving forward to the 2000s, a natural question is whether there is a quantum analogue of IP or MIP. The first attempt was given by John Watrous by defining the notion of a quantum interactive proof system (QIP), where the verifier is now given access to a polynomial-time quantum computer and allowed to share a quantum proof between the provers [Wat99]. However, it turns out this did not give a computational advantage to the verifier, as John Watrous, along

[8] The standard definition usually requires a (2/3, 1/3) completeness/soundness gap

with Rahul Jain, Zhengfeng Ji and Sarvagya Upadhyay showed that QIP = PSPACE [JJU+11]. Similarly, Hirotada Kobayashi and Keiji Matsumoto gave a notion of a quantum multiprover interactive proof system (QMIP), and also showed that this did not offer any advantage in comparison to classical MIP [KM02].

The more interesting scenario, and the one which we focus on in this thesis, is the case where the verifier is still classical, and is restricted to classical messages. The provers still can't communicate to each other, but they are allowed, in addition to coordinating answers, to prepare a shared entangled state between them. For players sharing the tensor product of entanglement, this is known as a *multiplayer interactive proof system with the tensor product model of entanglement* or MIP^* in the literature. If we restrict MIP^* to the scenario where the verifier is only allowed to make one round communication with two (entangled) provers, we see that this protocol, by definition, is extremely similar to the notion of a non-local game introduced back in Section 1.1.1, as the players (Alice and Bob) in a non-local game correspond to the provers in the interactive proof system model, and similarly, the referee (Richard) corresponds to the verifier. This connection was noticed by Richard Cleve, Peter Høyer, Ben Toner and John Watrous back in 2004 [CHT+04]. Due to this connection, we can describe the problems in MIP^* by the $(1, \frac{1}{2})$-tensor product value problem, which is a promise problem given as follows: given a non-local game $\mathcal{G}$, decide whether

- (Yes case): $\omega^*(\mathcal{G}) = 1$,
- (No case): $\omega^*(\mathcal{G}) \leq \frac{1}{2}$.

We remark that although a two-prover one-round MIP seems similar to a non-local game, there are some minor differences between these two complexity classes on how the problem is represented. A non-local game is represented by a tuple $(\mathcal{X}, \mathcal{A}, \mu, V)$, while an interactive proof system only requires a description of two Turing machines which implement μ and V respectively. Intuitively, μ, in the interactive proof system setting, is the function which implements the sampling procedure for the verifier to sample a pair of questions to the provers, and V corresponds to the decision function of whether the verifier will accept the given instance or not. We remark that although there are no question and answer sets in the interactive proof model, since both μ and V are assumed to run in polynomial-time, and hence only be able to read a polynomial number of bits as input, the question and answer size is effectively strings of size $\mathrm{poly}(n)$.

At the time, it was not immediately clear what the complexity of MIP^* should be, or even whether it was more powerful than NEXP at all! On the one hand, giving the provers entanglement could potentially let them "collude" more effectively in some instances, which weakens the power of the verifier. However, this also allows the verifier to create a new and inventive protocol which takes advantage of the entanglement they share, potentially giving the verifier more power.

The only trivial upper bound for MIP^* that we knew at the time was RE, the complexity class which is complete with respect to the halting problem. More precisely, a decision problem $\mathsf{L} \subseteq \{0,1\}^*$ is in RE if there exists an algorithm A such that

- if $x \in \mathsf{L}$, then A, on input x, halts and returns 1.
- Otherwise A, on input x, might never halt.

It is not hard to see that if one can solve the halting problem, one can decide L. We know that $\mathsf{MIP}^* \subseteq \mathsf{RE}$ due to the well-known search-from-above algorithm, a brute-search based algorithm which allows one to compute a sequence x_i such that $x_1 \leq x_2 \leq \cdots \leq \omega^*(\mathcal{G})$ and $\lim_{n\to\infty} x_n = \omega^*(\mathcal{G})$ for all games $\mathcal{G}$. Intuitively, each x_i corresponds to the best-performing strategy for the prover for $\mathcal{G}$ when restricted to finite-dimensional strategies defined within

$\mathbb{C}^i \otimes \mathbb{C}^i$ (up to an error of $\frac{1}{i}$). Since both $\mathscr{B}(\mathscr{H}_A)$ and $\mathscr{B}(\mathscr{H}_B)$ can be approximated by finite-dimensional matrices, the limit of this sequence converges to the tensor-product value of the game $\mathscr{G}$. Hence, if $\omega^*(\mathscr{G}) = 1$, then one can compute this sequence until $x_i > \frac{1}{2}$, thus correctly deciding the yes case for the $(1, \frac{1}{2})$-tensor product value problem. However, if $\omega^*(\mathscr{G}) \leq \frac{1}{2}$, then computing this sequence might never give a certificate which verifies $\omega^*(\mathscr{G}) \leq \frac{1}{2}$.

In 2008, Miguel Navascués, Stefano Pironio and Antonio Acín developed another algorithm which gives a computable sequence $y_1 \geq y_2 \geq \cdots \geq \omega^*(\mathscr{G})$, but $\lim_{n\to\infty} y_n = \omega^{co}(\mathscr{G})$ [NPA08]. This sequence, known as the NPA hierarchy in the literature, uses the commuting assumption from the commuting operator model of entanglement to define a sequence of semidefinite programs (SDPs) where the optimal value of the i-th SDP is precisely y_i. Assuming that the aforementioned Tsirelson's problem is true (i.e. $\omega^*(\mathscr{G}) = \omega^{co}(\mathscr{G})$), one can continuously compute $|y_i - x_i|$ until the difference is less than $\frac{1}{2}$ in order to estimate $\omega^*(\mathscr{G})$ with error bound less than $\frac{1}{2}$, thus giving a halting algorithm for the $(1, \frac{1}{2})$-tensor product value problem. In other words, if there is no algorithm which solves the $(1, \frac{1}{2})$-tensor product value problem (or $\mathsf{MIP}^* = \mathsf{RE}$), then Tsirelson's problem, along with Kirchberg's problem and Connes embedding problem, would admit a negative answer!

MIP^* was shown to be at least be as powerful as MIP in [IV12] by Thomas Vidick and Tsuyoshi Ito in 2012, where the authors showed that $\mathsf{NEXP} \subseteq \mathsf{MIP}^*$. However, it was unclear at the time whether the two complexity classes were equal. Numerous papers were published in this area, attempting to explore the complexity of MIP^* (see, for example [NV18], [NV17], [Gri20]). It was not until 2019 that Anand Natarajan and John Wright showed that NEEXP, the analogue of NP with doubly exponential proof size and running time, is contained in MIP^* [NW19], thus showing that MIP^* is strictly more powerful than MIP. Around the same time, Joseph Fitzsimons, Zhengfeng Ji, Thomas Vidick and Henry Yuen showed that for any $n \in \mathbb{N}$, the $(1, 1 - O(\exp(-C \cdot \exp(\cdots \exp(n)))))$-tensor product value problem (where the expression $1 - O(\exp(-C \cdot \exp(\cdots \exp(n))))$ contains $n-1$ nested exponentials) is complete for the n-ary $\mathsf{NE}\cdots\mathsf{EXP}$ complexity class (with n "E " behind the N) [FJV+19], thus showing that the $(1, \frac{1}{2})$-tensor product value problem could potentially be uncomputable.

Finally, in 2020, Anand Natarajan, Zhengfeng Ji, Thomas Vidick, John Wright and Henry Yuen, in a landmark result [JNV+22a], showed that $\mathsf{MIP}^* = \mathsf{RE}$ by combining the one-layer compression from [NW19] with the recursive compression technique from [FJV+19][9], thus giving a negative resolution to Tsirelson's problem, Kirchberg's problem and Connes embedding problem all at once.

Multiplayer interactive proof system with the commuting operator model of entanglement

From this story, there is an obvious missing complexity class, *multiplayer interactive proof system with the commuting operator model of entanglement*, or the complexity class $\mathsf{MIP}^{\mathrm{co}}$. There are many similarities between the complexity class MIP^* and $\mathsf{MIP}^{\mathrm{co}}$. For one, determining the complexity of both these classes started out as a master's project [Vid20; Lin22][10]. Another similarity is that the problems in $\mathsf{MIP}^{\mathrm{co}}$ are described by a similar problem known as the $(1, \frac{1}{2})$-commuting operator value problem, which is a promise problem given as follows: given a non-local game $\mathscr{G}$, decide whether

- (Yes case): $\omega^{co}(\mathscr{G}) = 1$,
- (No case): $\omega^{co}(\mathscr{G}) \leq \frac{1}{2}$.

From the NPA hierarchy described earlier, we know that $\mathsf{MIP}^{\mathrm{co}} \subseteq \mathsf{coRE}$, where, a decision problem $\mathsf{L} \subseteq \{0,1\}^*$ is in coRE if there exists an algorithm A such that

[9] Technically, this result was resolved in [JNV+22b] by fixing a previous bug regarding the quantum soundness of the low-degree test.

[10] Jokes aside, I think proving that $\mathsf{MIP}^* = \mathsf{RE}$ is a much more impressive masters project than my thesis, as a lot of the techniques in this thesis are based on techniques used to prove $\mathsf{MIP}^* = \mathsf{RE}$ result.

- if $x \notin \mathsf{L}$, then A, on input x halts and returns 0.

- Otherwise, A, on input x, might never halt.

It is well known that $\mathsf{RE} \neq \mathsf{coRE}$ and $\mathsf{RE} \cap \mathsf{coRE}$ contains the set of "computable" problems, i.e. the set of problems for which there exists a halting algorithm that decides them.

One of the main challenges in determining the complexity of $\mathsf{MIP}^{\mathrm{co}}$ is that the commuting operator model is infinite-dimensional in nature, and hence it is unclear whether similar properties in the tensor product setting would hold in this model. We remark that giving the provers more correlation does not necessarily give the verifier more power. By a previous result, if we give the provers a more general correlation model known as the non-signalling model, the computational power of the verifier actually decreases to EXP, which is even weaker than the classical model! In fact, if we allow the provers to communicate, the verifier naturally loses the power to cross-reference the different provers and hence the computation power actually drops down to PSPACE!

1.2 Main contribution of this doctoral thesis

Let's repeat the sentence from the beginning of this chapter once again! This thesis proves $\mathsf{MIP}^{\mathrm{co}} = \mathsf{coRE}$, or the complexity class defined by an interactive proof system with the commuting operator model of entanglement is equivalent to the non-halting problem. Along the way, I introduced many tools for both the quantum information and the computational complexity communities, which could be useful for other problems. I will start by first summarizing some of the main contributions of this doctoral thesis in this section.

1.2.1 Tracially embeddable strategies

One of the main contributions in this thesis related to quantum information is introducing the notion of tracially embeddable strategies into the literature. To be a bit more precise, for a non-local game $\mathcal{G}$, we use the term *strategy* to denote the entangled state and the measurement operators for the players/provers used to play $\mathcal{G}$. For a non-local game $\mathcal{G} = (\mu, D, \mathcal{X}, \mathcal{A})$, a tensor product strategy consists of a state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ on two (potentially infinite-dimensional) Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, and two sets of measurements $\{A_a^x\}_{x\in\mathcal{X}} \subseteq \mathcal{B}(\mathcal{H}_A)$ and $\{B_a^x\}_{x\in\mathcal{X}} \subseteq \mathcal{B}(\mathcal{H}_B)$ each with outcomes in $a \in \mathcal{A}$. Intuitively, A_a^x is the measurement performed by Alice and B_a^x is the measurement performed by Bob. Similarly, a commuting operator strategy consists of a state $|\psi\rangle \in \mathcal{H}$, and the measurement operators are defined by $\{A_a^x\}, \{B_a^x\} \subseteq \mathcal{B}(\mathcal{H})$ such that A_a^x and B_a^x pairwise commute.

Mathematically, a commuting operator strategy $(\mathcal{H}, |\psi\rangle, \{A_a^x\}, \{B_b^y\})$ is tracially embeddable if

- The underlying Hilbert space $\mathcal{H}$ arises from the standard form for some tracial von Neumann algebra $(\mathcal{A}, \tau)$, where the standard form for a tracial von Neumann algebra is when the algebra is represented under the GNS representation of the tracial state τ.

- The measurement state $|\psi\rangle$ is of the form $\sigma|\tau\rangle$, where $|\tau\rangle$ is the vector state representation of the tracial function τ,

- $\{A_a^x\} \subseteq \mathcal{A}$. In other words, the observable of one of the players/provers is defined within $\mathcal{A}$ under this representation.

- $\{B_a^x\} \subseteq \mathcal{A}'$. In other words, the observable of the other player/prover is defined within the commutant of $\mathcal{A}$ under this representation.

Intuitively, one should think of the standard form as the “vectorization" notation used in quantum information, and because of this, one can find many analogues of properties and techniques from finite-dimensional strategies when working with this class of strategies. I refer the readers to Example 4.1.2 for a more concrete example which illustrates the similarity between tracially embeddable strategies and finite-dimensional strategies.

One of the key theorems that I show in this thesis is that any behaviour by the provers/players in the commuting operator model of entanglement can be **well approximated** by provers/players restricted to using tracially embeddable strategies. To be a bit more precise, the set of *correlations*, or the set of probability distributions for outputting a certain answer pair given a question pair when the provers are given access to tracially embeddable strategies is dense within the set of correlations for provers with commuting operator strategies. One of the main ingredients needed for this result is Haagerup's L_p reduction [HJX10], which allows one to approximate any von Neumann algebra with a series of tracial von Neumann algebras.

Due to the structure of the tracially embeddable strategy, and the approximation theorem above, many known results in the finite-dimensional tensor product setting can be easily translated to the commuting operator model setting by a change of notation. Some examples in this thesis include the rigidity theorem for the Pauli Basis test given in [JNV+22a, Chapter 7], the rounding theorem which allows one to approximate any approximately synchronous correlations by exact synchronous correlations [Vid22; Pau22], as well as several soundness property results used in the proof of $\mathsf{MIP}^* = \mathsf{RE}$. In addition to the theorems listed above, I believe this class of strategies could be useful in translating other theorems about entanglement from the tensor product model to the commuting operator model.

1.2.2 The compression paradigm

Inspired by the proposed proof of $\mathsf{MIP}^{\mathrm{co}} = \mathsf{coRE}$ given in [MNY22], one of the other main contributions of this thesis is to introduce a new condition for a decision problem to be RE or coRE-complete, which I named the compressible condition. This condition was considered independently in an unpublished work[11] by Andrew Marks, Seyed Sajjad Nezhadi, and Henry Yuen [NMY25], and as far as I am aware, this condition is completely original in the literature.

Roughly speaking, a language $\mathsf{L} \subseteq \{0,1\}^*$ is compressible if there exists a computable map $\mathtt{Compress}^{\mathsf{L}}$ which takes as input a description of a function $\mathtt{Seq} : \mathbb{N} \to \{0,1\}^*$ which runs in $O(\mathrm{poly}(n))$ time (with respect to the input size) and outputs a description of a “compressed" version of the function $\mathtt{Seq}^{\mathrm{Comp}} : \mathbb{N} \to \{0,1\}^*$ which satisfies the following:

- (Runtime) $\mathtt{Seq}^{\mathrm{Comp}}$ runs in $O(\mathrm{polylog}(n))$ time.
- (Completeness) For every $n \in \mathbb{N}$, $\mathtt{Seq}(n) \in \mathsf{L} \implies \mathtt{Seq}^{\mathrm{Comp}}(n) \in \mathsf{L}$.
- (Soundness) For every $n \in \mathbb{N}$, $\mathtt{Seq}(n) \notin \mathsf{L} \implies \mathtt{Seq}^{\mathrm{Comp}}(n) \notin \mathsf{L}$.

It's worth pointing out that $\mathtt{Compress}^{\mathsf{L}}$ is language dependent, and should work for **all** maps $\mathtt{Seq}$. Furthermore, $\mathtt{Compress}^{\mathsf{L}}$ only gets the description of $\mathtt{Seq}$ as an input. Since $\mathtt{Seq}^{\mathrm{Comp}}$ is restricted to run in only $O(\mathrm{polylog}(n))$ time, the output $\mathtt{Seq}^{\mathrm{Comp}}$ actually cannot depend on running $\mathtt{Seq}$, and must rely on the properties of the language itself. Intuitively, one should think of the compression condition as a compression of the complexity of generating the problem instance itself. By [NMY25], it is known that the halting problem is compressible (Example 2.2.2), and by using a similar trick using Kleene's recursion theorem similarly to the original $\mathsf{MIP}^* = \mathsf{RE}$ proof [JNV+22a, Chapter 12], I show that if $\mathsf{L} \in \mathsf{RE}$ (i.e. L is decidable if one can solve the halting problem), then L being compressible is equivalent to L being RE-complete. If L is compressible, the language $\{0,1\}^* \setminus \mathsf{L}$ is also compressible, and hence the same condition can also be used to show coRE-completeness for languages and decision problems.

[11] unpublished at the time of writing this thesis.

Although we know that the halting problem is compressible, there are very few RE or coRE problems that we currently know to admit a "natural" compression map (i.e. a compression map that does not rely on the reduction back to the halting/non-halting problem). In practice, to use the compression paradigm, one often has to work with a succinct representation of a problem, and has to show that the problem is NP-hard in a specific manner so that one can actually simulate the succinct instance by a "smaller" instance of the same problem (in the $\mathsf{MIP}^*/\mathsf{MIP}^{co}$ case, self-testing of quantum states in combination with the PCP theorem gives the desired property). I personally believe that these techniques offer a new approach for proving uncomputability, and it is a very interesting open problem to see if one can use these techniques for problems outside of quantum information theory.

Compression and $\mathsf{MIP}^*/\mathsf{MIP}^{co}$. By leveraging this condition, I give a proof for both $\mathsf{MIP}^* = \mathsf{RE}$ and $\mathsf{MIP}^{co} = \mathsf{coRE}$ by showing the existence of a compression map for both the tensor product/commuting operator value problems. To be a bit more precise, a uniform problem instance for the tensor product/commuting operator value problem is defined as a Turing machine $\mathscr{V} : \mathbb{N} \to \mathscr{G}$, where to abuse notation, $\mathscr{G}$ in this case is the set of all possible descriptions for a non-local game. One could intuitively think of the uniform sequence as the "input Turing machine `Seq`" for the above example.

I show that a specific subclass of game sequences, known as the "conditionally linear verifier", admits a gap compression theorem described earlier in the introduction. The definition of the conditionally linear verifier is very much tailored to the proof of the compression theorem, and hence I will not go into detail here. Nevertheless, I give an informal version of the gap compression theorem below.

Theorem 1.2.1 (Gap compression theorem, informal version of)**.** *For every $\alpha \in \mathbb{N}$, there exists a polynomial time algorithm* `Gapcompress`*, that takes as input $\mathscr{V} : \mathbb{N} \to \mathscr{G}$, a conditionally linear verifier such that $\mathscr{V}(n)$ runs in $O(\mathrm{poly}(n))$ time, each game in the sequence can be sampled and decided in $O(\mathrm{poly}(n))$ time. The algorithm runs in $\mathrm{poly}(|\langle$`Gapcompress`$\rangle|, \alpha)$ time and outputs a conditionally linear verifier $\mathscr{V}^{Comp} : n \to \mathscr{G}$ such that, for $t \in \{*, co\}$:*

1. *(Runtime) $\mathscr{V}^{Comp}(n)$ runs in $O(\mathrm{polylog}(n))$ time, and each game $\mathscr{V}^{Comp}(n)$ can be sampled in $O(\mathrm{polylog}(n))$ time, and the decider function D_n runs in $O(\mathrm{polylog}(n))$ time.*
2. *(Completeness) If $\omega^t(\mathscr{V}(n)) = 1$, then $\omega^t(\mathscr{V}^{Comp}(n)) = 1$*
3. *(Soundness) If $\omega^t(\mathscr{V}(n)) \leq \frac{1}{2}$, then $\omega^t(\mathscr{V}^{Comp}(n)) \leq \frac{1}{2}$.*

Roughly speaking, by considering the `Gapcompress` guaranteed by the theorem, this shows that $\mathsf{MIP}^*/\mathsf{MIP}^{co}$, when restricted to games generated by a conditionally linear verifier, are compressible. Interestingly, even though the compression map for both MIP^* and MIP^{co} is the same, MIP^* is RE-complete due to the search-from-below algorithm putting $\mathsf{MIP}^* \subseteq \mathsf{RE}$, and MIP^{co} is coRE-complete due to the NPA Hierarchy [NPA08]. The proof of Theorem 1.2.1 is almost identical to the compression map given in the original $\mathsf{MIP}^* = \mathsf{RE}$ proof. The only significant difference is that I also show the completeness/soundness condition for the commuting operator model by using tracially embeddable strategies described in the last subsection. In comparison to the original theorem, using the compression framework actually gives a simpler proof for the $\mathsf{MIP}^* = \mathsf{RE}$ theorem, as one no longer has to keep track of the "entanglement dimension" property within the soundness clause in the compression theorem used in the original $\mathsf{MIP}^* = \mathsf{RE}$ proof [JNV+22a, Theorem 12.1].

Explicit separation between the models of entanglement. As a corollary of the $\mathsf{MIP}^* = \mathsf{RE}$ theorem, the tensor product model of entanglement is mathematically different from the commuting operator model of entanglement. A natural follow-up question is whether one can find

an experimental setup similar to the Bell test scenario to determine the correct way to model entanglement in our universe? [JNV+22a, Theorem 12.10] shows the existence of a game $\mathcal{G}$ such that $\omega^*(\mathcal{G}) \leq \frac{1}{2}$ and $\omega^{co}(\mathcal{G}) = 1$, which, in theory, could serve as the experiment to test which model better describes our universe. However, the question and answer sets for the given game have a magnitude of 10^{20}, making it impractical for experimental implementation.

I show that given a non-local game $\mathcal{G}$, the promise problem of deciding whether $\omega^*(\mathcal{G}) = \omega^{co}(\mathcal{G})$ or $|\omega^*(\mathcal{G}) - \omega^{co}(\mathcal{G})| \geq c$ for any fixed constant $c \in [0,1]$ is RE-complete. In other words, there is no algorithm which can be used to find an explicit separation between the tensor product model and the commuting operator model for general non-local games! The proof follows from a simple observation that Theorem 1.2.1 also gives a compression map for the above decision problem. Using the compression framework naturally gives a reduction for any non-halting Turing machine to an instance of a game such that $|\omega^*(\mathcal{G}) - \omega^{co}(\mathcal{G})| \geq c$, i.e. an experimental setup which separates the tensor product model of entanglement from the commuting operator model of entanglement; however, since the technique used is similar to the one given in [JNV+22a, Theorem 12.10], I suspect that the question size and answer size would be as large as those in [JNV+22a].

1.2.3 Relative entropy and Mutual information in the commuting operator model.

Tracially embeddable strategies are sufficient to generalize most theorems used in the proof of MIP* = RE to the commuting operator model. However, one spot that it completely fails is the anchored parallel repetition theorem from [BVY21]. One of the main gaps for proving a quantum parallel repetition theorem for non-local games for the commuting operator model is the lack of information-theoretic tools. These techniques are not known to generalize to the commuting operator setting. Building up from the works by [Ara77], I gave an analogue for a few information-theoretic tools, such as mutual information and Uhlmann's theorem, for states on a tracial von Neumann algebra. In addition to the anchored parallel repetition theorem, I believe that these tools can be used to prove other variants of quantum parallel repetition theorem for non-local games, and could potentially be independently interesting to the quantum information community.

I wanted to remark that these tools on quantum information theory were done as a part of my master's work back in Waterloo, along with Henry Yuen, William Slofstra and Hamoon Mousavi. I also want to remark that quantum information on von Neumann algebras has been considered independently in other literature (see, for example, the Ph.D thesis for Lauritz van Luijk [Lui25]).

1.3 Organization of this thesis

One of the main challenges that I had when reading the original MIP* = RE proof is the overwhelming background required to understand the paper. Although the same problem is presented in this thesis, the goal is to minimize some of the dependency on background and to give a clear picture about each individual piece for the MIP* = RE/MIP^{co} = coRE proof. With this goal in mind, I divided this thesis into four parts. I will describe the purpose of each of the four parts in this section, and some of the background and content each part covers in this thesis for the remainder of this section. Since the background notation is divided, I made an index of all nomenclature which should hopefully make this thesis more accessible.

1.3.1 Part 1: The compression condition for uncomputable problems.

In this part, the main goal is to discuss the compressible framework described in the last section. This part only consists of one chapter (Chapter 2), and it consists of my main contribution to complexity as a whole. I choose to put this in a separate section because understanding

this part only requires basic complexity theory, which makes it easy to write a self-contained section about it. This part introduces the compressible condition (Definition 2.2.1 and Definition 2.2.3) along with a self-contained proof showing that if a language is in RE, then the language being compressible also implies that the language is RE-complete (Theorem 2.2.5). Along the way, I also give a formulation for the compressible condition on succinct representation of languages Definition 2.3.2, a formulation which, in my opinion, is a lot more usable for showing RE or coRE-completeness for different decision problems.

1.3.2 Part 2: Tracially embeddable strategies for non-local games

In this part, the main goal is to discuss tracially embeddable strategies, and some general implications for the commuting operator model of entanglement. In a similar vein to Part 1 being focused on complexity theory, the theme of this part is on mathematical physics, and discusses some results about non-local games which are more oriented toward mathematical physics. In Chapter 3, I give some basic definitions and results on basic von Neumann algebra theory and quantum mechanics which I use for Part II and Part III. Essentially, this chapter contains the "basic math background" needed to understand the rest of the thesis. On the operator algebra side, I give a basic introduction to tracial von Neumann algebras and also give a finite-dimensional example for a tracial von Neumann algebra in standard form. On the quantum information side, I will give a formal definition of non-local games, as well as the different models of quantum entanglement in this chapter.

In Chapter 4, I define tracially embeddable strategies more formally Definition 4.1.1, and prove that the closure of the set of correlations generated by the tracially embeddable strategies forms the set of commuting operator correlations Theorem 4.1.3. Since the proof of Theorem 4.1.3 uses several advanced results in von Neumann algebra theory, we give a separate preliminary on these tools in this chapter. Based on the structure of the tracially embeddable strategies, I also define the commuting operator analogue for two classes of strategies, the symmetric and oracularizable strategies in this chapter.

Since tracially embeddable strategies are used quite extensively in this thesis, for readers without any background in von Neumann algebra theory, we give a finite-dimensional example in Example 4.1.2 for intuition. Alternatively, I also give a translation chart for notations from tracially embeddable strategies to finite-dimensional strategies in Table 4.1. I want to remark that, by using this translation chart in this thesis, one could obtain a proof based on finite-dimensional strategies for showing $\mathsf{MIP}^* = \mathsf{RE}$.

By using Chapter 4, I give a proof for the "rounding" theorem for the commuting operator model in Chapter 5. In essence, the rounding theorem states that any "almost synchronous correlation" can be approximated by exact synchronous correlations. This result was proven in the tensor product model in [Vid22; Pau22]. Due to the result on tracially embeddable strategies, one could mimic the proof from [Vid22] in order to obtain a similar result in the commuting operator model.

1.3.3 Part 3: Interactive proof systems and the compression theorem

In this part, the main goal is to introduce and prove the compression theorem for the complexity classes MIP^* and $\mathsf{MIP}^{\mathsf{co}}$, and combine this theorem with the compression paradigm introduced in Part I to show the main result of this thesis. The theme for this part is to treat a non-local game as a multiprover interactive proof system, and show some results about them.

In Chapter 6, I define the notion of an interactive proof system, and give a formal version of the compression theorem for $\mathsf{MIP}^*/\mathsf{MIP}^{\mathsf{co}}$ (Theorem 6.2.1). Furthermore, I give the algorithms "search from below" and "search from above" (Pseudocode 7 and Pseudocode 5) which put $\mathsf{MIP}^* \in \mathsf{RE}$ and $\mathsf{MIP}^{\mathsf{co}} \in \mathsf{coRE}$ respectively, both of these algorithms are well known in the literature. With these ingredients, I give a proof showing that $\mathsf{MIP}^* = \mathsf{RE}$ and $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$

in this chapter. I also give a proof for showing that giving an explicit separation between the models of entanglement is equivalent to the halting problem. In essence, this section is about how the compression theorem for non-local games is being used.

In Chapter 7, I formally define the notion of a conditionally linear distribution (Definition 7.2.5) and conditionally linear verifier (Definition 7.3.1). The conditionally linear distribution is originally introduced in [JNV+22a, Chapter 4], and the conditionally linear verifier is a more streamlined version of the detyped verifier in [JNV+22a, Definition 6.17]. Although a conditionally linear verifier is one of the key definitions used to define the compression theorem, the definition itself is not quite important for using the compression theorem. This is the main reason why I choose to define it in a separate chapter. I also give a proof outline for the compression theorem, which can be quickly summarized by a limerick I wrote back in my master's thesis [Lin22], presented below:

> Non-local games are Ar Ee complete,
> This means max tensors are not quite discrete,
> Introspect the sampler,
> P-C-P the player,
> and an anchor to parallel repeat!

As suggested by the limerick, the compression theorem is proven using three different transformations, the question reduction protocol known as Introspection, the answer reduction protocol which is based on a PCP construction, and the parallel repetition theorem which uses an anchor transformation. Defining each of the three transformations is very complicated, and each of them, in my opinion, requires different tools. The rest of Part 3 will be divided up in the following manner

- Chapter 8 and Chapter 9 will be focused on the question reduction transformation.
- Chapter 10 and Chapter 11 will be focused on the answer reduction transformation.
- Chapter 12 will be focused on the parallel repetition transformation.

Each of the three portions above is written in a way so that they can be read independently, assuming one understands Chapter 7. Similarly to the rounding theorem, the main contribution of this thesis is to show that the main property for these three transformations in the tensor product model also holds for the commuting operator model using tracially embeddable strategies. I will describe the key idea behind each of the three transformations below.

Question Reduction. The goal of the question reduction protocol is to force the provers to make an "ideal" measurement in order for them to sample a question pair following the conditionally linear verifier. The question reduction protocol is a combination of two tests: the Pauli Basis test and the Introspection protocol. The Pauli Basis test uses self-testing techniques in order to force the dishonest provers to prepare enough EPR pairs required to sample the input distribution, as well as perform an X or Z measurement on all the prepared EPR pairs. The Introspection protocol utilizes the EPR pairs guaranteed by the Pauli Basis test and forces the provers to make the correct Pauli Z measurement so that the provers can sample a question pair from the given conditionally linear distribution. The Pauli Basis test has a question sampling complexity of polylog(n), and the Introspection protocol has a sampling complexity independent of n (where both tests have a decision complexity of poly(n)), thereby reducing the complexity of $\mathtt{Q}$ from poly(n) to polylog(n). In this thesis, I use the simplification due to [dlS22b] for the Pauli Basis test, which circumvents the low-individual degree test subroutine used in the original Pauli Basis test [JNV+22a, Section 7]. As a result, the increase in soundness error in the question reduction theorem proven in this thesis is independent of the size of the game, an improvement over the polynomial dependency in [JNV+22a].

In Chapter 8, I will introduce the Pauli Basis test (Figure 8.2), and prove the soundness property needed for the question reduction protocol. In doing so, I also introduce several representation theory tools, such as the state Gowers-Hatami theorem (Theorem 8.1.1) and the Poincaré inequality (Theorem 8.1.2), which are used to prove the soundness of the Pauli Basis test. Finally, I will introduce the introspection protocol in Chapter 9. I want to remark that these two chapters are loosely based on [JNV+22a, Chapter 7, 8, Appendix A].

I want to remark that the idea of using self-testing to force an "honest sampling procedure" has been used in many earlier works in quantum complexity, and I refer the reader to an elegant paper by Alex Grilo [Gri20], which shows that QMA, the quantum analogue of NP, is in MIP^* for more details.

Answer Reduction. The goal of the answer reduction protocol is to reduce the complexity of the decider $\mathtt{D}$ from poly(n) time to polylog(n) time while retaining a polylog(n) runtime for the sampler. Similarly to [JNV+22a; NW19], a classical probabilistically checkable proof (PCP) is used on the verification Turing machine $\mathtt{D}_n$. In particular, we use the same tailor-made PCP procedure used within [JNV+22a]'s answer reduction protocol as a black box in this thesis. The PCP procedure reduces checking computation steps of $\mathtt{D}_n$ returning accept given the corresponding question/answer pair into checking whether the two provers share the same collection of low individual degree polynomials with specific properties, which can then be checked using the "quantum low-individual degree test" introduced in [JNV+22b].

There is an immediate problem with this approach. Recall that for a classical PCP which verifies an NP instance, the provers are intuitively trying to prove the following statement to the verifier: "given $x \in \mathsf{L}$ and a polynomial size circuit $\mathtt{C}$, there exists a proof string s such that $C(x,s) = 1$". Whereas in the non-local game setting, the provers are trying to prove: "given a validation function D_n for a non-local game, and a question pair (x,y), there exists an answer pair (a,b), which is generated by two entangled provers, such that $D_n(x,y,a,b) = 1$". In order to compute the second statement in a PCP instance, both provers need to somehow play the game using a predetermined entangled strategy and output their answers without communicating with each other. Then, the provers need to pass their answers so that they can encode the computation step of $D_n(x,y,a,b)$ as a PCP instance. In order to get around this issue, a transformation known as Oracularization is applied before computing the PCP instance (see Section 11.2). To ensure completeness is preserved through the oracularization transformation, we need an additional property in the completeness statement in the gap compression theorem: The perfect strategies in the original game must use a special kind of strategy known as an *oracularizable strategy*, which we defined back in Chapter 4.

I will introduce the quantum low-individual degree test (Figure 10.1) in Chapter 10, and prove several properties related to it. I will introduce both the Oracularization transformation (Figure 11.1) and give a summary of the answer reduction protocol (Theorem 11.3.1) in Chapter 11. Since the answer reduction protocol is exactly the same as the one used in the original $\mathsf{MIP}^* = \mathsf{RE}$ paper, I will not write down the transformation (which itself is based on the proof for $\mathsf{MIP} = \mathsf{NEXP}$ [BFL91]) in detail in this thesis. Instead I will refer readers to [JNV+22a, Chapter 10] for more details.

I want to remark that the PCP used in this thesis is vastly different from the more well-known and more modern construction due to Irit Dinur [DV07]. While Dinur's construction is interesting in its own right, I personally would not recommend this to be the starting point for learning the PCP construction for this thesis[12]. To get a basic understanding of this construction, a good starting point would be the excellent course notes by Prahladh Harsha [Har04] (which covers a substantial part of the answer reduction protocol).

[12] In fact, it is an interesting open problem whether one can make this construction work for the MIP^* model

Parallel repetition. Before introducing the parallel repetition theorem, I would like to briefly acknowledge that this part of the thesis originated from an early collaboration with Hamoon Mousavi, William Slofstra, and Henry Yuen.

By applying the above two subroutines to a conditionally linear verifier $\mathscr{V} : \mathbb{N} \to \mathscr{G}$ with complexity $O(\text{poly}(n))$, the resulting conditionally linear verifier $\mathscr{V}' : \mathbb{N} \to \mathscr{G}$ runs in $O(\text{polylog}(n))$ time, such that for $t \in \{*, co\}$

1. (Completeness) If $\omega^t(\mathscr{V}(n)) = 1$, then $\omega^t(\mathscr{V}'(n)) = 1$,
2. (Soundness) If $\omega^t(\mathscr{V}(n)) \leq \frac{1}{2}$, then $\omega^t(\mathscr{V}'(n)) \leq 1 - \text{polylog}(n)$.

Thus, in order to show Theorem 1.2.1, one would need to apply a logarithmic-fold parallel repetition transformation to the game in order to amplify the "soundness" condition for $\mathscr{V}'$ to $\leq \frac{1}{2}$ while retaining the "completeness" condition. Recall that a r-fold parallel repetition is a transformation for the game $\mathscr{G}$ in which r question pairs are sampled independently and sent to the provers, and the provers must treat each of the r question pairs as independent questions and reply with r corresponding answer pairs. The provers only win the r-fold parallel repetition game if they win on all r independent instances of the game. I want to remark that applying a logarithmic-fold parallel repetition to $\mathscr{V}'$ only increases the runtime by a logarithmic factor.

Unfortunately, a strong parallel repetition theorem, i.e. a parallel repetition theorem that shows an exponential decay in the optimal success rate for entangled provers is an open problem. However, [BVY21] shows that by applying a simple transformation, the anchored transformation, the resulting game would have a strong parallel repetition theorem, and this version of parallel repetition has been used in [JNV+22a]. we use a slight modification for the anchoring transformation[13] in this thesis. I give a more detailed summary on the proof of the anchored parallel repetition theorem in Section 12.2.1.

In Chapter 12, we introduce the new information-theoretic tools for the commuting operator model mentioned in the previous section. Although defining the information-theoretic tools uses some advanced theory in operator algebra (for example, the Tomita-Takesaki modular transformation [Tak70]), the tools are actually quite tame in comparison to Chapter 4. Using these tools, we translate the proof from [BVY21] into the commuting operator model.

I want to remark that the anchored parallel repetition for anchored games is actually based on the classical parallel repetition proof due to Thomas Holenstein [Hol09]. Furthermore, to understand why information-theoretic tools are needed in the quantum parallel repetition theorem, and some of the challenges in obtaining a general quantum parallel repetition theorem, I recommend reading the excellent paper by Rahul Jain, Attila Pereszényi and Penghui Yao which shows a quantum parallel repetition theorem for games with product distribution as their sampling procedure [JPY14].

1.3.4 Part 4: Implications and open problem.

In, we discuss some of the implications, and open problems which arise from this thesis. We give a brief summary of some of these implications and open problems below.

In a similar way that $\mathsf{MIP}^* = \mathsf{RE}$ gives a negative resolution to Tsirelson's problem, Kirchberg's problem, and the Connes embedding problem all at once, showing $\mathsf{MIP}^{co} = \mathsf{coRE}$ also gives a negative resolution to the same set of conjectures. However, there is a slight difference, from the operator algebraic standpoint, one should interpret $\mathsf{MIP}^{co} = \mathsf{coRE}$ as an uncomputable result about the type II_1 von Neumann algebra following its connection to the Connes embedding problem, and as an uncomputable result for the maximum tensor product norm following its connection to Kirchberg's conjecture. Hence, proving $\mathsf{MIP}^{co} = \mathsf{coRE}$ actually gives a few uncomputable results in continuous model theory and noncommutative

[13] The modification is designed to preserve synchronicity within the game.

polynomials that do not follow from $\mathsf{MIP}^* = \mathsf{RE}$ alone. We discuss some of these implications in Section 13.1.

I list several open problems which I found interesting in Section 13.2. One interesting problem I want to highlight is whether the compression paradigm can be used for other problems in operator algebra. To be more specific, the compression framework seems to be very good at answering the question of whether an object which is infinite in nature can be approximated by a finite object. One example is the Connes embedding problem, and another is the recently disproved Aldous-Lyons problem by Lewis Bowen, Michael Chapman, Alexander Lubotzky, Thomas Vidick [BCL+24; BCV24] which asks whether one can construct an infinite graph based on a set of local neighbourhood statistics. One famous open problem of this nature is to show the existence of a non-sofic group. This is a famous problem in operator algebra. Another interesting problem is whether non-local games are really needed for these uncomputable results, as both the disproof of the Connes embedding problem and the Aldous-Lyons problem require a very specific set of games. While non-local games do offer a nice structure which can be used to define a compression map, such as self-testing and PCPs, it does feel limiting when there are no longer natural analogues of these structures (for example, in the case of LinMIP^* when it comes to the existence of a non-hyperlinear group). Thus, it would be natural to ask: Is there another structure which falls nicely into the compression framework (as currently, the only two known examples are non-local games and Wang tiling [NMY25; DRS08]), and can one define a natural problem based on Dinur's PCP construction which falls into the compression framework, and thus be compressible? Can these new uncomputable problems potentially be useful for operator algebra? Can one apply the compression paradigm directly to operator algebraic objects?

Part I

The compression paradigm for uncomputable problems

CHAPTER 2

The compression paradigm

We start the technical portion of this thesis by discussing the compression paradigm. The compression paradigm is a new condition for uncomputability based on the proof approach for $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ given in [MNY22]. We remark that this condition is considered and formulated independently in [NMY25]. As observed in [NMY25], there are already several known examples of the compression paradigm in previous literature (and we refer to their paper for more details).

I choose to start my thesis on this technique not only because this is my favourite part of my Ph.D., but I personally believe that this is an alternative way to show uncomputability for problems even outside of quantum information. This part of the thesis has a higher emphasis on computability theory.

2.1 Turing Machine Preliminaries

We start this chapter by introducing some notation related to finite fields and Turing machines we use for the remainder of this thesis. We use $\mathbb{N}$ to denote the set of natural numbers. For integers n, we use $[n]$ to denote the set $\{0,\dots,n-1\}$, and for $n \leq m$ we use $[n,m]$ to denote the set $\{n,n+1,\dots,m-1\}$. We assume all log are in base 2 in this thesis. For $n \in \mathbb{N}$, we use $\mathbf{bin}(n) \in \{0,1\}^{\lceil \log(n) \rceil}$ to denote the binary representation for the number n, and for $s \in \{0,1\}^n$, we use $\mathbf{bininv}(s)$ to denote the unique integer i such that $\mathbf{bin}(i) = s$. For a string $s \in \{0,1\}^*$, we use $|s|_w$ to denote the length of the string s.

We use the Turing machine as the model of computation in this thesis. Let $\mathtt{A}(x_1,\dots,x_m)$ denote an m-input Turing machine. We assume that $\mathtt{A}$, in this case, consists of m input tapes, a single work tape, and a single output tape. When specifying the output of an m-input Turing machine, we might sometimes define it only for accepting fewer than m inputs; in this case, we assume the Turing machine only reads the first n of the input tapes during the computation step. We use $\langle \mathtt{A} \rangle \in \{0,1\}^*$ to denote the description of the Turing machine $\mathtt{A}$. To abuse notation, we use $\langle \mathtt{A}(x) \rangle$ to denote the description of the Turing machine $\mathtt{A}$ being hardcoded to run $x \in \{0,1\}^*$ as input. For $x \in \{0,1\}^*$, we use $\mathtt{A}(x)$ to denote the output of $\mathtt{A}$ on input x, and likewise $\mathtt{A}()$ to denote the output of $\mathtt{A}$ on the empty input. We note that $\mathtt{A}(x)$ might not be well defined as $\mathtt{A}$ does not necessarily halt on the input x. We use $|\mathtt{A}|$ to denote the minimum description length of $\mathtt{A}$, and we note that the description of the Turing machine is a constant that is independent of the input size. We use $\mathsf{TIME}_{\mathtt{A}}(x_1,\cdots,x_m)$ to denote the maximum of $|\mathtt{A}|$ and the runtime of the Turing machine $\mathtt{A}$ running with input $(x_1,\cdots,x_m)$. $\mathsf{TIME}_{\mathtt{A}}(x_1,\cdots,x_m)$ could potentially be ∞ if the Turing machine $\mathtt{A}$ does not halt on input $(x_1,\cdots,x_m)$.

Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of functions that map m finite sets $\{S_i^n\}_{i\in[m]}$ to $\{0,1\}^*$, i.e. $f_n : S_0^n \times \cdots \times S_{m-1}^n \to \{0,1\}^*$. We say the Turing machine $\mathtt{A}$ computes the sequence of functions f_n if $\mathtt{A}$ is an $(m+1)$-input Turing machine in which for all $n \in \mathbb{N}$

$$\mathtt{A}(\mathbf{bin}(n), x_1, \cdots x_m) = f_n(x_1, \cdots x_m),$$

where each element of S_i^n is encoded using a binary representation with $x_i \in S_i^n$. If $\mathtt{A}$ is the Turing machine which computes the functions $\{f_n\}$, to abuse notation, we use $\mathtt{A}_n$ to denote the function f_n. For $n \in \mathbb{N}$, we use the notation $\mathsf{TIME}_{\mathtt{A}}(n)$ to denote the maximum of the function $\mathsf{TIME}_{\mathtt{A}}(n, x_1, \cdots, x_m)$ over all input $x_1 \cdots x_m \in S_0^n \times \cdots \times S_{m-1}^n$, and we use $\mathsf{TIME}_{\mathtt{A}}(n, x_1, \cdots, x_m) = \infty$ if $\mathtt{A}$ does not halt on some input $(x_1, \cdots, x_m)$. In this thesis, if f takes an integer as input, the integer will always be represented under the binary representation, and hence any integer n is considered a $\log(n)$-bit input under this formulation.

We define a language L as a nonempty subset of $\{0,1\}^*$ and a decision problem $\mathsf{D} = (\mathsf{L}^{\mathsf{D}}_{\text{yes}}, \mathsf{L}^{\mathsf{D}}_{\text{no}})$ for two disjoint languages $\mathsf{L}^{\mathsf{D}}_{\text{yes}}, \mathsf{L}^{\mathsf{D}}_{\text{no}} \subseteq \{0,1\}^*$. We use coD to denote the complement of D (i.e. $\mathsf{L}^{\mathsf{coD}}_{\text{yes}} = \mathsf{L}^{\mathsf{D}}_{\text{no}}$ and $\mathsf{L}^{\mathsf{coD}}_{\text{no}} = \mathsf{L}^{\mathsf{D}}_{\text{yes}}$). To abuse notation, we write $x \in \mathsf{D}$ as $x \in \mathsf{L}^{\mathsf{D}}_{\text{yes}} \cup \mathsf{L}^{\mathsf{D}}_{\text{no}}$ and, for a set S, we write $f : S \to \mathsf{D}$ as $f : S \to \mathsf{L}^{\mathsf{D}}_{\text{yes}} \cup \mathsf{L}^{\mathsf{D}}_{\text{no}}$. For two decision problems D_1 and D_2, we write $\mathsf{D}_1 \leq \mathsf{D}_2$ if there exists a Karp reduction from D_1 to D_2 and $\mathsf{D}_1 \leq_p \mathsf{D}_2$ if furthermore the reduction is under polynomial time. For the standard definition of the Karp reduction, we refer the reader to [Sip06, Chapter 5.3] for more details. We define a *uniform problem instance* for D as a Turing machine $\mathtt{Seq} : \mathbb{N} \to \mathsf{D}$. Intuitively, this is a way to package a countable number of decision problems from D in a uniform manner.

We ended this section by introducing two complexity class which will be important for this thesis. Recall from the literature, the complexity class recursively enumerable languages, RE, corresponds to the class of decision problems in which there exists an algorithm that can decide any instance in L_{yes} in finite time (but the same algorithm could potentially run forever for instances in L_{no}). We say that a language $\mathsf{L} \subseteq \{0,1\}^*$ is in RE (or $\mathsf{L} \in \mathsf{RE}$) if there exists an algorithm that can correctly decide whether $x \in \mathsf{L}$ in a finite amount of time.

The halting problem is defined by $\mathsf{L}_{\text{yes}} = \mathsf{L}_{\text{halt}}$ and $\mathsf{L}_{\text{no}} = \mathsf{L}_{\text{nothalt}}$, with the definition of $\mathsf{L}_{\text{halt}}, \mathsf{L}_{\text{nothalt}}$ given below:

- L_{halt}: The set of Turing machines (represented under the binary description) that halt on the empty input,
- $\mathsf{L}_{\text{nothalt}}$: The set of Turing machines which does not halt on the empty input.

The halting problem is complete for RE, meaning that $(\mathsf{L}_{\text{halt}}, \mathsf{L}_{\text{nohalt}}) \in \mathsf{RE}$ and for any language $Ł \in \mathsf{RE}$, we have $Ł \leq \mathsf{RE}$.

Similarly, the complexity class coRE, or the complement of RE, corresponds to the class of decision problems in which there exists an algorithm which can decide all instances in L_{no} in finite time. We say that $\mathsf{L} \in \mathsf{coRE}$ if there exists an algorithm that can correctly decide whether $x \notin \mathsf{L}$ in a finite amount of time (but could potentially run forever if $x \in \mathsf{L}$). The non-halting problem is complete for coRE, where recall the non-halting problem is defined similarly as the halting problem but with the "yes" and "no" instances being swapped, or $\mathsf{L}^{\mathsf{coRE}}_{\text{yes}} = \mathsf{L}_{\text{nothalt}}$, $\mathsf{L}^{\mathsf{coRE}}_{\text{no}} = \mathsf{L}_{\text{halt}}$.

For a more comprehensive introduction on computability and complexity theory, we refer the reader to [Sip06].

2.2 Compressible condition and the compression theorem

In this section, we introduce the compression condition for general promise problems, and show the equivalence between a compressible language and the language being complete for RE or coRE. We introduce the notion of a compressible promise problem below.

Definition 2.2.1 (Compressible problems)**.** *Let* $\mathsf{D} = (\mathsf{L}^{\mathsf{D}}_{yes}, \mathsf{L}^{\mathsf{D}}_{no})$ *be a promise problem. We say that the promise problem* D *is compressible if there exists an algorithm* $\mathtt{Compress}^{\mathsf{D}}$ *which takes as input* $\langle \mathcal{S}eq_{\mathsf{D}} \rangle$*, a description of a uniform instance of* D*.* $\mathtt{Compress}^{\mathsf{D}}$ *outputs a description of a uniform instance of* D*,* $\langle \mathcal{S}eq^{Comp}_{\mathsf{D}} \rangle$*, such that the following holds:*

- *(Runtime):* $\mathsf{TIME}_{\mathtt{Compress}^{\mathsf{D}}} = O(\text{poly}(|\langle \mathtt{Seq}_{\mathsf{D}} \rangle|))$.
- *(Consistency of the output)* $\mathtt{Seq}_{\mathsf{D}}^{Comp}(n) \in \mathsf{D}$ *for all* $n \in \mathbb{N}$*, even if the initial input for* $\mathtt{Compress}^{\mathsf{D}}$ *is not a valid uniform instance of* D.
- *(Complexity bound for the output)* $\mathsf{TIME}_{\mathtt{Seq}_{\mathsf{D}}^{Comp}} = O(\text{polylog}(n))$.

Furthermore, if $\mathsf{TIME}_{\mathtt{Seq}_{\mathsf{D}}} = O(\text{poly}(n))$*, then for all* $n \in \mathbb{N}$*, the following holds:*

- *(Completeness)* $\mathtt{Seq}_{\mathsf{D}}(n) \in \mathsf{L}^{\mathsf{D}}_{yes} \implies \mathtt{Seq}_{\mathsf{D}}^{Comp}(n) \in \mathsf{L}^{\mathsf{D}}_{yes}$.
- *(Soundness)* $\mathtt{Seq}_{\mathsf{D}}(n) \in \mathsf{L}^{\mathsf{D}}_{no} \implies \mathtt{Seq}_{\mathsf{D}}^{Comp}(n) \in \mathsf{L}^{\mathsf{D}}_{no}$.

Intuitively, one should interpret the algorithm $\mathtt{Compress}^{\mathsf{D}}$ given in the above as a property associated with the promise problem rather than the uniform sequence itself. This means that $\mathtt{Compress}^{\mathsf{D}}$ is a single algorithm which works **for all** uniform instances of $\mathtt{Seq}_{\mathsf{D}}$. Intuitively, the compressible property allows one to generate a problem instance more efficiently (even though the new problem instance might be equally hard to decide). If D is compressible, then co$\mathsf{D}$ is also compressible by definition. We use the following clever example given in [NMY25] to show that the halting problem is compressible.

Example 2.2.2 (The Halting problem is compressible)**.** For the Halting problem, consider the compression algorithm $\mathtt{Compress}^{\mathtt{HALT}}$ for the halting problem defined by the following: given the input $\langle \mathtt{Seq}_{\mathtt{HALT}} \rangle$, $\mathtt{Compress}^{\mathtt{HALT}}$ returns a description of $\langle \mathtt{Seq}_{\mathtt{HALT}}^{\mathrm{Comp}} \rangle$, described by Pseudocode 1.

1 **Input**: Integer n
2 Compute and return $\langle \mathtt{Prog}_n^{\langle \mathtt{Seq}_{\mathtt{HALT}} \rangle} \rangle$, where the pseudocode for $\mathtt{Prog}_n^{\langle \mathtt{Seq}_{\mathtt{HALT}} \rangle}$ is given in Pseudocode 2

Pseudocode 1: The description of $\mathtt{Seq}_{\mathtt{HALT}}^{\mathrm{Comp}}$.

1 Halt and return ERROR if $\langle \mathtt{Seq}_{\mathtt{HALT}} \rangle$ is not a valid description for a Turing machine.
2 Compute the description of $\langle \mathtt{Seq}_{\mathtt{HALT}}(n) \rangle$ using $\langle \mathtt{Seq}_{\mathtt{HALT}} \rangle$ and n.
3 Run the program $\langle \mathtt{Seq}_{\mathtt{HALT}}(n) \rangle$ and return the corresponding output, halt and return ERROR if $\langle \mathtt{Seq}_{\mathtt{HALT}}(n) \rangle$ is not a valid description for a Turing machine.

Pseudocode 2: The description of the output for $\mathtt{Prog}_n^{\langle \mathtt{Seq}_{\mathtt{HALT}} \rangle}$. We remark that the above program depends on both $\mathtt{Seq}_{\mathtt{HALT}}$ and n.

In the above example, one should interpret the description of $\langle \mathtt{Prog}_n^{\langle \mathtt{Seq}_{\mathtt{HALT}} \rangle} \rangle$ given in Pseudocode 2 as an instance for the halting problem (and hence, the runtime of line 1 from Pseudocode 2 is irrelevant). The program $\mathtt{Seq}_{\mathtt{HALT}}^{\mathrm{Comp}}$ merely generates different instances of the halting problem by outputting the description of $\langle \mathtt{Prog}_n^{\langle \mathtt{Seq}_{\mathtt{HALT}} \rangle} \rangle$.

Since the Turing machine $\mathtt{Seq}_{\mathtt{HALT}}^{\mathrm{Comp}}(n)$ only returns the description for $\langle \mathtt{Prog}_n^{\langle \mathtt{Seq}_{\mathtt{HALT}} \rangle} \rangle$ which depends on n, the description $\langle \mathtt{Seq}_{\mathtt{HALT}}^{\mathrm{Comp}}(n) \rangle$ can be computed in $O(\text{poly}(|\langle \mathtt{Seq} \rangle|))$ time. Since any integer input is represented under the binary representation, $\mathsf{TIME}_{\mathtt{Seq}_{\mathsf{D}}^{\mathrm{Comp}}} = O(\text{polylog}(n))$. Furthermore, for all integer $n \in \mathbb{N}$

- If $\mathtt{Seq}_{\mathtt{HALT}}(n)$ returns a description of a halting program, then $\mathtt{Seq}_{\mathtt{HALT}}^{\mathrm{comp}}(n) = \langle \mathtt{Prog}_n^{\langle \mathtt{Seq}_{\mathtt{HALT}} \rangle} \rangle$ is a description of a halting program.
- Otherwise, if $\mathtt{Seq}_{\mathtt{HALT}}(n)$ returns a description of a non-halting program, then $\mathtt{Seq}_{\mathtt{HALT}}^{\mathrm{comp}}(n) = \langle \mathtt{Prog}_n^{\langle \mathtt{Seq}_{\mathtt{HALT}} \rangle} \rangle$ is a description of a non-halting program.

This shows that the halting problem is an example of a compressible promise problem.

The above example also shows that all RE or coRE-complete problems are also compressible. Although the compressible property offers a new condition for an RE or coRE-complete promise problem, it is often hard to construct the `Compress` algorithm and make this characterization practical. To this end, we give a weaker notion of compressibility below.

Definition 2.2.3 (Weakly compressible problems)**.** *Let* $\mathsf{D} = (\mathsf{L}^{\mathsf{D}}_{yes}, \mathsf{L}^{\mathsf{D}}_{no})$ *be a promise problem. We say that the promise problem* D *is weakly compressible if for every* $\alpha \in \mathbb{N}$*, there exists an algorithm* $\mathtt{Compress}^{\mathsf{D}}_{\alpha}$*, which takes, as input,* $\langle \mathcal{S}eq_{\mathsf{D}} \rangle$*, a description of a uniform instance of promise problemsfor* D*.* $\mathtt{Compress}^{\mathsf{D}}_{\alpha}$ *outputs a description for a uniform instance of promise problems for* D*,* $\langle \mathcal{S}eq^{Comp}_{\mathsf{D}} \rangle$*, such that the following holds: There exists an integer* $\gamma = O(\mathrm{poly}(\alpha))$ *such that*

1. *(Runtime):* $\mathsf{TIME}_{\mathtt{Compress}^{\mathsf{D}}_{\alpha}} = O(\mathrm{poly}(|\langle \mathcal{S}eq_{\mathsf{D},\alpha} \rangle|, \alpha))$.
2. *(Consistency of the output):* $\mathcal{S}eq^{Comp}_{\mathsf{D}}(n) \in \mathsf{D}$ *for all* $n \in \mathbb{N}$*, even if the initial input for* $\mathtt{Compress}^{\mathsf{D}}_{\alpha}$ *is not a valid uniform instance of* D*.*
3. *(Complexity bound for the output):* $\mathsf{TIME}_{\mathcal{S}eq^{Comp}_{\mathsf{D}}} = O(\log(n)^{\gamma})$.

Furthermore, if there exists some constant $n_0 \in \mathbb{N}$ *such that for all* $n \geq n_0$

$$\mathsf{TIME}_{\mathcal{S}eq_{\mathsf{D}}} \leq n^{\alpha}. \tag{2.1}$$

Then there exists some constant $n^{Comp}_0 = \mathrm{poly}(\gamma, n_0)$ *such that for all* $n \geq n^{Comp}_0$

- *(Completeness):* $\mathcal{S}eq_{\mathsf{D}}(n) \in \mathsf{L}^{\mathsf{D}}_{yes} \implies \mathcal{S}eq^{Comp}_{\mathsf{D}}(n) \in \mathsf{L}^{\mathsf{D}}_{yes}$.
- *(Soundness):* $\mathcal{S}eq_{\mathsf{D}}(n) \in \mathsf{L}^{\mathsf{D}}_{no} \implies \mathcal{S}eq^{Comp}_{\mathsf{D}}(n) \in \mathsf{L}^{\mathsf{D}}_{no}$.

Instead of requiring a single `Compress` algorithm which works for all uniform problem instances, the weakly compressible condition instead just requires a $\mathtt{Compress}_{\alpha}$ algorithm that compresses all uniform problem instances that run in $O(n^{\alpha})$ time for all $\alpha \in \mathbb{N}$ (i.e. these algorithms could be different depending on α). If D is compressible, then it is trivially weakly compressible. We have the following two theorems relating the compressible condition to RE or coRE-complete languages.

Theorem 2.2.4 (Compression criteria for RE-complete problems)**.** *Let* $\mathsf{D} = (\mathsf{L}^{\mathsf{D}}_{yes}, \mathsf{L}^{\mathsf{D}}_{no})$ *be a promise problem. If* $\mathsf{L}^{\mathsf{D}}_{no} \in \mathsf{coRE}$*, then the following are equivalent:*

1. D *is* RE*-complete.*
2. D *is a compressible promise problem.*
3. D *is a weakly compressible promise problem.*

Theorem 2.2.5 (Compression criteria for coRE-complete problems)**.** *Let* $\mathsf{D} = (\mathsf{L}^{\mathsf{D}}_{yes}, \mathsf{L}^{\mathsf{D}}_{no})$ *be a promise problem. If* $\mathsf{L}^{\mathsf{D}}_{yes} \in \mathsf{coRE}$*, then the following are equivalent:*

1. D *is* coRE*-complete.*
2. D *is a compressible promise problem.*
3. D *is a weakly compressible promise problem.*

Since whenever D is compressible/weakly compressible, co$\mathsf{D}$ is also compressible/weakly compressible, this implies that Theorem 2.2.4 implies Theorem 2.2.5 being true. Hence, we provide a proof for Theorem 2.2.4 below. We remark that this proof is inspired by the suggested approach for showing $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ in [MNY22, Conjecture 1.4].

Proof of Theorem 2.2.4. Note that (2) trivially implies (3), and by Example 2.2.2 (1) implies (2). Hence all that remains to show is (3) implies (1).

Fix a $\mathsf{D} = (\mathsf{L}^{\mathsf{D}}_{\text{yes}}, \mathsf{L}^{\mathsf{D}}_{\text{no}})$ as per Theorem 2.2.4, and assume that D is a weakly compressible problem. Since by assumption, $\mathsf{L}^{\mathsf{D}}_{\text{no}} \in \mathsf{coRE}$, let $\texttt{coREalgo}_{\mathsf{L}_{\text{no}}}$ denote the coRE algorithm that halts whenever $x \notin \mathsf{L}^{\mathsf{D}}_{\text{no}}$ and runs forever if $x \in \mathsf{L}^{\mathsf{D}}_{\text{no}}$ (this can be assumed by appending an infinite loop whenever $\texttt{coREalgo}_{\mathsf{L}_{\text{no}}}$ halts and correctly decides $x \in \mathsf{L}^{\mathsf{D}}_{\text{no}}$). The algorithm $\texttt{coREalgo}_{\mathsf{L}_{\text{no}}}$ already implies that $\mathsf{D} \in \mathsf{RE}$. Hence the only thing we need to show is that D can be reduced to the halting problem.

Hence, fix a Turing machine 🐟 (🐟 (pronounced “fish") is the only non-western character used in this thesis, and we use it to emphasize the fact that it is an instance of the Halting problem instead of a subroutine defined within this thesis[1]). We wish to find an instance $x_{🐟}$ such that whenever 🐟 halts, then $x_{🐟} \in \mathsf{L}^{\mathsf{D}}_{\text{yes}}$ and $x_{🐟} \in \mathsf{L}^{\mathsf{D}}_{\text{no}}$ otherwise. For every $\beta \in \mathbb{N}$, let $\texttt{Compress}^{\mathsf{D}}_{\beta}$ be the compression algorithm guaranteed by the weakly compressible condition given in Definition 2.2.3. Since we assume $\mathsf{L}^{\mathsf{D}}_{\text{yes}}$ and $\mathsf{L}^{\mathsf{D}}_{\text{no}}$ are non-empty in the preliminary, let $x_{\text{yes}} \in \mathsf{L}^{\mathsf{D}}_{\text{yes}}$ and $x_{\text{no}} \in \mathsf{L}^{\mathsf{D}}_{\text{no}}$ be an arbitrary element in these two sets. Let $\alpha, C_0 \in \mathbb{N}$ be two constants to be specified later in the proof. We construct the following uniform game sequence $\texttt{Seq}_{\mathsf{D},🐟}$ based on 🐟 in Pseudocode 3.

1 **Input**: Integer n.
2 Run 🐟 for n steps. If 🐟 halts in the given steps, **return** x_{yes}.
3 Compute the description of $\langle \texttt{Seq}_{\mathsf{D},🐟} \rangle$.
4 Compute $\langle \texttt{Seq}_{\mathsf{D},🐟}(C_0) \rangle$, the Turing machine which is hardcoded into computing $\texttt{Seq}_{\mathsf{D},🐟}(C_0)$.
5 Simulate line 6-7 for $\max\{0, n - C_0\}$ steps, if line 6-7 halts in the given steps, **return** x_{no}.
6 Run the Turing machine $\langle \texttt{Seq}_{\mathsf{D},🐟}(C_0) \rangle$ until $\texttt{Seq}_{\mathsf{D},🐟}(C_0)$ is computed.
7 Run $\texttt{coREalgo}_{\mathsf{L}_{\text{no}}}$ with $\texttt{Seq}_{\mathsf{D},🐟}(C_0)$ as input.
8 Otherwise, apply $\texttt{Compress}^{\mathsf{D}}_{\alpha}$ with the input $\langle \texttt{Seq}_{\mathsf{D},🐟} \rangle$ to obtain the description for $\langle \texttt{Seq}^{\text{comp}}_{\mathsf{D},🐟} \rangle$.
9 Compute and **return** $\texttt{Seq}^{\text{comp}}_{\mathsf{D},🐟}(n+1)$

Pseudocode 3: The description for $\texttt{Seq}_{\mathsf{D},🐟}$.

We point out that the length of the source code $|\langle \texttt{Seq}_{\mathsf{D},🐟} \rangle|$ only depends on the Turing machine 🐟. In line 3, we use Kleene's Recursion Theorem to allow $\langle \texttt{Seq}_{\mathsf{D},🐟} \rangle$ to perform computation on its own source code. This step can be performed in

$$O(\text{poly}(|\langle \texttt{Seq}_{\mathsf{D},🐟} \rangle|) = O(\text{poly}(|🐟|))$$

time. We also point out that $\texttt{Seq}_{\mathsf{D},🐟}$ is a valid uniform problem instance (i.e. its output range is in D), since it can only terminate in line 2 and 5 (or else the output is x_{yes} and x_{no} respectively, which are both in D), and in line 9 (which is in D by point 2 of Definition 2.2.3). We begin by deriving the runtime for $\texttt{Seq}_{\mathsf{D},🐟}$ on input $n \in \mathbb{N}$ by examining Pseudocode 3 line by line:

- For line 2, simulating 🐟 for n steps takes time

$$\text{poly}(|🐟|, n).$$

- For line 3, by Kleene's Recursion theorem, computing the description of $\langle \texttt{Seq}_{\mathsf{D},🐟} \rangle$ takes time

$$\text{poly}(|🐟|, |\langle \texttt{Seq}_{\mathsf{D},🐟} \rangle|, n) = \text{poly}(|🐟|, n).$$

[1] The original choice for this symbol is the Chinese character for Turing machine, or 图灵(pronounced Túlíng).

- For line 4, the Turing machine $\langle \mathtt{Seq}_{\mathsf{D},\text{🐢}}(C_0)\rangle$ is essentially $\langle \mathtt{Seq}_{\mathsf{D},\text{🐢}}\rangle$, but replacing every instance of n occurring after line 1 with C_0. Since C_0 is a constant, this takes time

$$O(\text{poly}(|\langle \mathtt{Seq}_{\mathsf{D},\text{🐢}}\rangle|, C_0)).$$

- For line 5-7, simulating line 6-7 for $\max\{0, n - C_0\}$ steps takes time

$$O(\text{poly}(|\mathscr{G}_{C_0}|, C_0, n)) = O(\text{poly}(n, C_0, |\text{🐢}|)).$$

- For line 8 and 9, applying $\mathtt{Compress}^{\mathsf{D}}_{\alpha}$ on the input $\langle \mathtt{Seq}_{\mathsf{D},\text{🐢}}\rangle$ and computing $\mathtt{Seq}^{\text{comp}}_{\mathsf{D},\text{🐢}}(n+1)$ takes time

$$O(\text{poly}(\alpha, \log(n)^{\text{poly}(\alpha)}, |\langle \mathtt{Seq}_{\mathsf{D},\text{🐢}}\rangle|))$$

  by condition 1 and 3 in Definition 2.2.3.

Although many parameters appear in the runtime analysis, the only variable which is not set to a constant is n! By combining the runtime analysis above, we have

$$\mathsf{TIME}_{\mathtt{Seq}_{\mathsf{D},\text{🐢}}}(n) = O(\text{poly}(n, \log(n)^{\text{poly}(\alpha)}, \alpha, |\text{🐢}|, C_0)).$$

Since C_0 does not depend exponentially on α in the above equation, there exists a choice for $\alpha, n_0 \in \mathbb{N}$ such that by setting $C_0 = g(n_0, \mathbf{f}(\alpha))$, where g is the polynomial used to define γ and $\mathbf{f}$ is the polynomial used to define n_0^{Comp} in Definition 2.2.1, we have

$$\mathsf{TIME}_{\mathtt{Seq}_{\mathsf{D},\text{🐢}}}(n) \leq n^{\alpha},$$

for all $n \geq n_0$. Fix α, C_0 and n_0 as the constant which satisfies the property above. By the completeness/soundness condition of Definition 2.2.1, for all $n \geq n_0$

- $\mathtt{Seq}_{\mathsf{D},\text{🐢}} \in \mathsf{L}^{\mathsf{D}}_{\text{yes}} \implies \mathtt{Seq}^{\text{Comp}}_{\mathsf{D},\text{🐢}}(n) \in \mathsf{L}^{\mathsf{D}}_{\text{yes}}$,
- $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(n) \in \mathsf{L}^{\mathsf{D}}_{\text{no}} \implies \mathtt{Seq}_{\mathsf{D},\text{🐢}}(n) \in \mathsf{L}^{\mathsf{D}}_{\text{no}}$.

We argue that $x_{\text{🐢}} = \mathtt{Seq}_{\mathsf{D},\text{🐢}}(C_0) \in \mathsf{D}$ is the instance needed for the proof (i.e. $x_{\text{🐢}} \in \mathsf{L}^{\mathsf{D}}_{\text{yes}}$ if 🐢 halts, and $x_{\text{🐢}} \in \mathsf{L}^{\mathsf{D}}_{\text{no}}$ otherwise). We first show that $\langle \mathtt{Seq}_{\mathsf{D},\text{🐢}}\rangle$ can never terminate at line 5 of Pseudocode 3. Consider $\mathtt{Seq}_{\mathsf{D},\text{🐢}}$ with input $n < C_0$; since line 5 does not perform any operation by definition, it also can never terminate in this spot. Similarly, if 🐢 halts in time T, for any input $n > T$, Pseudocode 3 terminates in line 2 before reaching line 5.

Hence, suppose for a contradiction that $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(n)$ terminates at line 5 for some input $n \geq C_0$, and suppose that 🐢 does not halt on step n, and without loss of generality assume that n is the smallest natural number such that $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(n)$ terminates at line 5. By definition, this means $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(n) = x_{\text{no}} \in \mathsf{L}^{\mathsf{D}}_{\text{no}}$.

Since $\langle \mathtt{Seq}_{\mathsf{D},\text{🐢}}(C_0)\rangle$ is a terminating program, this implies that for $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(n)$ to terminate at line 5, $\mathtt{coREalgo}_{\mathsf{L}_{\text{no}}}(\mathtt{Seq}_{\mathsf{D},\text{🐢}}(C_0))$ also would terminate. By the definition of $\mathtt{coREalgo}_{\mathsf{L}_{\text{no}}}$, we have $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(C_0) \notin \mathsf{L}^{\mathsf{D}}_{\text{no}}$. Since $\langle \mathtt{Seq}_{\mathsf{D},\text{🐢}}\rangle$ is a valid uniform sequence, we have $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(C_0) \in \mathsf{L}^{\mathsf{D}}_{\text{yes}}$.

Now consider $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(n-1)$, since n is the smallest natural number such that $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(n)$ terminates at line 5, and the Turing machine 🐢 does not halt in $n-1$ steps. $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(n-1)$, by default terminates on line 9 of Pseudocode 3. This means that $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(n-1) = \mathtt{Seq}_{\mathsf{D},\text{🐢}}(n)^{\text{comp}} \in \mathsf{L}^{\mathsf{D}}_{\text{no}}$ by the weakly compressible condition. By an inductive argument, since $n > C_0$, this implies that $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(C_0) \in \mathsf{L}^{\mathsf{D}}_{\text{no}}$, contradicting the fact that $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(C_0) \in \mathsf{L}^{\mathsf{D}}_{\text{yes}}$. Thus, $\langle \mathtt{Seq}_{\mathsf{D},\text{🐢}}\rangle$ cannot halt on line 5 of Pseudocode 3.

We first focus on the case where 🐢 terminates in T steps. If $T \leq C_0$, then $\mathtt{Seq}_{\mathsf{D},\text{🐢}}(T) = x_{\text{yes}} \in \mathsf{L}^{\mathsf{D}}_{\text{yes}}$ by line 1 of Pseudocode 3. Hence, assume $C_0 < T$. By line 2 of Pseudocode 3,

$\texttt{Seq}_{\mathsf{D},\circledcirc}(T) = x_{\text{yes}} \in \mathsf{L}^{\mathsf{D}}_{\text{yes}}$. Now, consider $\texttt{Seq}_{\mathsf{D},\circledcirc}(T-1)$, since it cannot terminate at line 2 and 5 of Pseudocode 3, this implies that the Turing machine $\langle \texttt{Seq}_{\mathsf{D},\circledcirc}(T-1)\rangle$ will terminate on line 9 of Pseudocode 3. Hence we have $\texttt{Seq}_{\mathsf{D},\circledcirc}(T-1) = \texttt{Seq}^{\text{comp}}_{\mathsf{D},\circledcirc}(T) \in \mathsf{L}^{\mathsf{D}}_{\text{yes}}$ by the weakly compressible condition. Again, by an inductive argument, we have $\texttt{Seq}_{\mathsf{D},\circledcirc}(T-1) \in \mathsf{L}^{\mathsf{D}}_{\text{yes}}$, which completes the proof for the case where $\circledcirc$ halts in finite time.

Now, suppose $\circledcirc$ does not halt. By the above claim, $\texttt{coREalgo}_{\mathsf{L}_{\text{no}}}$ also does not terminate when given the input $\texttt{Seq}_{\mathsf{D},\circledcirc}(C_0)$ (or else $\texttt{Seq}_{\mathsf{D},\circledcirc}$ terminates at line 5 of Pseudocode 3, deriving a contradiction). By the definition of $\texttt{coREalgo}_{\mathsf{L}_{\text{no}}}$, this implies that $\texttt{Seq}_{\mathsf{D},\circledcirc}(C_0) \in \mathsf{L}^{\mathsf{D}}_{\text{no}}$. Hence showing that D is RE-complete. This shows (3) implies (1) in Theorem 2.2.4, which completes the proof for Theorem 2.2.4. □

2.3 Compression on succinct representation of languages

The compressible condition is a powerful new technique for showing uncomputability for different sets of languages. However, since the compressible condition requires an efficient compression map for all uniform problem instances, it is often hard to construct such mappings in practice. To make this technique more usable, it is often more convenient to consider the compressible condition using the *succinct description* for a given decision problem. To this end, we introduce a notion of a uniform succinct problem instance below.

Definition 2.3.1 (Uniform succinct problem instance)**.** *For a Turing machine $\mathcal{Seq} : \mathbb{N}\times\mathbb{N} \to \{0,1,\bot\}$, we say that $\mathcal{Seq}$ is a uniform succinct problem for D if there exist a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathsf{D}$ such that for $n, i \in \mathbb{N}$, $\mathcal{Seq}$ terminates and*

$$\mathcal{Seq}(n,i) = \begin{cases} (x_n)_i & \text{if } |x_n|_w \le i \\ \bot & \text{Otherwise.} \end{cases} \tag{2.2}$$

Furthermore, we say that $\mathcal{Seq}$ runs in $f(n)$ time if $\mathsf{TIME}_{\mathcal{Seq}}(n,i) = f(n)$ (i.e. only depends on n) and $x_n \le f(n)$.

For a uniform succinct problem instance $\texttt{Seq}$ which runs in $O(f(n))$ time, since integers are represented in binary in this thesis, generating each instance x_n given the input $n \in \mathbb{N}$ takes $O(2^{f(n)})$ time. Similarly to Definition 2.2.1, we define a compressible condition for a uniform succinct problem instance below. We remark that the definition is more similar to the one given in [NMY25].

Definition 2.3.2 (Succinct compressible problems)**.** *Let $\mathsf{D} = (\mathsf{L}^{\mathsf{D}}_{yes}, \mathsf{L}^{\mathsf{D}}_{no})$ be a promise problem. We say that the promise problem D is succinct compressible if there exists an algorithm $\texttt{CompressSucc}^{\mathsf{D}}$, which takes, as input, $\langle \mathcal{Seq}_{\mathsf{D}}\rangle$, a description of a uniform succinct problem instance of D. $\texttt{Compress}^{\mathsf{D}}$ outputs a description of a uniform succinct problem instance of D, $\langle \mathcal{Seq}^{Comp}_{\mathsf{D}}\rangle$, such that the following holds:*

- *(Runtime):* $\mathsf{TIME}_{\texttt{CompressSucc}^{\mathsf{D}}} = O(\text{poly}(|\langle \mathcal{Seq}_{\mathsf{D}}\rangle|))$.
- *(Structure of the output) $\langle \mathcal{Seq}^{Comp}_{\mathsf{D}}\rangle$ is a uniform succinct problem for D with runtime of $O(\text{polylog}(n))$.*

Furthermore, if $\langle \mathcal{Seq}^{Comp}_{\mathsf{D}}\rangle$ is a uniform succinct problem for D with runtime of $O(\text{poly}(n))$, let $\{x_n\}$ and $\{x^{comp}_n\}$ be the two sequence decision problems given by $\langle \mathcal{Seq}_{\mathsf{D}}\rangle$ and $\langle \mathcal{Seq}^{Comp}_{\mathsf{D}}\rangle$ respectively. For all $n \in \mathbb{N}$, the following holds:

- *(Completeness) $x_n \in \mathsf{L}^{\mathsf{D}}_{yes} \implies x^{comp}_n \in \mathsf{L}^{\mathsf{D}}_{yes}$.*
- *(Soundness) $x_n \in \mathsf{L}^{\mathsf{D}}_{no} \implies x^{comp}_n \in \mathsf{L}^{\mathsf{D}}_{no}$.*

Similarly, one can define weakly succinctly compressible condition using a similar formulation as Definition 2.2.3. By replacing Pseudocode 3 with Pseudocode 4, one can show the following theorem:

Theorem 2.3.3 (Compression criteria for RE-complete problems)**.** *Let* $\mathsf{D} = (\mathsf{L}^{\mathsf{D}}_{yes}, \mathsf{L}^{\mathsf{D}}_{no})$ *be a promise problem. If* $\mathsf{L}^{\mathsf{D}}_{no} \in \mathsf{coRE}$*, then the following are equivalent:*

1. D *is* RE*-complete.*
2. D *is a succinct compressible promise problem.*
3. D *is a succinct weakly compressible promise problem.*

1 **Input**: Integer n and i
2 Run $\mathcal{M}$ for n steps. If $\mathcal{M}$ halts in the given steps, **return** $(x_{\text{yes}})_i$.
3 Compute the description of $\langle \mathtt{Seq}_{\mathsf{D},\mathcal{M}} \rangle$.
4 Compute $\langle \mathtt{Seq}_{\mathsf{D},\mathcal{M}}(C_0, \cdot) \rangle$, the Turing machine which is hardcoded into computing $\mathtt{Seq}_{\mathsf{D},\mathcal{M}}$ with the first input fixed to C_0.
5 Simulate line 6-7 for $\max\{0, n - C_0\}$ steps, if line 6-7 halts in the given steps, **return** $(x_{\text{no}})_j$.
6 Compute x_{C_0} by enumerating through $\mathtt{Seq}_{\mathsf{D},\mathcal{M}}(C_0, j)$ from $j = 1$ until $\mathtt{Seq}_{\mathsf{D},\mathcal{M}}(C_0, j) = \perp$.
7 Run $\mathtt{coREalgo}_{\mathsf{L}_{\text{no}}}$ with x_1 as input.
8 Otherwise, apply $\mathtt{Compress}^{\mathsf{D}}_{\alpha}$ with the input $\langle \mathtt{Seq}_{\mathsf{D},\mathcal{M}} \rangle$ to obtain the description for $\langle \mathtt{Seq}^{\text{comp}}_{\mathsf{D},\mathcal{M}} \rangle$.
9 Compute and **return** $\mathtt{Seq}^{\text{comp}}_{\mathsf{D},\mathcal{M}}(n+1, i)$

Pseudocode 4: The description for $\mathtt{Seq}_{\mathsf{D},\mathcal{M}}$ for proving Theorem 2.3.3, note that for all $n \in \mathbb{N}$, x_n will always be finite (and bounded) due to the "structure of the output" clause from Definition 2.3.2.

In practice, the succinct description could vary depending on the language, or even a particular type of succinct description is sufficient to argue uncomputability. Intuitively, the uniform succinct problem instance allows one to extract a circuit which computes partial information about the problem instance itself, which could potentially be useful in defining a more concrete compression map. For example, as observed in [NMY25], the proof of the uncomputability for the Wang tiling proof by Durand, Romashchenko and Shen [DRS08] can be phrased in this framework, but the succinct description used is actually a Turing machine which, given four colours, decides whether the tile is in the tile set. As seen later in this thesis, the MIP^* or the MIP^{co} compression relies on the specific succinct representation, whereby the sample distribution and the decision function are encoded as a Turing machine. This encoding allows one to use a variant of the low-individual degree based PCP theorem to construct such a compression map. Interestingly, both of these proof approaches build upon the problem being NP-hard in a succinct manner.

We end this chapter by listing a few of the failed attempts in formulating the compression paradigm for succinct problem. Let $\mathsf{D} = (\mathsf{L}^{\mathsf{D}}_{\text{yes}}, \mathsf{L}^{\mathsf{D}}_{\text{no}})$ be a decision problem, and consider the decision problem $\mathtt{Succ}\mathsf{poly}\mathsf{D} = (\mathtt{Succ}\mathsf{poly}\mathsf{L}^{\mathsf{D}}_{\text{yes}}, \mathtt{Succ}\mathsf{poly}\mathsf{L}^{\mathsf{D}}_{\text{no}})$ defined as the following

- $\mathtt{Succ}\mathsf{poly}\mathsf{L}^{\mathsf{D}}_{\text{yes}} = \{\langle \mathtt{Seq} \rangle | \langle \mathtt{Seq} \rangle \text{ runs in polynomial time}, \mathtt{Seq}() \in \mathsf{L}^{\mathsf{D}}_{\text{yes}}\}$
- $\mathtt{Succ}\mathsf{poly}\mathsf{L}^{\mathsf{D}}_{\text{no}} = \{\langle \mathtt{Seq} \rangle | \langle \mathtt{Seq} \rangle \text{ runs in polynomial time}, \mathtt{Seq}() \in \mathsf{L}^{\mathsf{D}}_{\text{no}}\}$

In other words, $\mathtt{Succ}\mathsf{poly}\mathsf{D}$ consists of the set of Turing machines that run in poly(n) time, and output an instance of the original decision problem. By using a similar compression map

as Example 2.2.2, one can show that SuccpolyD is compressible, however, one can just decide the problem instance by running Seq⟩ until it halts.

It turns out, $\texttt{Succpoly}\mathsf{L}^{\mathsf{D}}_{\text{no}}$ is not in coRE. By Rice's theorem, one can reduce the halting problem into deciding whether ⟨Seq⟩ runs in polynomial time (and outputs something from $\mathsf{L}^{\mathsf{D}}_{\text{no}}$)[2]. $\texttt{Succpoly}\mathsf{L}^{\mathsf{D}}_{\text{no}} \in \mathsf{coRE}$ would imply that $\mathsf{RE} \in \mathsf{coRE}$ which is a clear contradiction. For a similar reason, $\texttt{Succpoly}\mathsf{L}^{\mathsf{D}}_{\text{yes}}$ as a language is also compressible, because $\texttt{Succpoly}\mathsf{L}^{\mathsf{D}}_{\text{yes}}$ is uncomputable.

2.4 Remarks about the generalized compression framework

In this section, we give some remarks about the compression framework. If we take a Turing machine 🚗 which halts in time $T \in \mathbb{N}$ and apply the argument from this chapter, the resulting instance $x_{🚗}$ would be in $\mathscr{L}$ by definition. However, it would take the "RE-algorithm" $\texttt{Algo}_{\mathscr{L}}$ at least T steps before concluding that $x \in \mathscr{L}$. By constructing a series of halting Turing machine which takes an increasing amount of time to terminate, one can use the argument given in this chapter to construct a sequence of problem from L_{yes} with increasingly number of runtime to decide using the given RE-algorithm.

The instance $x_{🚗}$, after the reduction given by the proof of Theorem 2.2.4, is not being represented as a string, but rather, the output for Seq(1), where Seq is the uniform sequence being define in Pseudocode 3, which might be the reason why it takes at least time T for $\texttt{Algo}_{\mathscr{L}}$ to find a certificate against the no case. So an interesting question is whether $x_{🚗}$, when represented as a string, would still take time T for the algorithm $\texttt{Algo}_{\mathscr{L}}$ to find that $x_{🚗}$? Or $\texttt{Algo}_{\mathscr{L}}$ running with at least T steps on the input $x_{🚗}$ simply a by-product of $x_{🚗}$ having a dependency on the Turing machine 🚗?

Interestingly, in the proof of Pseudocode 3, since C_0 is a constant, computing $\texttt{Seq}_{\mathsf{D},🚗}(C_0)$ takes $O\big(\text{poly}(|🚗|)\big)$ time. This shows that the reduction from the Halting problem to D is a polynomial-time reduction. Interestingly, for the argument given above, independent of whether 🚗 halts or not, we can never observe $\texttt{Seq}_{\mathsf{D},🚗}$ halting at line 5. Thus, one might be tempted to remove line 5-7 from Pseudocode 3 on the proof above. However, without these three lines, we cannot argue that $x_{🚗} \in \mathsf{L}^{\mathsf{D}}_{\text{no}}$ whenever 🚗 does not halt. Interestingly, although x_{no} is never actually returned, it still plays a critical role in the above argument. We highlight this as an open problem: can one remove line 5 to 7, or remove the "no instances" (and retain line 5 to 7 of Pseudocode 3) and still repeat the same argument for the proof above.

As pointed out before, every compressible promise problem is also a weakly compressible problem. We show the converse of the statement above assuming that $\mathsf{D} \in \mathsf{RE}$ or $\mathsf{D} \in \mathsf{coRE}$. However, it is an interesting problem if the converse is true in general. The issue is that given a description of a uniform instance $\langle \texttt{Seq}_{\mathsf{D}} \rangle$, there is no algorithm which can determine the smallest $\alpha \in \mathbb{N}$ such that $\mathsf{TIME}_{\texttt{Seq}_{\mathsf{D}}} = O(n^{\alpha})$ (or whether $\mathsf{TIME}_{\texttt{Seq}_{\mathsf{D}}} = O(\text{poly}(n))$ at all). Hence, there is no clear way to construct a universal $\texttt{Compress}^{\mathsf{D}}$ algorithm given the $\texttt{Compress}^{\mathsf{D}}_{\alpha}$ algorithm.

[2] This is a well known classical result in computability theory.

Part II

Tracially embeddable strategies for non-local games

CHAPTER 3

Operator algebra and Quantum Information Preliminaries

Part two of the thesis focuses on tracially embeddable strategies, a class of quantum strategies for non-local games. Intuitively, this class of strategies can be used to approximate *any* quantum commuting correlation. We will establish this result, along with some of its implications for quantum correlation and quantum information in Chapter 4.

Before turning to our contributions in this area, it will be appropriate to first give general preliminaries introducing some notation that will be used for the rest of this thesis.

3.1 Sets and strings

We begin this section by introducing some notation regarding finite sets and strings we use for this thesis. For a finite set S, we use $|S|$ to denote the number of elements in S, and for $a, b \in S$, we use $\delta_{a,b}$ for the Kronecker delta between the two elements, i.e.

$$\delta_{a,b} = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{otherwise} \end{cases}.$$

Given a (potentially infinite) set T and $n \in \mathbb{N}$, we use $\mathcal{M}_n(T)$ to denote the set of n by n matrices over the set T. Given a distribution μ, we use $\mathbb{E}_{x\sim\mu}$ to denote the expectation over the distribution μ and for a set S, we use $\mathbb{E}_{x\in S}$ to denote the expectation over the set S.

For a bit string $a, s, t \in \{0,1\}^n$, we use $|s|$ to denote the Hamming weight of s and $s \cdot t$ to denote the inner product between s and t, or $\sum s_i t_i \bmod 2$. We use $s|_a \in \{0,1\}^n$ to denote the string

$$(s|_a)_i = \begin{cases} s_i & \text{if } a_i = 1 \\ 0 & \text{otherwise} \end{cases}.$$

We use $\pi_{>j}(s)$ to denote the function which zeros out the first j entries of the string s and we use $\pi_{\leq j}$ to denote the function which zeros out everything except for the first j entries of the string. In other words

$$\begin{aligned} \pi_{>j}(s_0, \cdots, s_{n-1}) &= (0, \cdots, 0, s_j, \cdots, s_{n-1}) \\ \pi_{<j}(s_0, \cdots, s_{n-1}) &= (s_0, \cdots, s_{j-1}, 0, \cdots, 0), \end{aligned} \tag{3.1}$$

and we take the convention that $\pi_{>j} = \pi_{\geq j+1}$.

3.2 Von Neumann algebras

In this section, we give some basic notation of von Neumann algebras we use for this thesis. Recall, a Hilbert space is an inner product space which is complete with respective to the norm introduced by the inner product [1]. Furthermore, we adopt the braket notation for this thesis, i.e. for a Hilbert space $\mathscr{H}$, we use $|\psi\rangle \in \mathscr{H}$ to denote a vector within the Hilbert space and $\langle\psi|$ to denote the of $|\psi\rangle$.

Let $\mathscr{H}$ be a Hilbert space, and $\mathscr{B}(\mathscr{H})$ denote the set of bounded operators on $\mathscr{H}$. Given $|\psi\rangle \in \mathscr{H}$, we use $||\psi\rangle|$ to denote the vector norm. Recall, a (concrete, unital) C^*-algebra $\mathscr{A} \subseteq \mathscr{B}(\mathscr{H})$ is a normed $*$-algebra with $\mathbb{I}_{\mathscr{H}} = \mathbb{I}_{\mathscr{A}}$ and closed in the norm topology. We use $\mathscr{A}^+$ to denote the set of positive elements within $\mathscr{A}$ (i.e. elements of the form $s^* s$ for $s \in \mathscr{A}$).

A *state* on a C^*-algebra is a linear function $\psi : \mathscr{A} \to \mathbb{C}$, which is *positive*, meaning that $\psi(a) \geq 0$ for all $a \in \mathscr{A}^+$ and satisfies $\psi(\mathbb{I}) = 1$. We use $\|\psi\|$ to denote the operator norm of ψ, or

$$\|\psi\| := \sup\{\psi(z) | z \in \mathscr{A}^+, \|z\|_{op} = 1\},$$

and we write $\|\psi\|_{\mathscr{A}}$ in order to emphasize the underlying algebra that the norm is taken over. A state ψ on $\mathscr{A}$ is said to be *faithful* if, for all $a \in \mathscr{A}^+$, we have $\psi(a) = 0$ if and only if $a = 0$. Furthermore, we say that the state is a *tracial state* if $\psi(st) = \psi(ts)$ for all $s, t \in \mathscr{A}$. The famed GNS representation theorem states that every state ψ on a C^*-algebra $\mathscr{A}$ induces a *representation* (a $*$-homomorphism to some $\mathscr{B}(\mathscr{H})$) π_ψ onto $\mathscr{B}(\mathscr{H}_\psi)$, and a unit vector $|\psi\rangle \in \mathscr{H}_\psi$ such that $\psi(z) = \langle\psi|\pi_\psi(z)|\psi\rangle$ for all $z \in \mathscr{A}$, and $\overline{\mathscr{A}|\psi\rangle} = \mathscr{H}$ (we refer to [KR97a, Theorem 4.5.2] for more details about the GNS representation). This representation is specified with the triplet $(\pi_\psi, \mathscr{H}_\psi, |\psi\rangle)$.

An element $P \in \mathscr{A}$ is a projector if $P^2 = P$. An element $V \in \mathscr{A}$ is a partial isometry if and only if VV^* and V^*V are both a projector, and unitary if furthermore $VV^* = V^*V = \mathbb{I}_{\mathscr{A}}$. For two projectors P and Q, P is said to be *equivalent* to Q if there exist an isometry V if $V^*V = P$ and $VV^* = Q$. We use $\mathscr{U}(\mathscr{A})$ to denote the set of unitary elements in $\mathscr{A}$ in this thesis. For $\mathscr{A} \subseteq \mathscr{B}(\mathscr{H})$, the *commutant* $\mathscr{A}'$ of $\mathscr{A}$ is defined to be the set of all elements which commute with $\mathscr{A}$, or $\mathscr{A}' := \{z \in \mathscr{B}(\mathscr{H}) : zw = wz \text{ for all } w \in \mathscr{A}\}$. A C^*-algebra $\mathscr{A} \subseteq \mathscr{B}(\mathscr{H})$ is said to be a *von Neumann algebra* if $\mathscr{A} = \mathscr{A}''$. By the von Neumann bicommutant theorem, an equivalent definition for von Neumann algebra $\mathscr{A}$ is for $\mathscr{A}$ to be closed in the weak $*$-topology, where recall, given $S \subseteq \mathscr{B}(\mathscr{H})$, the weak-operator topology consist of neighbourhoods of the form

$$\mathrm{N}(a, \{|\psi_i\rangle\}_{i\in[n]}, \cdots |\phi_1\rangle, \cdots |\phi_n\rangle, \varepsilon) = \{b : b \in \mathscr{B}(\mathscr{H}), |\langle\psi_i|(b-a)|\phi_i\rangle| < \varepsilon \text{ for all } i \in [n]\}$$

for $a \in S, n \in \mathbb{N}$, $|\psi_i\rangle, |\phi_i\rangle \in \mathscr{H}$ and $\varepsilon > 0$. It is well known that the weak-operator topology is a finer topology than the norm topology. Since the weak $*$-topology is more coarse than the norm topology, not every C^*-algebra is a von Neumann algebra. Unless stated otherwise, $\mathscr{A}$ is assumed to be a concrete von Neumann algebra for the remainder of this thesis.

In this thesis, we work with *separable* von Neumann algebra, meaning the underlying Hilbert space $\mathscr{H}$ is separable. By [Tak01, Proposition 3.19] and the remark after [KR97a, Proposition 13.1.4], a separable von Neumann algebras always admits a faithful normal state, where a state ψ on $\mathscr{A}$ is said to be *normal* ψ is continuous with respect to the weak operator topology.

Finally, we recall the following lemma, which states that every positive element inside a von Neumann algebra admits a left and right inverse. We remark that this statement is true even if $\mathscr{A}$ is a C^*-algebra, as the proof below only uses the norm closure property of C^*-algebras.

Lemma 3.2.1 (Existence of a left and right inverse)**.** *Let $A, B \in \mathscr{A}^+$ such that $B \leq A$, then there exists some element $R \in \mathscr{A}$ with $R^*R \leq \mathbb{I}$ and $A^{\frac{1}{2}}R = B$. Furthermore, there exists some element $L \in \mathscr{A}$ with $L^*L \leq \mathbb{I}$ and $LA^{\frac{1}{2}} = B$.*

[1] Just a good reminder for Stacey! Since she always forgets the definition of a Hilbert space in a thesis defence.

Proof. Fix $n \in \mathbb{N}$, we see that $0 < \frac{\mathbb{I}_{\mathscr{A}}}{n} \leq A + \frac{\mathbb{I}_{\mathscr{A}}}{n} \leq (\|A\|_{op} + \frac{1}{n})\mathbb{I}_{\mathscr{A}}$, which implies that $A + \frac{\mathbb{I}_{\mathscr{A}}}{n}$ is positive and invertible in $\mathscr{A}$. Let $R_n = (A + \frac{\mathbb{I}_{\mathscr{A}}}{n})^{-\frac{1}{2}} B^{\frac{1}{2}}$. We note that L_n forms a Cauchy sequence in the operator norm, and since $B \leq A \leq A + \frac{\mathbb{I}_{\mathscr{A}}}{n}$, we have $0 \leq R_n^* R_n \leq \mathbb{I}_{\mathscr{A}}$. Hence, since $\mathscr{A}$ is closed in the operator norm, the sequence R_n converges to some positive element $R \in \mathscr{A}^+$ with $R^* R \leq \mathbb{I}$, and by continuity, $A^{\frac{1}{2}} R = B^{\frac{1}{2}}$. The second part of the proof follows from considering the sequence $L_n = B^{\frac{1}{2}}(A + \frac{\mathbb{I}_{\mathscr{A}}}{n})^{-\frac{1}{2}}$. □

3.2.1 Tracial von Neumann algebras

We refer to a von Neumann algebra to be *tracial* if it admits a faithful normal tracial state τ, and we use $(\mathscr{A}, \tau)$ to emphasize the existence of τ. Whenever $\mathscr{A}$ is finite-dimensional, we use $\mathrm{Tr}(\cdot)$ to denote the trace function for clearly[2]. The faithful trace τ naturally gives the notion of a "Hilbert-Schmidt" norm on $\mathscr{A}$, defined to be

$$||A||_2 := \sqrt{\tau(A^* A)}.$$

Recall that the *standard form* for a tracial von Neumann algebra $(\mathscr{A}, \tau)$ is the GNS representation triplet $(\chi_\tau, \mathscr{L}^2(\mathscr{A}, \tau), |\tau\rangle)$ of $\mathscr{A}$ for the tracial state τ, where $\mathscr{L}^2(\mathscr{A}, \tau)$ denotes the Hilbert space for the representation. Note that the standard form for a tracial von Neumann algebra is unique up to canonical isomorphism [AP10, Proposition 7.5.1]. For simplicity of notation, if $(\mathscr{A}, \tau)$ is in standard form, for each $a \in \mathscr{A}$, we use a to denote $\chi_\tau(a)$ as the operator defined within $\mathscr{B}(\mathscr{L}^2(\mathscr{A}, \tau))$. In this representation, the vector $|\tau\rangle$ is *cyclic*, meaning that $\overline{\chi_\tau(\mathscr{A})|\tau\rangle} = \mathscr{L}^2(\mathscr{A}, \tau)$, and *separating*, meaning that for all $z \in \mathscr{A}$, we have $z|\tau\rangle = 0$ if and only if $z = 0$. This means that each $\sigma \in \mathscr{A}$ gives a unique vector $\sigma|\tau\rangle \in \mathscr{L}^2(\mathscr{A}, \tau)$, and we can specify the action of a acting on the Hilbert space $\mathscr{L}^2(\mathscr{A}, \tau)$ by its *left regular representation*:

$$a(\sigma|\tau\rangle) = (a\sigma)|\tau\rangle$$

for all $\sigma \in \mathscr{A}$.

Recall, given a von Neumann algebra $\mathscr{A}$, the *opposite algebra* $\mathscr{A}^{op} := \{a^{op} : a \in \mathscr{A}\}$ is a von Neumann algebra which has the same linearity as $\mathscr{A}$, but has the opposite multiplication structure, or more precisely $(ab)^{op} = (b)^{op}(a)^{op}$. The algebra $\mathscr{A}^{op}$ can also be faithfully embeddable onto $\mathscr{B}(\mathscr{L}^2(\mathscr{A}, \tau))$ by

$$\chi_\tau^{op}(a^{op})(\sigma|\tau\rangle) = (\sigma a)|\tau\rangle. \tag{3.2}$$

This is known as the right regular representation for $\mathscr{A}$. Clearly, $\chi_\tau^{op}(\mathscr{A}) \subseteq \mathscr{A}'$, and in fact, $\mathscr{A}' = \mathscr{A}^{op}$ [AP10, Theorem 7.1.1]. For simplicity of notation, we use a^{op} to denote $\chi_\tau^{op}(a)$ in this thesis. The map $op : a \to a^{op}$ forms a $*$-anti-isomorphism from $\mathscr{A} \to \mathscr{A}'$, meaning

$$(\lambda a + b)^{op} = \lambda a^{op} + b^{op}, \quad (ab)^{op} = b^{op} a^{op} \quad (a^*)^{op} = ((a)^{op})^*$$

for all $a, b \in \mathscr{A}$, $\lambda \in \mathbb{C}$.

3.2.2 Finite dimension example

To make the definition above more concrete, let $n \in \mathbb{N}$ and $\mathscr{A} = \mathscr{M}_n(\mathbb{C})$, we define the maximally entangled state $|\mathrm{ME}_n\rangle$ as the vector state on $\mathbb{C}^n \otimes \mathbb{C}^n$ as

$$|\mathrm{ME}_n\rangle := \frac{1}{\sqrt{n}} \sum_{i \in n} |i\rangle \otimes |i\rangle$$

[2] We remark that the notation for the trace function for a finite-dimensional matrix is the same as the trace function for a finite field. In the context of the Tr function in this thesis, we typically use lower case letter for finite field element and upper case letter for a matrix element.

We remark that this is precisely the vector state which arises from applying the GNS theorem on the normalized matrix trace $\mathrm{Tr}_n(A)$ on the algebra $\mathscr{M}_n(\mathbb{C})$, the resulting vector representation for Tr are $|\mathrm{ME}_n\rangle$. Under this representation, elements of $\mathscr{M}_n(\mathbb{C})$ gets mapped to $\mathscr{M}_n(\mathbb{C})\otimes\mathbb{I}_n$, with the commutant being $\mathbb{I}_n\otimes\mathscr{M}_n(\mathbb{C})$. For all vectors $|\psi\rangle\in\mathbb{C}^n\otimes\mathbb{C}^n$, we can always find some positive element $\sigma\in\mathscr{M}_n(\mathbb{C})\otimes\mathbb{I}_n$ such that $\sigma|\mathrm{ME}_n\rangle$, mainly, the canonical square root (the unique square root which is also positive) of the reduced density on the first register. The opposite algebra map, in this case, is the map $op:\mathscr{A}\otimes\mathbb{I}_n\to\mathbb{I}_n\otimes\mathscr{A}$ defined by $op(A\otimes\mathbb{I}_n)=\mathbb{I}_n\otimes A^T$, where A^T is the transpose of A. As a further sanity check, for all $A,B\in\mathscr{M}_n(\mathbb{C})$

$$op(A\otimes\mathbb{I}_n)(B\otimes\mathbb{I}_n)|\mathrm{ME}_n\rangle=(B\otimes A^T)|\mathrm{ME}_n\rangle=(BA\otimes\mathbb{I}_n)|\mathrm{ME}_n\rangle, \tag{3.3}$$

consistent with the definition given above. For a more comprehensive introduction on von Neumann algebra, we refer the reader to [Bla06].

3.3 Quantum measurements

Let $\mathscr{H}$ be a (potentially infinite-dimensional) Hilbert space, and let $\mathscr{B}(\mathscr{H})$ denote the set of the bounded operators on $\mathscr{H}$. If $\mathscr{A}$ is a finite set, then a *positive operator-valued measure (POVM)* on $\mathscr{H}$ with outcome set $\mathscr{A}$ is a collection of positive operators $\{A_a\}_{a\in\mathscr{A}}\subseteq\mathscr{B}(\mathscr{H})$, such that $\sum_{a\in\mathscr{A}}A_a=\mathbb{I}_{\mathscr{H}}$. A *projection-valued measure (PVM)*, or *projective measurement*, is a POVM where each of the operators A_a is a projection operator (i.e. $A_a^2=A_a$).

Let $\mathbf{f}:\mathscr{S}\to\mathscr{A}$ be a function mapping a finite set $\mathscr{S}$ to another finite set $\mathscr{A}$, and let $\{A_t\}_{t\in\mathscr{A}}$ be a POVM with measurement outcome in T. For all $s\in\mathscr{S}$, we denote

$$A_{[f|s]}:=\sum_{a:\mathbf{f}(a)=s}A_a, \tag{3.4}$$

and $A_{[f|s]}=0$ if s is not in the image for f. Intuitively, this corresponds to performing the measurement which first samples an element from A_t, and apply the map f through the measurement outcome. This is known as a "data processed measurement" in the literature. Before ending this section, we recall the orthogonalization lemma, which is used to approximate a set of POVMs on a von Neumann algebra by a PVM being defined on the same algebra.

Lemma 3.3.1 (Orthogonalization lemma, Theorem 1.2 of [dlS22a])**.** *Let $\mathscr{A}\subseteq\mathscr{B}(\mathscr{H})$ be a von Neumann algebra and let $|\psi\rangle\in\mathscr{H}$ be a unit vector. For any POVM $\{A_a\}\subseteq\mathscr{A}$ such that $\sum_a\langle\psi|A_a^2|\psi\rangle>1-\epsilon$, there exists a PVM $\{P_a\}\subseteq\mathscr{A}$ such that*

$$\sum_a\langle\psi|(A_a-P_a)^2|\psi\rangle<9\epsilon.$$

If $(\mathscr{A},\tau)$ is a tracial von Neumann algebra in standard form, we can replace $|\psi\rangle$ by $\sigma|\tau\rangle$ for some $\sigma\in\mathscr{A}$ for the lemma above in order to obtain a Hilbert-Schmidt norm approximation of the original POVM.

3.3.1 Distance between quantum measurements

In this subsection, we introduce some distance between measurements which will be useful for the analysis of non-local games. Let $\mathscr{X}$ be a finite set, μ be a probability measurement, $(\mathscr{A},\tau)$ be a tracial von Neumann algebra represented under the standard form $(\chi_\tau,\mathscr{L}^2(\mathscr{A},\tau),|\tau\rangle)$ and $|\psi\rangle\in\mathscr{L}^2(\mathscr{A},\tau)$. We say that the two sets of POVM, $\{A_a^x\}_{x\in\mathscr{X}}\subseteq\mathscr{A}$ and $\{B_a^x\}_{x\in\mathscr{X}}\subseteq\mathscr{A}'$, are δ-*consistent* with each other with respect to $|\psi\rangle$ and μ if

$$\mathop{\mathbb{E}}_{x\sim\mu}\sum_{a\neq b}\langle\psi|A_a^xB_b^x|\psi\rangle\leq O(\delta), \tag{3.5}$$

and we write

$$A_a^x \simeq_\delta B_a^x$$

if $\{A_a^x\}$ and $\{B_a^x\}$ are δ-consistent with each other and $|\psi\rangle$ and μ are clear from context.

For two sets of POVM $\{A_a^x\}_{x\in\mathcal{X}}, \{B_a^x\}_{x\in\mathcal{X}} \subseteq \mathcal{B}(\mathcal{H})$, we say that $\{A_a^x\}$ and $\{B_a^x\}$ are δ-close with each other with respect to $|\psi\rangle$ and μ if

$$\mathop{\mathbb{E}}_{x\sim\mu} \sum_a \| \left(A_a^x - B_a^x\right)|\psi\rangle \|^2 \leq O(\delta), \tag{3.6}$$

and we write

$$A_a^x \approx_\delta B_a^x$$

if $\{A_a^x\}$ and $\{B_a^x\}$ are δ-close with each other and $|\psi\rangle$ and μ are clear from context. We also use the same notation to denote distances between matrices if the superscript "x" are omitted when describing $\simeq$ or $\approx$ distances. By definition, $A_a^x \approx_\varepsilon B_a^x$ is the same as writing $\left(A_a^x - B_a^x\right) \approx_\varepsilon 0$. We remark that both measurements distance defined above are analogous to [JNV+22a, Definition 5.15, 5.16].

For three sets of POVM indexed by $x \in \mathcal{X}$, $\{A_a^x\}_{a\in\mathcal{A}}, \{B_a^x\}_{a\in\mathcal{A}}, \{C_a^x\}_{a\in\mathcal{A}} \subseteq \mathcal{B}(\mathcal{H})$, if $A_a^x \approx_\delta B_a^x$ and $B_a^x \approx_\varepsilon C_a^x$ over $|\psi\rangle \in \mathcal{H}$ and μ by the triangle inequality, this implies that $A_a^x \approx_{\delta+\varepsilon} C_a^x$ over μ and $|\psi\rangle$. Since applying a function to the measurement output cannot decrease the probability of two measurement outcome agree with each other, we get the following analogue of [NW19, Fact 4.26].

Fact 3.3.2 (Data processing)**.** *Let $\mathcal{X}$ be a finite set, $\{A_a^x\}_{a\in\mathcal{A}}$ and $\{B_b^y\}_{(a,b)\in\mathcal{A}^2}$ be two sets of POVMs, and $\boldsymbol{f}\colon \mathcal{A} \to \mathcal{B}$ be a function. Then $A_a^x \simeq_\varepsilon B_a^x$ implies that $A_{[\boldsymbol{f}|a]}^x \simeq_\varepsilon B_{[\boldsymbol{f}|a]}^x$*

We remark that the above fact does not work for the $\approx_\varepsilon$ measurement outcome and we refer to [NW19, Fact 4.26] for more details. We show the following trivial lemmas about distances of POVM measurements. The first lemma converts between closeness and distance for a pair of commuting measurement. We remark that this is an analogue of [NW19, Fact 4.13, 4.14]

Lemma 3.3.3 (Conversion between closeness and distance)**.** *Let $\mathcal{X}$ be a finite set, μ be a distribution over $\mathcal{X}$, $|\psi\rangle \in \mathcal{H}$ and $\{A_a^x\}_{a\in\mathcal{A}}$ and $\{B_b^y\}_{(a,b)\in\mathcal{A}^2}$ be two sets of POVMs in $\mathcal{B}(\mathcal{H})$ such that $A_a^x B_b^y = B_b^y A_a^x$ for all $(x,y,a,b) \in \mathcal{X}^2 \times \mathcal{A}^2$. Then the following holds:*

- *If $A_a^x \simeq_\delta B_b^y$ over μ and $|\psi\rangle$, then $A_a^x \approx_\delta B_b^y$ over μ and $|\psi\rangle$.*
- *If $A_a^x \approx_\delta B_b^y$ over μ and $|\psi\rangle$ and additionally both $\{A_a^x\}, \{B_b^y\}$ are PVMs, then $A_a^x \simeq_\delta B_b^y$ over μ and $|\psi\rangle$.*
- *If $A_a^x \approx_\delta B_b^y$ over μ and $|\psi\rangle$ and additionally either $\{A_a^x\}$ or $\{B_b^y\}$ are PVMs, then $A_a^x \simeq_{\sqrt{\delta}} B_b^y$ over μ and $|\psi\rangle$.*

Proof. By definition $A_a^x \simeq_\delta B_b^y$, we have

$$\mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\psi|A_a^x B_a^x|\psi\rangle \geq 1 - O(\delta)$$

By expanding the definition of closeness, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{x\sim\mu} \sum_a \| \left(A_a^x - B_a^x\right)|\psi\rangle \|^2 &= \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\psi|(A_a^x)^2 + (B_b^y)^2 - 2A_a^x B_a^x|\psi\rangle \\
&\leq \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\psi|A_a^x + B_b^y - 2A_a^x B_a^x|\psi\rangle \\
&\leq 2 - 2 \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\psi|A_a^x B_a^x|\psi\rangle
\end{aligned}$$

and item 1 follows accordingly. For item 2, if both $\{A_a^x\}, \{B_b^y\}$ are PVMs, then the above inequality becomes an equality and the statement follows accordingly. For item 3, without lost of generality assume that $\{A_a^x\}$ is projective, then

$$\begin{aligned}
1 - \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\psi| A_a^x B_a^x |\psi\rangle &= \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\psi| (A_a^x)^2 |\psi\rangle - \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\psi| A_a^x B_a^x |\psi\rangle \\
&= \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\psi| A_a^x \cdot (A_a^x - B_a^x) |\psi\rangle \\
&\leq \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \| A_a^x |\psi\rangle \| \cdot \| (A_a^x - B_a^x) |\psi\rangle \| \\
&\leq \sqrt{\mathop{\mathbb{E}}_{x\sim\mu} \sum_a \| A_a^x |\psi\rangle \|^2} \sqrt{\mathop{\mathbb{E}}_{x\sim\mu} \sum_a \| (A_a^x - B_a^x) |\psi\rangle \|^2}
\end{aligned}$$

where in line 2, we use the fact that $\{A_a^x\}$ is projective and a PVM, and the third line follows from Cauchy-Schwarz and the forth line follows from Jensen's inequality. Bounding the first term in line 4 by 1 completes the claim for the lemma. □

The second lemma gives a way to combine measurements while preserving distances between the measurements, we remark that this is an analogue of [NW19, Fact 4.20].

Lemma 3.3.4 (Combination of measurement preserves distance)**.** *Let $\mathcal{X}, \mathcal{A}, \mathcal{C}$ be finite sets, μ be a distribution over $\mathcal{X}^2$ with marginal distribution $\mu_X \sim \mathcal{X}$ and $\mu_Y \sim \mathcal{X}$ over the first and second coordinates respectively. For each $(x, y) \in \mathcal{X}^2$, let $\{A_{a,b}^x\}_{(a,b)\in\mathcal{A}^2}$ and $\{B_{a,b}^x\}_{(a,b)\in\mathcal{A}^2}$ be two sets of POVMs in $\mathcal{B}(\mathcal{H})$, and let $\{C_{a,c}^{x,y}\}_{(a,c)\in\mathcal{A}\times\mathcal{C}}$ be a set of POVM in $\mathcal{B}(\mathcal{H})$. If $A_{a,b}^x \approx_\delta B_{a,b}^x$ with respect to $|\psi\rangle$ and either μ_X or μ_Y, then*

$$C_{a,c}^{x,y} A_{a,b}^x \approx_\delta C_{a,c}^{x,y} B_{a,b}^x, \qquad \text{and} \qquad A_{a,b}^x C_{a,c}^{x,y} \approx_\delta B_{a,b}^x C_{a,c}^{x,y}$$

where $\approx_\delta$ is over the state $|\psi\rangle$ and the distribution $(x, y) \sim \mu$.

Proof. Since both implication follows a similar proof, we only show the first one below. Fix $(x, y) \in \mathcal{X}^2$ and $(a, b) \in \mathcal{A}^2$. By expanding the vector state, we see that

$$\begin{aligned}
\sum_c \| \left(C_{a,c}^{x,y} A_{a,b}^x - C_{a,c}^{x,y} B_{a,b}^x \right) |\psi\rangle \|^2 &= \sum_c \langle\psi| (A_{a,b}^x - B_{a,b}^x)^* (C_{a,c}^{x,y})^* C_{a,c}^{x,y} (A_{a,b}^x - B_{a,b}^x) |\psi\rangle \\
&\leq \langle\psi| (A_{a,b}^x - B_{a,b}^x)^* (A_{a,b}^x - B_{a,b}^x) |\psi\rangle \\
&= \| \left(A_{a,b}^x - B_{a,b}^x \right) |\psi\rangle \|^2
\end{aligned}$$

where the second inequality follows from C being a POVM. The lemma follows accordingly. □

3.4 Quantum correlations

In this section, we introduce different notions of quantum information that will be used in this thesis. Given two finite sets $\mathcal{X}$ and $\mathcal{A}$, a (bipartite) correlation set with question set $\mathcal{X}$ and answer set $\mathcal{A}$ is the set $\{C_{x,y,a,b}\}_{(x,y)\in\mathcal{X}^2,(a,b)\in\mathcal{A}^2} \subseteq [0,1]^{\mathcal{X}^2\times\mathcal{A}^2}$ such that

$$\sum_{(a,b)\in\mathcal{A}^2} C_{x,y,a,b} = 1$$

for all $(x, y) \in \mathcal{X}^2$. For fixed question pair $(x, y) \in \mathcal{X}^2$, $\mu(a, b) = C_{x,y,a,b}$ forms a probability distribution over $\mathcal{A}^2$. We remark that a correlation set could be defined with two different question set and two different answer set, but the formulation above is equivalent by setting some of the $C_{x,y,a,b} = 0$. In this thesis, we are primarily concerned with two sets of correlations, the quantum tensor correlations and the quantum commuting correlations, which we introduce below:

3.4.1 Quantum tensor correlations

A correlation set $\{C_{x,y,a,b}\}_{(x,y)\in\mathcal{X}^2,(a,b)\in\mathcal{A}^2}$ is a *quantum tensor correlation*, if there exist two collections of POVM, $\{A_a^x\}_{a\in\mathcal{A}} \subseteq \mathbf{M}_m(\mathbb{C})$ and $\{B_b^y\}_{b\in\mathcal{A}} \subseteq \mathbf{M}_n(\mathbb{C})$, along with an entangled state $|\psi\rangle \in \mathbb{C}^m \otimes \mathbb{C}^n$ such that

$$C_{x,y,a,b} = \langle\psi|A_a^x \otimes B_b^y|\psi\rangle$$

for all $(x,y) \in \mathcal{X}^2$ and $(a,b) \in \mathcal{A}^2$. In this case, we refer to the set

$$\mathcal{S} = (\mathbb{C}^m \otimes \mathbb{C}^n, \{A_a^x\}_{a\in\mathcal{A}}, \{B_b^y\}_{b\in\mathcal{A}}, |\psi\rangle)$$

as the *tensor product strategy* (or a $*$ strategy) which realizes the correlation $C_{x,y,a,b}$. We use $C_q(\mathcal{X},\mathcal{A})$ to denote the set of quantum tensor correlations with input set $\mathcal{X}$ and output set $\mathcal{A}$ in this thesis or simply C_q if $\mathcal{X}$ and $\mathcal{A}$ is clear from context. To abuse notation, we write C_q as C_* so that it is consistent with the non-local games notation. For an integer n, we use C_q^n to denote the set of quantum correlations achievable by a tensor product strategies with dimension n

$$C_q^n := \left\{\{\langle\psi|A_a^x \otimes B_b^y|\psi\rangle\}_{(x,y,a,b)} \,|\, |\psi\rangle \in \mathbb{C}^n \otimes \mathbb{C}^n, \{A_a^x\}, \{B_b^y\} \text{ are POVMs in } \mathbf{M}_n(\mathbb{C})\right\}.$$

For $m < n$, we have $C_q^m \subseteq C_q^n$ and furthermore

$$C_q = \bigcup_{n\in\mathbb{N}^+}^{\infty} C_q^n.$$

3.4.2 Quantum commuting correlations

A correlation $\{C_{x,y,a,b}\}_{(x,y)\in\mathcal{X}^2,(a,b)\in\mathcal{A}^2}$ is a *quantum commuting correlation* if there exist a (potentially infinite-dimensional) Hilbert space $\mathcal{H}$, two sets of POVM $\{A_a^x\}_{a\in\mathcal{A}}, \{B_b^y\}_{b\in\mathcal{A}} \subseteq \mathcal{B}(\mathcal{H})$ such that $[A_a^x, B_b^y] = A_a^x \cdot B_b^y - B_b^y \cdot A_a^x = 0$ for all $(x,y) \in \mathcal{X}^2$ and $(a,b) \in \mathcal{A}^2$, and a vector state $|\psi\rangle \in \mathcal{H}$ such that

$$C_{x,y,a,b} = \langle\psi|A_a^x C B_b^y|\psi\rangle$$

for all $(x,y) \in \mathcal{X}^2$ and $(a,b) \in \mathcal{A}^2$. In this case, we refer to the set

$$c\mathcal{S} = (\mathcal{H}, \{A_a^x\}_{a\in\mathcal{A}}, \{B_b^y\}_{b\in\mathcal{A}}, |\psi\rangle)$$

as the *commuting operator strategy* (or a qc strategy) which realizes the correlation $C_{x,y,a,b}$. We refer to a strategy (for both tensor product and commuting operator) to be a projective strategy if both the measurement operator $\{A_a^x\}$ and $\{B_b^y\}$ are PVMs.

We use $C_{qc}(\mathcal{X},\mathcal{A})$ to denote the set of quantum commuting correlations with input set $\mathcal{X}$ and output set $\mathcal{A}$ and C_{qc} if $\mathcal{X}$ and $\mathcal{A}$ are clear from context. Since any tensor product strategy is a commuting operator strategy by definition (as $[A_a^x \otimes \mathbb{I}_n, \mathbb{I}_m \otimes B_b^y]$), we have $C_q \subseteq C_{qc}$. As a consequence of the MIP* = RE theorem, we know that this inclusion is strict, or $\overline{C_q} \subsetneq C_{qc}$ [JNV+22a].

3.4.3 Synchronous correlations

In this thesis, we work with a set of correlations known as synchronous correlations. For $t \in \{*, qc\}$ and finite set $\mathcal{X}$ and $\mathcal{A}$, a correlation $\{C_{x,y,a,b}\} \in C^t(\mathcal{X},\mathcal{A})$ is synchronous iff for all $x \in \mathcal{X}$ and $(a,b) \in \mathcal{A}^2$

$$C_{x,x,a,b} = \delta_{a,b},$$

and we use C_t^s to denote the set of synchronous correlations for models t. Intuitively, a synchronous correlation corresponds to the behaviour where the experimenters will always give

the same question when given the same question. Synchronous correlations were first studied in [PSS+16], and have been used in [MNY22] to study the complexity of zero gap MIP*. We call a strategy that realizes a synchronous correlation to be a *synchronous strategy*. The set of synchronous strategy for both the tensor product model and the community operator model admits a nice structure, and we refer to Section 4.3 for more details.

In this thesis, we also consider near-synchronous correlations. One of the main contribution is the so called "rounding theorem" for the commuting operator models, which intuitively states that any near synchronous correlation is "close" (in the l_1 norm sense) to a synchronous correlation. Since the notion of a near-synchronous correlations is closely tie to the "rounding theorem", we refer to Chapter 5 for the definition when introducing the rounding theorem.

3.5 Non-local games

A *two-prover one-round (non-local) game* is described by a tuple $\mathcal{G} = (\mathcal{X}^2, \mathcal{A}^2, \mu, D)$, where $\mathcal{X}$ is a finite set denoting the list of potential questions, $\mathcal{A}$ is another finite set denoting the list of potential answers, μ is a distribution over $\mathcal{X}^2$ which corresponds to the question distribution, and $D : \mathcal{X}^2 \times \mathcal{A}^2 \to \{0,1\}$ is the evaluation map[3]. The game is played between two cooperating *provers*, Alice and Bob, and a *verifier*[4]. In this game, the verifier first samples a question pair (x, y) according to the distribution μ and sends x to Alice and y to Bob. Upon receiving their questions, Alice (and resp. Bob) must, without communicating with the other prover, respond with answers a (resp. b) in $\mathcal{A}$ back to the referee, and the provers win if and only if $D(x, y, a, b) = 1$. Conventionally, non-local games are usually expressed with provers sharing different question and answer sets. However, by forcing the probability distribution μ to be zero on certain question pair, the formulations we give are equivalent to the conventional formulation. In this thesis, we consider non-local games where the provers have access to the two models of entanglement described in the previous subsection, which gives two different *values*, or the optimal success probability for the provers, for a given game $\mathcal{G}$. We introduce these notions in the remainder of this subsection:

3.5.1 Values of a game

Under the quantum tensor product model, the provers first prepare a joint entangled quantum state $|\psi\rangle \in \mathbb{C}^n \otimes \mathbb{C}^m$ between them. After receiving the question from the verifier, the provers then perform localized measurements on their respective register based on the question they receive. The provers then respond to the verifier with the measurement output as their answers. In this case, the behaviour of the provers precisely describes a tensor product strategy defined in the previous subsection, and the probability of outputting the answer pair (a, b) given the question pair (x, y) is $C_{x,y,a,b}$, where $\{C_{x,y,a,b}\}_{x,y,a,b}$ is the correlation generated by the said strategy.

Given a quantum tensor correlation $C = \{C_{x,y,a,b}\} \in C_q(\mathcal{X}, \mathcal{A})$, we define the success rate for the correlation, or the *value* of the correlation to be

$$\omega(\mathcal{G}, C) := \sum_{(x,y)\in\mathcal{X}^2} \mu(x, y) \sum_{(a,b)\in\mathcal{A}} D(x, y, a, b) C_{x,y,a,b}. \tag{3.7}$$

Similarly, for a tensor product strategy $\mathcal{S}$, we use $\omega(\mathcal{G}, \mathcal{S})$ to denote the value of the correlation realized by the strategy $\mathcal{S}$. The optimal success rate given quantum tensor model of

[3] We remark that this is slightly different from how we define the evaluation map in Chapter 1, in this case 0 maps to "loses" and 1 maps to "win".

[4] We are adopting the notation from an interactive proof setting in this thesis. In other non-local games literature, the provers might be referred to as "players" and the verifier might be referred to as the "referee"

entanglement, or the *tensor product value* of a game $\mathcal{G}$ is

$$\omega^*(\mathcal{G}) := \sup_{\{C_{x,y,a,b} \in C_q\}} \omega(\mathcal{G}, C). \tag{3.8}$$

We remark that since C_q is not a closed set [Slo19b], the supremum in the above equation might not be realizable by a tensor product correlation.

Similarly, if the provers are allowed to use the commuting model of entanglement. The set up is similar as above, except the provers use commuting operator strategies instead. Given a quantum commuting operator correlation $C = \{C_{x,y,a,b}\} \in C_q(\mathcal{X}, \mathcal{A})$, the value for the correlation is the same as (3.7). The optimal success rate given quantum commuting model of entanglement, or the *commuting operator value* of a game $\mathcal{G}$ is

$$\omega^{co}(\mathcal{G}) := \sup_{\{C_{x,y,a,b} \in C_{qc}\}} \omega(\mathcal{G}, C). \tag{3.9}$$

For model $t \in \{*, co\}$, we call a strategy $\mathcal{S}$ in model t a perfect strategy for game $\mathcal{G}$ if $\omega(\mathcal{G}, \mathcal{S}) = 1$. While the value of a game can be defined in terms of either quantum correlations or quantum strategies, there is a distinction between correlations and strategies. From the verifier's point of view, he can only "detect" the correlation by sampling a pair of questions and getting a pair of response from the provers. However, the provers can choose different strategies (which the verifier cannot detect) in order to realize the same correlation.

3.5.2 Synchronous games

Given a game $\mathcal{G}$, we call a game *synchronous* iff $D(x, x, a, b) = \delta_{a,b}$. In other words, the provers must provide the same answer pair when given the same question pair. For a synchronous game $\mathcal{G}$, we call a question pair (x, y) to be synchronous iff $x = y$. For model $t \in \{*, qc\}$ and a synchronous game $\mathcal{G}$, we define the synchronous value for $\mathcal{G}$ in model t to be

$$\omega_s^t(\mathcal{G}) := \sup_{\{C_{x,y,a,b} \in C_t^s\}} \omega(\mathcal{G}, C).$$

Intuitively, a synchronous strategy corresponds to the set of strategies in which the provers always give the same answer when given the same questions. Since $C_q^s \subseteq C_q$ (resp. $C_{qc}^s \subseteq C_{qc}$), we have $\omega_s^*(\mathcal{G}) \leq \omega^*(\mathcal{G})$ (resp. $\omega_s^{co}(\mathcal{G}) \leq \omega^{co}(\mathcal{G})$). However, as seen by the following theorem, these two values are equivalent whenever $\mathcal{G}$ admits a perfect strategy.

Theorem 3.5.1 (Perfect quantum value implies perfect synchronous value, Theorem 3.2 of [MNY22])**.** *Let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$ be a synchronous game such that $\mu(x, x) > 0$ for all $x \in \mathcal{X}$. For model $t \in \{0, 1\}$, $\omega^t(\mathcal{G}) = 1 \to \omega_s^t(\mathcal{G}) = 1$.*

As seen in Section 5.3, for a specific class of synchronous games, the above theorem is true even if the optimal value of a game is not perfect.

CHAPTER 4

Tracially embeddable, symmetric and oracularizable strategies

In this chapter, we first introduce a special class of commuting operator strategies, tracially embeddable strategies. Intuitively, one should interpret this as a class of strategies with many of the properties of finite-dimensional strategies. One of the key observations that we made is that the correlation set from these strategies can be used to approximate *all* commuting operator correlations! Using this definition, we introduce two classes of strategies, symmetric strategies and oracularizable strategies in Section 4.3 and Section 4.4 respectively, and prove some useful lemmas about these strategies. Both classes of strategies play an important role in proving the existence of a compression map for $\mathsf{MIP}^*/\mathsf{MIP}^{\mathrm{co}}$.

4.1 Tracially embeddable strategies

We start this chapter by introducing a special class of commuting operator strategies known as *tracially embeddable strategies,* motivated by the standard form of a tracial von Neumann algebra. We give the definition below.

Definition 4.1.1 (Tracially embeddable strategy)**.** *Let $\mathcal{G} = (\mathcal{X}^2, \mathcal{A}^2, \mu, D)$ be a non-local game. A quantum commuting strategy $\mathcal{S} = (\mathcal{H}, |\psi\rangle, \{A_a^x\}_{a\in\mathcal{A}}, \{B_b^y\}_{b\in\mathcal{A}})$ for game $\mathcal{G}$ is called **tracially embeddable** if there exists a tracial von Neumann algebra $(\mathcal{A}, \tau)$ with standard form $(\pi_\tau, \mathcal{L}^2(\mathcal{A}, \tau), |\tau\rangle)$ and $\sigma \in \mathcal{A}^+$ such that $\mathcal{H} = \mathcal{L}^2(\mathcal{A}, \tau)$, $|\psi\rangle = \sigma|\tau\rangle$, $\{A_a^x\}_{a\in\mathcal{A}} \subseteq \mathcal{A}$ and $\{B_b^y\}_{b\in\mathcal{A}} \subseteq \mathcal{A}'$.*

In this thesis, we use $\mathcal{S} = (\mathcal{L}^2(\mathcal{A}, \tau), \sigma|\tau\rangle, \{A_a^x\}, \{(B_b^y)^{op}\})$ to denote a tracially embeddable strategy on the tracial von Neumann algebra $(\mathcal{A}, \tau)$[1]. Furthermore, we call the set of possible correlations generated by a tracially embeddable strategy as *tracially embeddable correlations.* Similarly, we use $\mathcal{C}_{qc}^{Tr}(\mathcal{X}, \mathcal{A}) \subseteq \mathcal{C}_{qc}(\mathcal{X}, \mathcal{A})$ to denote the set of tracially embeddable correlations with input set $\mathcal{X}$ and output set $\mathcal{A}$. Intuitively, tracially embeddable strategies have a similar structure as finite-dimensional tensor product strategies. Let $\mathcal{S} = (\mathcal{L}^2(\mathcal{A}, \tau), \sigma|\tau\rangle, \{A_a^x\}, \{(B_b^y)^{op}\})$ be a tracially embeddable strategy on the tracial von Neumann algebra $(\mathcal{A}, \tau)$. We can think of $\mathcal{A}$(resp. $\mathcal{A}'$) to be Alice's (resp. Bob's) algebra, similar to the local Hilbert space structure in the tensor product model. Since $1 = \langle\tau|\sigma^2|\tau\rangle = \|\sigma\|_2^2$ and $\sigma \in \mathcal{A}^+$, we can intuitively think of σ^2 as the reduced density matrix on Alice's side. Furthermore, we can use the opposite algebra map introduced in the previous section to perform "observable switching", or to move Bob's measurement operator to Alice's algebra using the vector state $|\tau\rangle$. To make the discussion more concrete, we consider the following example:

[1] Note that the POVM $\{B_b^y\}$ is defined in $\mathcal{A}'$ under this formulation.

Example 4.1.2. Let $\mathscr{A} = \mathbf{M}_n(\mathbb{C})$, recall, the GNS representation using the trace $\mathrm{Tr}(\cdot)$ maps $\mathscr{A}$ to $\mathscr{A} \otimes \mathbb{I}_n \in \mathbf{M}_{n^2}(\mathbb{C})$, where the state $\mathrm{Tr}(\cdot)$ is mapped to the maximally entangled state $|\tau\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} |ii\rangle$. As a sanity check, we see that $\mathrm{Tr}(A) = \langle \tau | (A \otimes \mathbb{I}_n) | \tau \rangle$ for all $A \in \mathbf{M}_n(\mathbb{C})$. We note that $|\tau\rangle$ is a tracial state for $\mathbf{M}_n(\mathbb{C}) \otimes \mathbb{I}_n$ by definition, and for all $|\psi\rangle \in \mathbb{C}^{n^2}$, there exists some $\sigma \in \mathbf{M}_n(\mathbb{C})$ such that $(\sigma \otimes \mathbb{I}_n)|\tau\rangle = |\psi\rangle$, showing that $|\tau\rangle$ is also cyclic. In this case, σ^2 is also the reduced density matrix for $|\psi\rangle$ on the first register. For all $\sigma \otimes \mathbb{I}_n \in \mathbf{M}_n(\mathbb{C})$, the operator $\mathbb{I}_n \otimes \mathscr{A}^T \in \mathscr{A}'$ maps

$$(\mathbb{I}_n \otimes A^T)\left((\sigma \otimes \mathbb{I}_n)|\tau\rangle\right) = (\sigma A \otimes \mathbb{I}_n)|\tau\rangle .$$

Hence, we can define the opposite algebra in a similar fashion as Equation (3.2), where

$$(A \otimes \mathbb{I}_n)^{op} = \mathbb{I}_n \otimes \mathscr{A}^T.$$

For a non-local game $\mathscr{G} = (\mathscr{X}^2, \mathscr{A}^2, \mu, D)$ and a finite dimensional strategy $(\{A_a^x \otimes \mathbb{I}_n\}, \{\mathbb{I}_n \otimes B_b^y\}, |\psi\rangle)$ defined on $\mathbb{C}^n \otimes \mathbb{C}^n$. The correlation generated by $(\{A_a^x \otimes \mathbb{I}_n\}, \{\mathbb{I}_n \otimes B_b^y\}, |\psi\rangle)$ can be rewritten as

$$C_{x,y,a,b} = \langle \tau | (\sigma \otimes \mathbb{I}_n)(A_a^x \otimes \mathbb{I}_n)\left((B_b^y)^T \otimes \mathbb{I}_n\right)^{op} (\sigma \otimes \mathbb{I}_n)|\tau\rangle ,$$

for some $\sigma \in \mathbf{M}_n(\mathbb{C})$. Although σ does not necessarily have to be positive, by the polar decomposition of $\sigma = \sigma U$ for some unitary element U, by replacing Bob's observable by $(U^T)^* B_b^y U^T$, we can, without loss of generality, assume that σ is positive regardless.

Finally, for audiences with no prior background in operator algebra, the reference chart given in Table 4.1 may be helpful when reading some of the proofs in this thesis.

	TES	FD strategy (over $\mathbb{C}^n \otimes \mathbb{C}^n$)
Algebra	$\mathscr{A}$	$\mathscr{M}_n(\mathbb{C}) \otimes \mathbb{I}_n$
Commutant	$\mathscr{A}'$	$\mathbb{I}_n \otimes \mathscr{M}_n(\mathbb{C})$
Hilbert space ($\mathscr{H}$)	$\mathscr{L}^2(\mathscr{A}, \tau)$	$\mathbb{C}^n \otimes \mathbb{C}^n$
Tracial state	$\|\tau\rangle$	$\|\mathrm{ME}_n\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n} \|ii\rangle$
Reduced density matrices	σ^2	$\mathrm{Tr}_B(\|\psi\rangle\langle\psi\|)$
POVMs for prover 1 (Alice)	A_a^x	$A_a^x \otimes \mathbb{I}_n$
POVMs prover 2 (Bob)	B_b^y	$\mathbb{I}_n \otimes B_b^y$
Observable switching trick	$A\|\tau\rangle = A^{op}\|\tau\rangle$	$A \otimes \mathbb{I}_n \|\mathrm{ME}_n\rangle = \mathbb{I}_n \otimes A^T \|\mathrm{ME}_n\rangle$

Table 4.1: A diagram translating components of a tracially embeddable strategy to its finite-dimensional counterpart. We assume the finite-dimensional strategy is defined over registers A and B. Where TES above stands for Tracially embeddable strategies.

As one of the main results for this thesis, we see that every correlation set C_{qc} can be approximated by a sequence of tracially embeddable strategies through the following theorem.

Theorem 4.1.3. *Let $\mathscr{X}$ and $\mathscr{A}$ be arbitrary finite sets. Then*

$$\overline{\mathscr{C}_{qc}^{Tr}(\mathscr{X}, \mathscr{A})} = \mathscr{C}_{qc}(\mathscr{X}, \mathscr{A}),$$

where the closure above is in the l_1-norm sense.

We give a proof for Theorem 4.1.3 in the next section. We remark that this is a surprising result, as shown in our (well, my) masters thesis years ago, I personally would expect the above theorem to be false (see [Lin22] for the embarrassing call I made back in the years).

4.2 Approximation of tracially embeddable strategies

In this section, we give a proof for Theorem 4.1.3. The proof relies on Haagerup's reduction theorem [HJX10] which, roughly speaking, allows one to approximate any von Neumann algebra by a sequence of tracial von Neumann algebras. Intuitively, since every commuting operator strategy can be realized in a von Neumann algebra, one can use this approximation theorem to define a series of tracially embeddable strategies which can be used to approximate the original strategy.

We start by introducing additional tools required to prove the main theorem in this subsection. Since introducing Haagerup's reduction theorem requires introducing deeper theories in von Neumann algebra, we choose to include these theories below. We remark that understanding the rest of the thesis, which is dedicated to potential applications of the main theorem, does not depend on understanding the technical details from this subsection.

4.2.1 Additional theorems about von Neumann algebras

Recall, for a set S, the discrete topology is generated by open sets of the form $\{s\}$ for $s \in S$. Let $\mathfrak{G}$ be a countable subgroup of $(\mathbb{R},+)$ equipped with the discrete topology. Let $\boldsymbol{\alpha}$ be a *continuous automorphic representation* of $\mathfrak{G}$ on $\mathscr{A}$, or a continuous map $\boldsymbol{\alpha}: \mathfrak{G} \to \boldsymbol{\alpha}_g$ where each α_g is a $*$-automorphism acting on $\mathscr{A}$ indexed by $g \in \mathfrak{G}$. Let $\ell^2(\mathfrak{G},\mathbb{C})$ denote the set of square integrable functions from $\mathfrak{G}$ to $\mathbb{C}$. Since $\mathscr{G}$ is a subgroup of $\mathbb{R}$, for any function $\psi \in \ell^2(\mathfrak{G},\mathbb{C})$, we have $\sum_{g\in\mathfrak{G}} \|\psi(g)\|^2 < \infty$. This means that any functions in $\ell^2(\mathfrak{G},\mathbb{C})$ can be represented as elements in the Hilbert space $\oplus_{g\in\mathfrak{G}}\mathbb{C}_g$ by $\psi \to \oplus_{g\in\mathfrak{G}}\psi(g)$.

We say a von Neumann algebra $\mathscr{A} \subseteq \mathscr{B}(\mathscr{H})$ is in standard form if it admits a cyclic and separating vector in $\mathscr{H}$ for the algebra $\mathscr{A}$ (see [Ara74] for more details). We remark that in the case where $\mathscr{A}$ is tracial, $\mathscr{A}$ being in standard form is equivalent to $\mathscr{A}$ represented under the GNS representation of the trace function, which is consistent with the definition given in Section 3.2.1.

In order to define the modular automorphism group, we need to define the well known Tomita-Takesaki construction [Tak70]. The Tomita-Takesaki construction will be used more extensively later on in this thesis in proving a parallel repetition theorem for the commuting operator model (in particular, the modular operator will be an important component for defining a notion of relative entropy between two states in the commuting operator model). Hence, we list some properties necessary for the proof of Theorem 4.1.3 in this section, and give a more detailed explanation and a finite dimensional example in Section 12.1.1.

Recall from [Tak70] that any von Neumann algebra in standard form admits a modular operator Δ and a modular conjugation operator $\mathfrak{J}$ acting on $\mathscr{H}$ such that

$$\Delta^{it}\mathscr{A}\Delta^{-it} = \mathscr{A}, \quad \mathfrak{J}\mathscr{A}\mathfrak{J} = \mathscr{A}', \quad t \in \mathbb{R}.$$

The above relation can be used to define a *modular automorphism group* for any subgroup $\mathfrak{G}$ of $\mathbb{R}$. More specifically, let $\mathrm{Aut}(\mathscr{A})$ define the set of automorphisms on $\mathscr{A}$. For each $t \in \mathfrak{G}$, we define the function $\boldsymbol{\sigma}_t \in \mathrm{Aut}(\mathscr{A})$ as $\boldsymbol{\sigma}_t(A) = \Delta^{it}A\Delta^{-it}$. The modular automorphism group is defined to be the continuous map $\boldsymbol{\sigma}: \mathfrak{G} \to \mathrm{Aut}(\mathscr{A})$ given by

$$\boldsymbol{\sigma}(t) = \boldsymbol{\sigma}_t,$$

for all $t \in \mathfrak{G}$. We remark that since $\mathfrak{G}$ is assumed to be equipped with the discrete topology in this thesis, $\boldsymbol{\sigma}$ is continuous by default.

Recall from the literature that the *crossed product* of $\mathscr{A}$ by $\mathfrak{G}$ with respect to $\boldsymbol{\alpha}$, denoted by $\mathscr{R} = \mathscr{A} \rtimes_{\alpha} \mathfrak{G}$ is a von Neumann algebra acting on $\ell^2(\mathfrak{G},\mathscr{H})$ generated by $\{\boldsymbol{\pi}_{\boldsymbol{\alpha}}(A)\}_{A\in\mathscr{A}}$ and $\{\boldsymbol{\lambda}_{\boldsymbol{\alpha}}(h)\}_{h\in\mathfrak{G}}$, where

$$(\boldsymbol{\pi}_{\boldsymbol{\alpha}}(A)\zeta)(g) = \boldsymbol{\sigma}_g^{-1}(A)\zeta(g) \quad (\boldsymbol{\lambda}_{\boldsymbol{\alpha}}(h)\zeta)(g) = \zeta(g-h), \quad \text{for} \quad \zeta \in \ell^2(\mathfrak{G},\mathscr{H}), \quad g \in \mathfrak{G}.$$

Since we assume $\mathfrak{G}$ to be a countable subgroup of $\mathbb{R}$ in this thesis, if $\mathscr{A}$ is a separable von Neumann algebra, then the von Neumann algebra $\mathscr{A} \rtimes_{\boldsymbol{\alpha}} \mathfrak{G}$ is also separable. Under this assumption, the generator for $\mathscr{A} \rtimes_{\boldsymbol{\sigma}} \mathfrak{G}$ can be specified by the operator $\{\boldsymbol{\pi}_{\boldsymbol{\sigma}}(A)\}_{A\in\mathscr{A}}$ and $\{\boldsymbol{\lambda}_{\boldsymbol{\sigma}}(h)\}_{h\in\mathfrak{G}}$, where

$$\boldsymbol{\pi}_{\boldsymbol{\sigma}}(A) = \sum_{t\in\mathfrak{G}} \Delta^{it} A \Delta^{-it} \otimes |t\rangle\langle t|, \quad \boldsymbol{\lambda}_{\boldsymbol{\sigma}}(h) = \mathbb{I}_{\mathscr{A}} \otimes \sum_{t\in\mathfrak{G}} |t-h\rangle\langle t|,$$

for $A \in \mathscr{A}$ and $h \in \mathfrak{G}$. In this case, the map $\boldsymbol{\theta} : \mathscr{A} \to \mathscr{A} \rtimes_{\boldsymbol{\sigma}} \mathfrak{G}$ given by

$$\boldsymbol{\theta}(A) = \boldsymbol{\pi}_{\boldsymbol{\sigma}}(A) \tag{4.1}$$

gives a $*$-isomorphism from $\mathscr{A}$ to its crossed product construction with $\mathfrak{G}$. The following proposition extends a normal state from $\mathscr{A}$ to $\mathscr{A} \rtimes_{\boldsymbol{\sigma}} \mathfrak{G}$.

Proposition 4.2.1 (Proposition 13.1.4 of [KR97b])**.** *Let $\mathfrak{G}$ be a subgroup of $\mathbb{R}$, $\mathscr{A}$ be a von Neumann algebra and $\boldsymbol{\sigma}$ be the modular automorphism group over $\mathfrak{G}$. For every normal state ρ acting on $\mathscr{A}$, there is a normal state ω acting on $\mathscr{A} \rtimes_{\boldsymbol{\sigma}} \mathfrak{G}$ such that*

$$\omega(A) = \rho\left((\mathbb{I}_{\mathscr{A}} \otimes \langle 0_{\mathfrak{G}}|) A (\mathbb{I}_{\mathscr{A}} \otimes |0_{\mathfrak{G}}\rangle)\right) \tag{4.2}$$

for all $A \in \mathscr{A} \rtimes_{\boldsymbol{\sigma}} \mathfrak{G}$, where $\boldsymbol{\theta}$ is the isomorphic map defined in (4.1), and $0_{\mathfrak{G}}$ should be interpreted as the identity of the group $(\mathfrak{G}, +)$.

Since $\Delta^{i0} A \Delta^{-i0} = A$ for all $A \in \mathscr{A}$, this implies that $(\mathbb{I}_{\mathscr{A}} \otimes \langle 0_{\mathfrak{G}}|)\boldsymbol{\theta}(A)(\mathbb{I}_{\mathscr{A}} \otimes |0_{\mathfrak{G}}\rangle) = A$ for all $A \in \mathscr{A}$. This means that we can replace (4.2) with

$$\omega(A) = \rho\left((\mathbb{I}_{\mathscr{A}} \otimes \langle 0_{\mathfrak{G}}|)\boldsymbol{\theta}(A)(\mathbb{I}_{\mathscr{A}} \otimes |0_{\mathfrak{G}}\rangle)\right).$$

In other words, the normal state given by Proposition 4.2.1 also preserves the action given by $\boldsymbol{\theta}$.

For a von Neumann subalgebra $\mathscr{B} \subseteq \mathscr{A}$, we say a completely positive contraction map $\boldsymbol{\Phi} : \mathscr{A} \to \mathscr{B}$ is a *conditional expectation* from $\mathscr{A}$ to $\mathscr{B}$ if for all $X \in \mathscr{A}$, $A, B \in \mathscr{A}$, we have

$$\boldsymbol{\Phi}(AXB) = A\boldsymbol{\Phi}(X)B.$$

Furthermore, a conditional expectation is *normal* if it preserves the weak topology. We recall the following reduction theorem below, which states that any separable von Neumann algebra under the crossed product of $\mathfrak{G} = \bigcup_{n\geq 1} 2^{-n}\cdot\mathbb{Z} \subseteq \mathbb{R}$ with respect to the modular automorphism group can be approximated by a sequence of tracial von Neumann algebras.

Theorem 4.2.2 (Haagerup's reduction, Theorem 2.1 of [HJX10])**.** *Let $\mathscr{A}$ be a separable von Neumann algebra in standard form, let $\mathfrak{G} = \bigcup_{n\geq 1} 2^{-n}\cdot\mathbb{Z} \subseteq \mathbb{R}$ and let $\boldsymbol{\sigma}$ be the corresponding modular automorphism group of $\mathfrak{G}$. There exists a sequence $(\mathscr{R}_n, \tau_n)_{n\geq 1}$ of subalgebras of $\mathscr{R} = \mathscr{A} \rtimes_{\boldsymbol{\sigma}} \mathfrak{G}$ such that*

- *$\mathscr{R}_n \subseteq \mathscr{R}_{n+1}$ for all $n \in \mathbb{N}$.*
- *Each $\mathscr{R}_n$ is a tracial von Neumann algebra with the normal trace being τ_n.*
- *For each n, there exists a normal faithful conditional expectation $\boldsymbol{\Phi}_n$ from $\mathscr{R} \to \mathscr{R}_n$ with $\boldsymbol{\Phi}_n(A) + \boldsymbol{\Phi}_n(B) = \boldsymbol{\Phi}_n(A+B)$;*
- *$\bigcup_{n\geq 1} \mathscr{R}_n$ is w^* dense within $\mathscr{R}$.*

We remark that Theorem 4.2.2 actually holds for σ-finite von Neumann algebras according to the original formulation. However, as seen in the remark after [KR97a, Definition 5.5.14], all separable von Neumann algebra are also σ-finite. Finally, we recall two theorems from the literature about tracial von Neumann algebras in standard form which we use in the proof of Theorem 4.1.3. The first is an approximation theorem about normal linear functionals for tracial von Neumann algebra in standard form.

Proposition 4.2.3 (Proposition 7.3.4 and Theorem 7.3.8 of [AP10])**.** *For every positive normal linear functional ψ on a tracial von Neumann algebra $\mathscr{A} \subseteq \mathscr{B}(\mathscr{H})$ in standard form, there exists some vector $|\psi\rangle \in \mathscr{H}$ with $\psi(A) = \langle\psi|A|\psi\rangle$. Furthermore, there exists a sequence of positive operators $A_n \in \mathscr{A}^+$ such that*

$$\lim_{n\to\infty} ||\,|\psi\rangle - A_n\,|\tau\rangle\,|| = 0$$

Recall, given two linear functionals ψ_1 and ψ_2 acting on a C^*-algebra $\mathscr{A}$, we say $\psi_1 \leq \psi_2$ iff $\psi_2 - \psi_1$ is positive. The second theorem is a variant of the Radon-Nikodým theorem.

Theorem 4.2.4 (Proposition 7.3.5. of [KR97a])**.** *Let $\mathscr{A} \subseteq \mathscr{B}(\mathscr{H})$ be a von Neumann algebra, and let ω be a positive linear functional acting on $\mathscr{A}$. Furthermore, there exists a vector state $|\psi\rangle$ such that $\omega \leq \omega_{|\psi\rangle}$, where $\omega_{|\psi\rangle}$ is the linear functional acting on $\mathscr{A}$ such that $\omega_{|\psi\rangle}(A) = \langle\psi|A|\psi\rangle$. Then there is a positive operator $H^{'}$ in the unit ball $(\mathscr{A}^{'})_1$ of $\mathscr{A}^{'}$ such that*

$$\omega(A) = \langle\psi|H^{'}A|\psi\rangle$$

for all $A \in \mathscr{A}$.

4.2.2 The proof

We prove Theorem 4.1.3 below.

Proof. Let $C_{x,y,a,b} \in C_{qc}$ be an arbitrary commuting operator correlation, and let

$$\mathscr{S} = (\{\widetilde{A}_a^x\}, \{\widetilde{B}_b^y\}, |\widetilde{\psi}\rangle, \mathscr{H})$$

be the commuting operator algebra strategy that realizes $C_{x,y,a,b}$. We first show that $\{\widetilde{A}_a^x\}$ can be realized by a separable algebra. By Fritz's characterization [Fri12, Proposition 3.4] of commuting operator correlation, since both $\mathscr{X}$ and $\mathscr{A}$ are finite sets, we can, without loss of generality, assume $\mathscr{H}$ is separable. Let $\mathscr{N}$ be the von Neumann algebra generated by $\{\widetilde{A}_a^x\}$, since $\mathscr{H}$ is separable, $\mathscr{N}$ is also separable by definition. $\mathscr{N}$ being separable implies that $\mathscr{N}$ admits a faithful normal state, and this gives a σ-weakly continuous isomorphism which maps $\mathscr{N}$ to its standard form.

We first apply the cross product construction to $\mathscr{N}$ in order to apply Theorem 4.2.2. Let $\mathfrak{G} = \cup_{n\geq 1} 2^{-n}\cdot\mathbb{Z} \subseteq \mathbb{R}$ as stated in Theorem 4.2.2, let $\boldsymbol{\sigma}$ be the corresponding modular automorphism group for $\mathscr{N}$ and let $\mathscr{R} = \mathscr{N} \rtimes_{\boldsymbol{\sigma}} \mathfrak{G}$. For each $(y,b) \in \mathscr{X}\times\mathscr{A}$, let $\widetilde{\psi}_b^y : \mathscr{N} \to \mathbb{C}$ be the normal linear functional given by $\widetilde{\psi}_b^y(A) = \langle\widetilde{\psi}|A\cdot\widetilde{B}_b^y|\widetilde{\psi}\rangle$ and let $\widetilde{\psi} : \mathscr{N} \to \mathbb{C}$ be the normal linear functional defined by $\widetilde{\psi}(A) = \langle\widetilde{\psi}|A|\widetilde{\psi}\rangle$. Since each $\{\widetilde{B}_b^y\}_{b\in\mathscr{A}}$ is a set of POVMs, we have $\sum_b \widetilde{\psi}_b^y = \widetilde{\psi}$ for all $y \in \mathscr{A}$. By Proposition 4.2.1, each $\widetilde{\psi}_b^y$ (resp. $\widetilde{\psi}$) extends to a normal state $\widetilde{\psi}_b^y$ (resp. $\widetilde{\psi}$) on $\mathscr{R}$ such that $\widetilde{\psi}_b^y(A) = \widetilde{\psi}_b^y(\boldsymbol{\theta}(A))$ for all $A \in \mathscr{N}$, where the map $\boldsymbol{\theta}$ is the map defined in (4.1).

Let $(\mathscr{R}_n, \tau_n)_{n\geq 1}$ be the sequence of tracial von Neumann sub-algebras of $\mathscr{R}$ guaranteed by Theorem 4.2.2. We wish to construct a sequence of strategies $C_{x,y,a,b}^n \in \overline{\mathscr{C}_{qc}^{Tr}(\mathscr{X},\mathscr{A})}$ such that

- For each $n \geq 1$, each $C_{x,y,a,b}^n$ can be approximated by a series of tracially embeddable strategies defined in the standard form of $(\mathscr{R}_n, \tau_n)$,
- The limit of $C_{x,y,a,b}^n$ approaches to $C_{x,y,a,b}$.

For each $(y,b) \in \mathscr{X}\times\mathscr{A}$ and $n \in \mathbb{N}$, define $\{\widehat{C}_{x,y,a,b}^n\}_{(a,x)\in\mathscr{X}\times\mathscr{A}}$ to be

$$\widehat{C}_{x,y,a,b}^n = \widetilde{\psi}_b^y \circ \boldsymbol{\Phi}_n \circ \boldsymbol{\theta}(\widetilde{A}_a^x). \tag{4.3}$$

We see that

$$\begin{aligned}\sum_{a,b}\widehat{C}^n_{x,y,a,b} &= \sum_{a,b}\widetilde{\psi}^y_b\circ\mathbf{\Phi}_n\circ\boldsymbol{\theta}(\widetilde{A}^x_a)\\ &= \sum_b \widetilde{\psi}^y_b\circ\mathbf{\Phi}_n\circ\boldsymbol{\theta}(\sum_a \widetilde{A}^x_a)\\ &= \widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})\end{aligned}$$

where the second line follows from the functional $\widetilde{\psi}^y_b$, $\mathbf{\Phi}_n$ and $\boldsymbol{\theta}$ being linear. By property 1 and 4 of Theorem 4.2.2, the sequence $\{\mathbf{\Phi}_n\circ\boldsymbol{\theta}(X)\}_{n\in\mathbb{N}}$ is an increasing sequence of operators within $\mathscr{R}$ for all $X\in\mathscr{R}$ with $\lim_{n\to\infty}\mathbf{\Phi}_n\circ\boldsymbol{\theta}(X)=\boldsymbol{\theta}(X)$. In particular, we have

$$\lim_n \widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}}) = \widetilde{\psi}(\mathbb{1}_{\mathscr{N}}) = 1.$$

Let

$$C^n_{x,y,a,b} = \frac{1}{\widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})}\widehat{C}^n_{x,y,a,b}.$$

Since $\sum_{a,b}C^n_{x,y,a,b}=1$ for all $(x,y)\in\mathscr{X}^2$ and $n\in\mathbb{N}$, this shows that $\{C^n_{x,y,a,b}\}$ is a correlation set for each $n\in\mathbb{N}$.

We first show $\lim_{n\to\infty}C^n_{x,y,a,b}=C_{x,y,a,b}$. Since for each $(x,y,a,b)\in\mathscr{X}^2\times\mathscr{A}^2$, the linear functional ψ^a_x is normal, we have

$$\lim_n C^n_{x,y,a,b} = \frac{\lim_n \widetilde{\psi}^a_x\circ\mathbf{\Phi}_n\circ\boldsymbol{\theta}(\widetilde{A}^x_a)}{\lim_n\widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})} = \widetilde{\psi}^a_x\circ\boldsymbol{\theta}(\widetilde{A}^x_a) = \langle\widetilde{\psi}|\widetilde{A}^x_a\cdot\widetilde{B}^y_b|\widetilde{\psi}\rangle = C_{x,y,a,b},$$

hence showing the sequence of correlations $C^n_{x,y,a,b}$ approximates $C_{x,y,a,b}$.

Now, we wish to show that each of the $C^n_{x,y,a,b}$ can be approximated by a sequence of tracially embeddable strategies defined within $(\mathscr{R}_n,\tau_n)$ to complete the proof. Let $(\pi_{\tau_n},\mathscr{L}^2(\mathscr{R}_n,\tau_n),|\tau_n\rangle)$ be the standard form for $(\mathscr{R}_n,\tau_n)$. Fix $n\in\mathbb{N}$, since Φ_n is normal, $\frac{1}{\widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})}\widetilde{\psi}\circ\mathbf{\Phi}_n$ is a normal state on $\mathscr{R}$ (and hence for $\mathscr{R}_n\subseteq\mathscr{R}$). By [Ara74, Theorem 6], there exists some vector state $|\psi_n\rangle\in\mathscr{L}^2(\mathscr{R}_n,\tau_n)$ such that for all $X\in\mathscr{R}_n$,

$$\frac{\widetilde{\psi}_n\circ\mathbf{\Phi}_n(X)}{\widetilde{\psi}_n\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})} = \langle\psi_n|X|\psi_n\rangle.$$

Let $A^{x,n}_a=\mathbf{\Phi}_n\circ\boldsymbol{\theta}(\widetilde{A}^x_a)\in\mathscr{R}_n$, since $\mathbf{\Phi}_n$ is linear

$$\sum_a A^{x,n}_a = \mathbf{\Phi}_n\circ\boldsymbol{\theta}(\sum_a\widetilde{A}^x_a) = 1$$

and hence each $\{A^{x,n}_a\}_{a\in\mathscr{A}}$ is a POVM defined within $(\mathscr{R}_n)$. For each $(y,b)\in\mathscr{X}\times\mathscr{A}$, the state $\frac{1}{\widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})}\widetilde{\psi}^y_b\circ\mathbf{\Phi}_n(A)$ is a normal linear functional with $\frac{1}{\widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})}\widetilde{\psi}^y_b\circ\mathbf{\Phi}_n\leq\frac{1}{\widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})}\widetilde{\psi}\circ\mathbf{\Phi}_n$. Hence, by Theorem 4.2.4, there exists a positive element $B^{y,n}_b\in\mathscr{R}'_n$ such that for all $X\in\mathscr{R}$

$$\frac{\widetilde{\psi}^y_b\circ\mathbf{\Phi}_n(X)}{\widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})} = \langle\psi_n|B^{y,n}_b X|\psi_n\rangle.$$

Since for all $y\in\mathscr{X}$, $X\in\mathscr{N}$, we have

$$\sum_b\frac{\widetilde{\psi}^y_b\circ\mathbf{\Phi}_n\circ\boldsymbol{\theta}(X)}{\widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})} = \frac{\widetilde{\psi}\circ\mathbf{\Phi}_n\circ\boldsymbol{\theta}(X)}{\widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})},$$

this implies that

$$\sum_b\langle\psi_n|B^{y,n}_b X|\psi_n\rangle = \langle\psi_n|X|\psi_n\rangle$$

for all $X \in \mathscr{R}$. This implies that $\sum_b B_b^{y,n} \leq \mathbb{1}_{\mathscr{R}_n}$ with $\langle\psi_n|\left(\mathbb{1}_{\mathscr{R}_n} - \sum_b B_b^{y,n}\right)|\psi_n\rangle = 0$. By adding $(\mathbb{1}_{\mathscr{R}_n} - \sum_b B_b^{y,n})$ to one of the $B_b^{y,n}$, we can further assume that each $\{B_b^{y,n}\}_{b\in\mathscr{A}}$ forms a set of POVMs in $(\mathscr{R}_n)'$. Finally, by Proposition 4.2.3, there exists a sequence of positive operators $\sigma_m \in \mathscr{R}_m$ such that

$$\lim_{m\to\infty} \| |\psi_n\rangle - \sigma_m|\tau_n\rangle \| = 0. \tag{4.4}$$

Since $|\psi\rangle$ is a vector state, without loss of generality, we can also assume that $\tau_n(\sigma_m^2) = \|\sigma_m|\tau_n\rangle\| = 1$ by normalizing each of the σ_m. We see that for each m, $(\mathscr{L}^2(\mathscr{R}_n,\tau_n), \sigma_m|\tau_n\rangle, \{A_a^{x,n}\}, \{B_b^{y,n}\})$ is a tracially embeddable strategy, and by (4.4), we have

$$\begin{aligned}
&\lim_{m\to\infty} \sum_{x,y,a,b} |C^n_{x,y,a,b} - \langle\psi_n|\sigma_m A_a^{x,n} B_b^{y,n} \sigma_m|\psi_n\rangle| \\
&= \lim_{m\to\infty} \sum_{x,y,a,b} |\frac{1}{\widetilde{\psi}\circ\mathbf{\Phi}_n(\mathbb{1}_{\mathscr{R}})} \widetilde{\psi}_b^y \circ \mathbf{\Phi}_n \circ \boldsymbol{\theta}(\widetilde{A}_a^{x,n}) - \langle\psi_n|\sigma_m A_a^{x,n} B_b^{y,n}\sigma_m|\psi_n\rangle| \\
&= \lim_{m\to\infty} \sum_{x,y,a,b} |\langle\psi_n|A_a^{x,n}B_b^{y,n}|\psi_n\rangle - \langle\psi_n|\sigma_m A_a^{x,n}B_b^{y,n}\sigma_m|\psi_n\rangle| = 0,
\end{aligned}$$

showing that $C^n_{x,y,a,b} \in \overline{\mathscr{C}_{qc}^{Tr}(\mathscr{X},\mathscr{A})}$, thus concluding the proof of the main theorem.

□

4.3 Symmetric strategies

In this section, we introduce a natural generalization of symmetric strategy [Vid22, Definition 2.3] for the commuting operator model using the structure of tracially embeddable strategies.

Definition 4.3.1 (Symmetric strategies)**.** *Let $\mathscr{G} = (\mathscr{X}^2, \mathscr{A}^2, \mu, D)$ be a synchronous game and let $(\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{A_a^x\}, \{(B_b^y)^{op}\})$ be a tracially embeddable strategy. We call this strategy symmetric if $A_a^x = B_a^x$ for all $x \in \mathscr{X}$ and $a \in \mathscr{A}$.*

Symmetric strategies will be written as $(\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{A_a^x\})$ in this thesis. If a symmetric strategy is also projective (i.e. $\{A_x^a\}$ are projective) with $\sigma = \mathbb{1}$, then the given strategy is also a synchronous strategy. By using [PSS+16, Theorem 5.5], and taking the double commutant of the corresponding C^*-algebra for Alice's measurement operator, we can always write any synchronous strategies as a projective, symmetric strategy of the form $(\mathscr{L}^2(\mathscr{A},\tau), |\tau\rangle, \{A_a^x\})$.

The following theorem shows that all synchronous correlations can be realized by a projective, symmetric strategy for which the underlying state is the tracial state of the algebra (i.e. the maximally entangled state in the finite dimensional setting).

Lemma 4.3.2 (Synchronous correlations can be realized using a symmetric strategy)**.** *Let $C_{x,y,a,b} \in C_{qc}^s$, then there exists a projective, symmetric strategy $\mathscr{S} = (\mathscr{L}^2(\mathscr{A},\tau), |\tau\rangle, \{A_a^x\})$ which realizes $C_{x,y,a,b}$. Furthermore, if $C_{x,y,a,b} \in C_q^s$, then $\mathscr{S}$ is finite-dimensional.*

The above lemma can be proven by first taking the double commutant of $\{A_a^x\}$ from [PSS+16, Theorem 5.5], then applying point iii) of [PSS+16, Theorem 5.5] to get the desired result. Due to this lemma, we can assume that all synchronous correlations are realized using synchronous strategies guaranteed by Lemma 4.3.2. Hence we denote every synchronous strategy as a projective strategy $\mathscr{S} = (\mathscr{L}^2(\mathscr{A},\tau), |\tau\rangle, \{A_a^x\})$ in this thesis.

We prove some useful lemmas related to symmetric strategies for the remainder of this section.

4.3.1 Lemma related to symmetric strategies

The first lemma regarding symmetric strategies is bounding the synchronicity of any tracially embeddable strategy by the synchronicity of the two symmetric strategies defined by their respective measurement operators.

Lemma 4.3.3. *Let $\mathcal{G} = (\mathcal{X}^2, \mathcal{A}^2, \mu, D)$ be a synchronous game, $\mathcal{S} = (\mathcal{L}^2(\mathcal{A}, \tau), \sigma|\tau\rangle, \{A_a^x\}, \{(B_b^y)^{op}\})$ be a tracially embeddable strategy for the game $\mathcal{G}$, then*

$$1 - \delta_{sync}(\mathcal{G}, \mathcal{S}) \leq \sqrt{1 - \delta_{sync}(\mathcal{G}, \mathcal{S}_{(A_a^x, \sigma|\tau\rangle)})}\sqrt{1 - \delta_{sync}(\mathcal{G}, \mathcal{S}_{(B_b^y, \sigma|\tau\rangle)})},$$

where $\mathcal{S}_{(A_a^x, \sigma|\tau\rangle)}$ is the symmetric strategy defined by $(\mathcal{L}^2(\mathcal{A}, \tau), \sigma|\tau\rangle, \{A_a^x\})$, and likewise with $\mathcal{S}_{(A_a^x, \sigma|\tau\rangle)}$.

Proof. By the Cauchy-Schwarz's inequality and Jensen's inequality,

$$\begin{aligned} 1 - \delta_{\text{sync}}(\mathcal{G}, \mathcal{S}) &= \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|\sigma A_x^a \sigma B_x^a \sigma|\tau\rangle \leq \sqrt{\mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|\sigma A_a^x \sigma A_a^x|\tau\rangle}\sqrt{\mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|\sigma B_a^x \sigma B_a^x|\tau\rangle} \\ &= \sqrt{1 - \delta_{\text{sync}}(\mathcal{G}, \mathcal{S}_{(A_a^x, \sigma|\tau\rangle)})}\sqrt{1 - \delta_{\text{sync}}(\mathcal{G}, \mathcal{S}_{(B_b^y, \sigma|\tau\rangle)})}. \end{aligned}$$

□

The following lemma is an analogue of [Vid22, Lemma 2.10] with tracially embeddable strategies, and the proof follows a similar structure.

Lemma 4.3.4. *Let $\mathcal{G} = (\mathcal{X}^2, \mathcal{A}^2, \mu, D)$ be a synchronous game, and let $\mathcal{S} = (\mathcal{L}^2(\mathcal{A}, \tau), \sigma|\tau\rangle, \{A_a^x\}, \{(B_b^y)^{op}\})$ be a tracially embeddable strategies for the game $\mathcal{G}$. Let $\{C_x^a\} \subseteq \mathcal{A}$ be another set of POVMs, with*

$$\gamma = \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|\sigma(A_a^x - C_a^x)^2\sigma|\tau\rangle,$$

for some $\gamma \geq 0$. Then for the symmetric strategy $\mathcal{S}_{(A_a^x, \sigma|\tau\rangle)} = (\mathcal{L}^2(\mathcal{A}, \tau), \sigma|\tau\rangle, \{A_a^x\})$, we have

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} |\langle\tau|\sigma A_a^x (B_y^b)^{op}\sigma|\tau\rangle - \langle\tau|\sigma C_a^x (B_b^y)^{op}\sigma|\tau\rangle| \leq 6\left(\delta_{sync}(\mathcal{G}, \mathcal{S}_{(A_a^x, \sigma|\tau\rangle)}) + \sqrt{\gamma}\right).$$

Proof. For simplicity of notation, let $\delta_A = \delta_{\text{sync}}(\mathcal{G}, \mathcal{S}_{(A_a^x, \sigma|\tau\rangle)})$. We show this lemma using the following three inequalities, followed by the triangle inequality:

$$\mathop{\mathbb{E}}_{x,y\sim\mu} \sum_{a,b} |\langle\tau|\sigma A_a^x (B_b^y)^{op}\sigma|\tau\rangle - \langle\tau|\sigma (A_a^x)^2 (B_b^y)^{op}\sigma|\tau\rangle| \leq 2\delta_A, \tag{4.5}$$

$$\mathop{\mathbb{E}}_{x,y\sim\mu} \sum_{a,b} |\langle\tau|\sigma (A_a^x)^2 (B_b^y)^{op}\sigma|\tau\rangle - \langle\tau|\sigma (C_a^x)^2 (B_b^y)^{op}\sigma|\tau\rangle| \leq 2\sqrt{\gamma}, \tag{4.6}$$

$$\mathop{\mathbb{E}}_{x,y\sim\mu} \sum_{a,b} |\langle\tau|\sigma C_a^x (B_b^y)^{op}\sigma|\tau\rangle - \langle\tau|\sigma (C_a^x)^2 (B_b^y)^{op}\sigma|\tau\rangle| \leq 4\left(\delta_A + \sqrt{\gamma}\right). \tag{4.7}$$

For (4.5),

$$\begin{aligned} \mathop{\mathbb{E}}_{x,y\sim\mu} \sum_{a,b} |\langle\tau|\sigma A_a^x (B_b^y)^{op}\sigma|\tau\rangle - \langle\tau|\sigma (A_a^x)^2 (B_b^y)^{op}\sigma|\tau\rangle| &= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} \langle\tau|\sigma (A_a^x - (A_a^x)^2)(B_b^y)^{op}\sigma|\tau\rangle \\ &= 1 - \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|\sigma (A_a^x)^2 \sigma|\tau\rangle. \end{aligned} \tag{4.8}$$

By the Cauchy-Schwarz's inequality and Jensen's inequality,

$$\begin{aligned} 1 - \delta_A = \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|\sigma A_a^x (A_a^x)^{op}\sigma|\tau\rangle &\leq \sqrt{\mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|\sigma (A_a^x)^2 \sigma|\tau\rangle}\sqrt{\mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|\sigma ((A_a^x)^{op})^2 \sigma|\tau\rangle} \\ &\leq \sqrt{\mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|\sigma (A_a^x)^2 \sigma|\tau\rangle}. \end{aligned}$$

Solving for $\mathbb{E}_{x\sim\mu} \sum_a \langle\tau|\sigma (A_a^x)^2 \sigma|\tau\rangle$, we have

$$1 - 2\delta_A \leq 1 - 2\delta_A + \delta_A^2 \leq \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|\sigma (A_a^x)^2 \sigma|\tau\rangle, \tag{4.9}$$

and hence (4.5) follows accordingly. For (4.6), we bound the inequality through the following two inequalities:

$$\begin{aligned}\mathop{\mathbb{E}}_{x,y\sim\mu}\sum_{a,b}\langle\tau|\sigma((A_a^x)^2-A_a^xC_a^x)(B_b^y)^{op}\sigma|\tau\rangle\le\sqrt{\gamma},\\ \mathop{\mathbb{E}}_{x,y\sim\mu}\sum_{a,b}\langle\tau|\sigma((C_a^x)^2-A_a^xC_a^x)(B_b^y)^{op}\sigma|\tau\rangle\le\sqrt{\gamma}.\end{aligned}\tag{4.10}$$

Since the inequalities are proven in a similar manner, we will only show the proof for (4.10): by Cauchy-Schwarz's inequality and Jensen's inequality,

$$\begin{aligned}&\mathop{\mathbb{E}}_{x,y\sim\mu}\sum_{a,b}\langle\tau|\sigma((A_a^x)^2-A_a^xC_a^x)(B_b^y)^{op}\sigma|\tau\rangle\\ &\le\sqrt{\mathop{\mathbb{E}}_{x,y\sim\mu}\sum_{a,b}\langle\tau|\sigma(A_a^x)^2((B_b^y)^2)^{op}\sigma|\tau\rangle}\sqrt{\mathop{\mathbb{E}}_{x,y\sim\mu}\sum_{a,b}\langle\tau|\sigma(A_a^x-C_a^x)^2\sigma|\tau\rangle}\\ &\le 1\cdot\sqrt{\mathop{\mathbb{E}}_{x,y\sim\mu}\sum_{a}\langle\tau|\sigma(A_a^x-C_a^x)^2\sigma|\tau\rangle}\le\sqrt{\gamma}.\end{aligned}$$

For (4.7), this is analogous to (4.5), except we are estimating $\delta_C=\delta_{\text{sync}}(\mathcal{G},\mathcal{S}_{(C_a^x,\sigma|\tau\rangle)})$ instead, where $\mathcal{S}_{(C_a^x,\sigma|\tau\rangle)}$ is the symmetric strategy $(\mathcal{L}^2(\mathcal{A},\tau),\sigma|\tau\rangle,\{C_a^x\})$. To bound δ_C, let η denote the synchronicity for the tracially embeddable strategy $(\mathcal{L}^2(\mathcal{A},\tau),\sigma|\tau\rangle,\{C_a^x\},\{(A_b^y)^{op}\})$. By Cauchy-Schwarz's inequality and Jensen's inequality,

$$\begin{aligned}\eta-\delta_A&\le\mathop{\mathbb{E}}_{x\sim\mu}\sum_a|\langle\tau|\sigma(C_a^x)(A_a^x)^{op}\sigma|\tau\rangle-\langle\tau|\sigma(A_a^x)(A_a^x)^{op}\sigma|\tau\rangle|\\ &=\mathop{\mathbb{E}}_{x\sim\mu}\sum_a|\langle\tau|\sigma(C_a^x-A_a^x)(A_a^x)^{op}\sigma|\tau\rangle|\\ &\le\sqrt{\mathop{\mathbb{E}}_{x\sim\mu}\sum_a\langle\tau|\sigma(C_a^x-A_a^x)^2\sigma|\tau\rangle}\sqrt{\mathop{\mathbb{E}}_{x\sim\mu}\sum_a\langle\tau|\left((A_a^x)^2\right)^{op}\sigma|\tau\rangle}\le\sqrt{\gamma}.\end{aligned}$$

By solving for η and by Lemma 4.3.3, $\delta_A\le 2\eta\le 2\left(\delta_A+\sqrt{\gamma}\right)$. Proceeding the same way as (4.5)

$$\begin{aligned}\mathop{\mathbb{E}}_{x,y\sim\mu}\sum_{a,b}|\langle\tau|\sigma C_a^x(B_b^y)^{op}\sigma|\tau\rangle-\langle\tau|\sigma(C_a^x)^2(B_b^y)^{op}\sigma|\tau\rangle|&\le 1-\mathop{\mathbb{E}}_{x\sim\mu}\sum_a\langle\tau|\sigma(C_a^x)^2\sigma|\tau\rangle\\ &\le 2\delta_C\le 4\left(\delta_A+\sqrt{\gamma}\right).\end{aligned}$$

□

Combining Lemma 4.3.3 and Lemma 4.3.4, we have the following lemma.

Lemma 4.3.5. *Let $\mathcal{G}=(\mathcal{X}^2,\mathcal{A}^2,\mu,D)$ be a synchronous game, and let $\mathcal{S}=(\mathcal{L}^2(\mathcal{A},\tau),\sigma|\tau\rangle,\{A_a^x\},\{B_b^y\})$ be a δ-synchronous strategy for $\mathcal{G}$. Then the symmetric strategy $\mathcal{S}'=(\mathcal{L}^2(\mathcal{A},\tau),\sigma|\tau\rangle,\{A_a^x\})$ is a $O(\delta)$-synchronous strategy with*

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma A_a^x(B_b^y)^{op}\sigma|\tau\rangle-\langle\tau|\sigma A_a^x(A_b^y)^{op}\sigma|\tau\rangle|\le O(\sqrt{\delta}),$$

Proof. Let δ_A be the synchronicity of the strategy $\mathcal{S}'$, by Lemma 4.3.3, $\delta_A\le 2\delta=O(\delta)$. For the second part of the lemma, we see that by triangle inequality, we have

$$\mathop{\mathbb{E}}_{x\sim\mu}\sum_a\langle\tau|\sigma((A_a^x-B_a^x)^{op})^2\sigma|\tau\rangle\le\mathop{\mathbb{E}}_{x\sim\mu}\sum_a(||(A_a^x-(B_a^x)^{op})\sigma|\tau\rangle||^2+||(A_a^x-(A_a^x)^{op})\sigma|\tau\rangle||^2).\tag{4.11}$$

For the first part of (4.11), since $(A_a^x)^2\le A_a^x\le\mathbb{I}$,

$$\begin{aligned}\mathop{\mathbb{E}}_{x\sim\mu}\sum_a\|(A_a^x-(B_a^x)^{op})\sigma|\tau\rangle\|^2&=\mathop{\mathbb{E}}_{x\sim\mu}\sum_a\left(\langle\tau|\sigma(A_a^x)^2\sigma|\tau\rangle+\langle\tau|\sigma((B_a^x)^2)^{op}\sigma|\tau\rangle-2\langle\tau|\sigma A_a^x(B_a^x)^{op}\sigma|\tau\rangle\right)\\ &\le\mathop{\mathbb{E}}_{x\sim\mu}\sum_a\left(2-2\langle\tau|\sigma A_a^x(B_a^x)^{op}\sigma|\tau\rangle\right)=2\delta.\end{aligned}$$

For the second part of (4.11), by a similar calculation

$$\mathop{\mathbb{E}}_{x\sim\mu}\sum_a \|(A_a^x-(A_a^x)^{op})\sigma|\tau\rangle\|^2 \le \mathop{\mathbb{E}}_{x\sim\mu}\sum_a (2-2\langle\tau|\sigma A_x^a(A_x^a)^{op}\sigma|\tau\rangle) \le 2\delta_A \le O(\delta). \tag{4.12}$$

Hence, by combining the above two equations, we have

$$\mathop{\mathbb{E}}_{x\sim\mu}\sum_a \langle\tau|\sigma((A_a^x-B_a^x)^{op})^2\sigma|\tau\rangle \le O(\delta).$$

By Lemma 4.3.4 and as well as Lemma 4.3.3 apply on the symmetric strategy $\mathscr{S}^B=(\mathscr{L}^2(\mathscr{A},\tau),\sigma|\tau\rangle,\{B_b^y\}_{b\in\mathscr{A}})$, we have

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma|A_a^x(B_b^y)^{op}\sigma|\tau\rangle-\langle\tau|\sigma A_a^x(A_b^y)^{op}\sigma|\tau\rangle| \le O(\sqrt{\delta})+O(\delta_B)\le O(\sqrt{\delta})$$

where δ_B denotes the synchronicity of $\mathscr{S}^B$ this concludes the second claim of the lemma. □

The following corollary shows that every δ-synchronous correlation can be approximated by some correlations generated by a symmetric projective strategy defined on the standard form of a tracial von Neumann algebra $(\mathscr{A},\tau)$.

Corollary 4.3.6. *Let $\mathcal{G}=(\mathscr{X}^2,\mathscr{A}^2,\mu,D)$ be a synchronous game, and let $(\mathscr{L}^2(\mathscr{A},\tau),\sigma|\tau\rangle,\{A_a^x\},\{B_b^y\})$ be a δ-synchronous strategy for $\mathcal{G}$. There exists a symmetric, projective and $\delta^{\frac{1}{4}}$-synchronous strategy $(\mathscr{L}^2(\mathscr{A},\tau),\sigma|\tau\rangle,\{P_a^x\})$ with*

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}\|\left(A_a^x-P_a^x\right)\sigma|\tau\rangle\|^2\le O(\delta),$$

such that

$$\mathop{\mathbb{E}}_{x\sim\mu}\sum_{a\in A}|\langle\tau|\sigma A_a^x(B_b^y)^{op}\sigma|\tau\rangle-\langle\tau|\sigma P_a^x(P_b^y)^{op}\sigma|\tau\rangle|\le O(\delta^{\frac{1}{4}}), \tag{4.13}$$

Proof. Let $\mathscr{S}^{sym}=(\mathscr{L}^2(\mathscr{A},\tau),\sigma|\tau\rangle,\{A_a^x\}_{a\in\mathscr{A}})$ be a symmetric strategy. By Lemma 4.3.5, $\delta_{\text{sync}}(\mathcal{G},\mathscr{S}^{sym})=O(\delta)$ and

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma A_a^x(B_b^y)^{op}\sigma|\tau\rangle-\langle\tau|\sigma A_a^x(A_b^y)^{op}\sigma|\tau\rangle|\le O(\sqrt{\delta}). \tag{4.14}$$

By the Cauchy-Schwarz's inequality,

$$1-O(\delta)=\mathop{\mathbb{E}}_{x\sim\mu}\sum_a\langle\tau|\sigma A_a^x(A_a^x)^{op}\sigma|\tau\rangle\le\sqrt{\mathop{\mathbb{E}}_{x\sim\mu}\sum_a\langle\tau|\sigma(A_a^x)^2\sigma|\tau\rangle},$$

or $1-O(\delta)\le\mathbb{E}_{x\sim\mu}\sum_a\langle\psi|(A_a^x)^2|\psi\rangle$. Hence, for all $x\in\mathscr{X}$, by applying Lemma 3.3.1 on $\{A_a^x\}$, there exists a set of projective measurements $\{P_a^x\}$ such that

$$\mathop{\mathbb{E}}_{x\sim\mu}\|\left(A_a^x-P_a^x\right)\sigma|\tau\rangle\|^2=\mathop{\mathbb{E}}_{x\sim\mu}\langle\tau|\sigma(A_a^x-P_a^x)^2\sigma|\tau\rangle\le O(\delta). \tag{4.15}$$

By applying Lemma 4.3.4 on $\{P_a^x\}$ for $\mathscr{S}^{sym}$, we have

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma A_a^x(A_b^y)^{op}\sigma|\tau\rangle-\langle\tau|\sigma P_a^x(A_b^y)^{op}\sigma|\tau\rangle|\le O(\sqrt{\delta}).$$

Let $\mathscr{S}^2=(\mathscr{L}^2(\mathscr{A},\tau),\sigma|\tau\rangle,\{A_a^x\}_{a\in\mathscr{A}},\{(P_a^x)^{op}\}_{a\in\mathscr{A}})$. By Cauchy-Schwarz's inequality, and $|\tau\rangle$ being a tracial state, we have

$$\begin{aligned}
\delta_{\text{sync}}(\mathcal{G},\mathscr{S}^2)-\delta_{\text{sync}}(\mathcal{G},\mathscr{S}^{sym}) &\le \mathop{\mathbb{E}}_x\sum_a\langle\tau|\sigma A_a^x(P_a^x)^{op}\sigma|\tau\rangle-\langle\tau|\sigma A_a^x(A_a^x)^{op}\sigma|\tau\rangle\\
&\le\sqrt{\mathop{\mathbb{E}}_x\sum_a\langle\tau|\sigma(A_a^x)^2\sigma|\tau\rangle}\sqrt{\mathop{\mathbb{E}}_x\sum_a\langle\tau|\sigma((A_x^a-P_a^x)^2)^{op}\sigma|\tau\rangle}\\
&\le 1\cdot\sqrt{(\mathop{\mathbb{E}}_x\sum_a\langle\tau|(\sigma)^2((A_a^x-P_a^x)^2)^{op}|\tau\rangle}\\
&=\sqrt{\mathop{\mathbb{E}}_x\sum_a\langle\tau|\sigma(A_a^x-P_a^x)^2\sigma|\tau\rangle}=O(\sqrt{\delta}),
\end{aligned}$$

and hence $\delta_2\le O(\sqrt{\delta})$. By repeating this argument again, on $(A_a^x)^{op}$ with $\mathscr{S}_2$ and together with (4.14) and (4.15) gives (4.13). □

4.4 Oracularizable strategy

Before we introduce the rounding theorem, we introduce a special class of strategies known as an *oracularizable strategy* for synchronous games below. We remark that this set of strategies follows an analogue of the "commuting" property, a property for finite-dimensional strategies, given in [JNV+22a, Definition 5.8].

Definition 4.4.1 (Oracularizable strategy)**.** *A tracially embeddable strategy* $\mathscr{S} = (\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{A_a^x\}, \{B_b^y\})$ *for a synchronous game* $\mathscr{G} = (\mathscr{X},\mathscr{A},\mu,D)$ *is oracularizable if whenever* $\mu(x,y) > 0$*, then for all* $(a,b) \in \mathscr{A}^2$

$$[A_a^x, B_b^y] = A_a^x B_b^y - B_b^y A_a^x = 0.$$

We remark that in the above theorem, both $\{A_a^x\}$ and $\{\left(B_b^y\right)^{op}\}$ are defined in $\mathscr{A}$. Hence the above condition does not follow immediately from the definition for a commuting operator strategy.

The notion of an oracularizable strategy itself is pretty interesting in the overall literature about $\mathsf{MIP}^*/\mathsf{MIP}^{\mathrm{co}}$. In particular, it is an integral part of the "completeness clause" for the compression theorem, and as will be seen in Part III, one has to carefully check that a perfect oracularizable strategy must exist in each of the subroutine for the compression theorem. However, the only purpose of maintaining such conditions is to prove the "completeness" clause within the oracularization transformation within the answer reduction procedure (see Section 11.2 for more details).

Given the above summary on how oracularizable strategies are used, one might be quickly tempted to think that they are not important to the uncomputability of non-local games. This, however, is not the case, when one tries to generalizes the unique games conjecture [Kho02] for $\mathsf{MIP}^*/\mathsf{MIP}^{\mathrm{co}}$ protocols. In an earlier work, Kempe, Regev and Toner shows that the unique games [KRT09] in the entanglement setting is in P, hence showing that the unique games conjecture to be false under the $\mathsf{MIP}^*/\mathsf{MIP}^{\mathrm{co}}$ model. However, by the recent work of Mousavi and Spirig [MS24b], if one were to restrict to the classes of allowed strategy to be oracularizable, the landscape all of the sudden changes, and the "quantum oracularizable unique games conjecture" could all of a sudden be true (and the complexity of unique games with perfect oracularizable strategies is uncomputable)! Since this is not the main focus of this thesis, we refer interested readers to the reference given for more details.

CHAPTER 5

Rounding almost synchronous correlations

In this chapter, we introduce and prove the "rounding" lemma in the commuting operator setting. This theorem was proven for the tensor product model by both [Vid22] and [Pau22], and it was proven independently in [dlSM23] for the commuting model. Interestingly, the proof given in [dlSM23] uses advanced techniques from von Neumann algebra theory (namely the Tamita-Takesaki theorem). In comparison, our proof only relies on Theorem 4.1.3 and the structure of tracially embeddable strategies, which allows us to mimic the proof structure of [Vid22].

5.1 The rounding lemma

Before introducing this lemma, we first introduce the notion of synchronicity given a correlation. Given a distribution $\mu \sim \mathcal{X}^2$, we denote the *synchronicity* of the correlation set with respect to μ as

$$\delta_{\text{sync}}(\mu, C) := \max\left\{ \mathop{\mathbb{E}}_{x\sim\mu_x} \sum_{a\neq b} C_{x,x,a,b}, \mathop{\mathbb{E}}_{y\sim\mu_y} \sum_{a\neq b} C_{y,y,a,b} \right\}, \tag{5.1}$$

where μ_x (resp. μ_y) denotes the marginal distribution of x (resp. y) over μ. For a quantum strategy $\mathcal{S}$, we use $\delta_{\text{sync}}(\mu, \mathcal{S})$ to denote the synchronicity for the correlation generated by $\mathcal{S}$ with respect to μ. For a tensor product/commuting operator strategy $\mathcal{S} = \mathcal{S} = (\mathcal{H}, \{A_a^x\}_{a\in\mathcal{A}}, \{B_b^y\}_{b\in\mathcal{A}}, |\psi\rangle)$, by definition we have

$$A_a^x \simeq_{\delta_{\text{sync}}(\mu,\mathcal{S})} B_a^x \tag{5.2}$$

where $\simeq_{\delta_{\text{sync}}(\mu,\mathcal{S})}$ is over the state $|\psi\rangle$ and both the distribution μ_x and μ_y. We define a correlation C to be *δ-synchronous* with respect to μ if $\delta_{\text{sync}}(\mu, C) \leq \delta$ and δ-synchronous if the underlying distribution μ is clear from context.

We are now ready to introduce the "rounding theorem" below.

Theorem 5.1.1 (Rounding for synchronous correlations)**.** *There exist a universal polynomial $s^{Rd} : [0,1] \to [0,1]$ such that $\boldsymbol{s}^{Rounding}(\delta) = O(\delta^{\frac{1}{8}})$ such that the following holds: Let $\mu \sim \mathcal{X}^2$ be a distribution and $t \in \{q, qc\}$, and let $\{C_{x,y,a,b}\} \in \mathcal{C}_t(\mathcal{X}, \mathcal{A})$ be a δ-synchronous correlation. Then there exist a collection of $\{C^{\lambda}_{x,y,a,b}\}_{\lambda\in[0,\infty]} \subseteq \mathcal{C}^s_{qc}(\mathcal{X}, \mathcal{A})$ and a probability distribution P over $[0,\infty]$, such that*

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} \left| C_{x,y,a,b} - \int_0^{\infty} P(\lambda) \cdot C^{\lambda}_{x,y,a,b} d\lambda \right| \leq \boldsymbol{s}^{Rounding}(\delta).$$

Furthermore, if $C_{x,y,a,b}$ is realized by a tracially embeddable strategy $(\mathcal{L}^2(\mathcal{A}, \tau), \sigma|\tau\rangle, \{A_a^x\}, \{(B_b^y)^{op}\})$, then each of the $C^{\lambda}_{x,y,a,b}$ can be realized by a synchronous strategy defined on the standard form of some subalgebra of $\mathcal{A}$.

Intuitively, since the convex combination of synchronous correlations is still synchronous, the above theorem essentially states that the set of δ-synchronous can always be approximated by the set of synchronous correlations.

5.2 Proof of Theorem 5.1.1

The proof of Theorem 5.1.1 relies on a lemma originating from the seminal paper of Connes [Con76]. We remark that the finite dimensional variant of this lemma is an important tool used in both [Vid22] and [Pau22], and [dlSM23] generalized this lemma to the type III von Neumann algebra setting using the core of the algebra.

For $\lambda \in \mathbb{R}$, we define the characteristic function $\chi_{\geq\lambda} : \mathbb{R} \to \mathbb{R}$ as

$$\chi_{\geq\lambda}(x) := \begin{cases} 1 & \text{if } x \geq \lambda \\ 0 & \text{otherwise.} \end{cases}$$

We extend, via functional calculus, the above function acting on $\mathscr{A}^+$. Since $\int_0^\infty \chi_{\geq\sqrt{\lambda}}(x)d\lambda = x^2$ for all $x \in \mathbb{R}$, by extension, for all $\sigma \in \mathscr{A}^+$

$$\int_0^\infty \chi_{\geq\sqrt{\lambda}}(\sigma)d\lambda = \sigma^2. \tag{5.3}$$

We recall the following lemma due to Connes [Con76, Lemma 1.2.6].

Lemma 5.2.1 (Connes' joint distribution lemma)**.** *Let $(\mathscr{A},\tau)$ be a tracial von Neumann algebra and ρ,σ be two positive elements in $\mathscr{A}$. Then*

$$\int_0^\infty ||\chi_{\geq\sqrt{\lambda}}(\rho) - \chi_{\geq\sqrt{\lambda}}(\sigma)||_2^2 d\lambda \leq ||\rho-\sigma||_2 \cdot ||\rho+\sigma||_2.$$

To show Theorem 5.1.1, we first show the following proposition. This proposition is analogous to [Vid22, Theorem 3.1] in the infinite-dimensional case. It states that any projective, symmetric, almost synchronous strategy defined on the standard form of some tracial von Neumann algebra $(\mathscr{A},\tau)$ can be approximated by a set of synchronous strategies, each on some subalgebra of $\mathscr{A}$.

Proposition 5.2.2. *Let $\mathscr{G} = (\mathscr{X}^2, \mathscr{A}^2, \mu, D)$ be a synchronous game and let $(\mathscr{A},\tau)$ be a tracial von Neumann algebra in standard form. Furthermore, let $(\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{A_a^x\})$ be a projective, symmetric and δ-synchronous strategy. There exists a set of projectors $\{P_\lambda\}_{[0,\infty]} \subseteq \mathscr{A}$ such that*

$$\int_0^\infty P_\lambda d\lambda = \sigma^2,$$

and for each λ, there exists a set of PVMs $\{A_a^{\lambda,x}\}_{\lambda\in[0,\infty]} \subseteq P_\lambda \mathscr{A} P_\lambda$ such that

$$\int_0^\infty \mathop{\mathbb{E}}_{x\sim\mu} \sum_a ||(A_a^x - A_a^{\lambda,x})P_\lambda||_2^2 d\lambda \leq O(\delta^{\frac{1}{4}}).$$

Proof. Let $P_\lambda := \chi_{\geq\sqrt{\lambda}}(\sigma)$, by (5.3) $\int_0^\infty P_\lambda = \sigma^2$. We first show

$$\int_0^\infty \mathop{\mathbb{E}}_{x\sim\mu} \sum_a ||(A_a^x - P_\lambda A_a^x P_\lambda)P_\lambda||_2^2 d\lambda \leq 2\sqrt{\delta}. \tag{5.4}$$

By Hölder's inequality

$$\int_0^\infty \mathop{\mathbb{E}}_{x\sim\mu} \sum_a ||(A_a^x - P_\lambda A_a^x P_\lambda)P_\lambda||_2^2 d\lambda \leq \mathop{\mathbb{E}}_{x\sim\mu} \int_0^\infty \sum_a ||(A_a^x P_\lambda - P_\lambda A_a^x)||_2^2 d\lambda.$$

Let $m = |\mathscr{A}|$ and associate the answer set with $\mathbb{Z}_m$ and define $U_b^x := \sum_a^m e^{\frac{2i\pi ab}{m}} A_a^x$. Since A_a^x is a PVM, U_b^x defines a unitary in $\mathscr{A}$ for each $a \in \mathscr{A}$ and $b \in \mathbb{Z}_m$. By substituting A_a^x with U_b^x

$$\begin{aligned}\mathop{\mathbb{E}}_{x\sim\mu}\int_0^\infty \sum_{a\in\mathbb{Z}_m} ||A_a^x P_\lambda - P_\lambda A_a^x||_2^2 d\lambda &= \mathop{\mathbb{E}}_{x\sim\mu}\int_0^\infty \left(\sum_{a\in\mathbb{Z}_m} 2(\langle\tau|P_\lambda A_a^x|\tau\rangle - \langle\tau|P_\lambda A_a^x P_\lambda A_a^x|\tau\rangle)\right) d\lambda \\ &= \mathop{\mathbb{E}}_{x\sim\mu}\int_0^\infty \left(2\langle\tau|P_\lambda|\tau\rangle - 2\mathop{\mathbb{E}}_b \langle\tau|P_\lambda (U_b^x)^* P_\lambda U_b^x|\tau\rangle\right) d\lambda \\ &= \mathop{\mathbb{E}}_{x\sim\mu}\mathop{\mathbb{E}}_b \int_0^\infty ||U_b^x P_\lambda - P_\lambda U_b^x||_2^2 d\lambda.\end{aligned} \tag{5.5}$$

By Lemma 5.2.1, since $\chi_{\geq\sqrt{\lambda}}\left((U_b^x)^*\sigma U_b^x\right) = (U_b^x)^*\left(\chi_{\geq\sqrt{\lambda}}(\sigma)\right)U_b^x$ for all λ, it follows that

$$\begin{aligned}\mathop{\mathbb{E}}_{x\sim\mu}\mathop{\mathbb{E}}_b \int_0^\infty ||U_b^x P_\lambda - P_\lambda U_b^x||_2^2 d\lambda &= \mathop{\mathbb{E}}_{x\sim\mu}\mathop{\mathbb{E}}_b \int_0^\infty ||P_\lambda - (U_b^x)^* P_\lambda U_b^x||_2^2 d\lambda \\ &\leq \mathop{\mathbb{E}}_{x\sim\mu}\mathop{\mathbb{E}}_b ||\sigma - (U_b^x)^*\sigma U_b^x||_2 ||\sigma + (U_b^x)^*\sigma U_b^x||_2. \\ &\leq \sqrt{\mathop{\mathbb{E}}_{x\sim\mu}\mathop{\mathbb{E}}_b ||\sigma - (U_b^x)^*\sigma U_b^x||_2^2} \cdot \sqrt{\mathop{\mathbb{E}}_{x\sim\mu}\mathop{\mathbb{E}}_b ||\sigma + (U_b^x)^*\sigma U_b^x||_2^2} \\ &\leq \sqrt{2}\cdot\sqrt{\mathop{\mathbb{E}}_{x\sim\mu}\mathop{\mathbb{E}}_b ||\sigma - (U_b^x)^*\sigma U_b^x||_2^2}.\end{aligned}$$

By a similar calculation as (5.5)

$$\sqrt{2}\cdot\sqrt{\mathop{\mathbb{E}}_{x\sim\mu}\mathop{\mathbb{E}}_b ||\sigma - (U_b^x)^*\sigma U_b^x||_2^2} = \sqrt{2}\cdot\sqrt{2 - 2\mathop{\mathbb{E}}_{x\sim\mu}\sum_a \langle\tau|\sigma A_a^x \sigma A_a^x|\tau\rangle} = \sqrt{2}\cdot\sqrt{2\delta} = 2\sqrt{\delta},$$

showing (5.4). Now, for each λ, the algebra $P_\lambda \mathscr{A} P_\lambda$ is a tracial von Neumann algebra with the corresponding tracial state being $\frac{1}{\sqrt{\tau(P_\lambda)}} P_\lambda|\tau\rangle$. Furthermore, $\{P_\lambda A_a^x P_\lambda\}$ defines a set of POVMs within the algebra $P_\lambda \mathscr{A} P_\lambda$, and we can use Lemma 3.3.1 to construct a set of PVMs to approximate the above POVM. To this end, observe that

$$\begin{aligned}1 - 2\delta \leq \mathop{\mathbb{E}}_{x\sim\mu}\sum_a \langle\tau|\sigma(A_a^x)^2\sigma|\tau\rangle &= \mathop{\mathbb{E}}_{x\sim\mu}\sum_a \int_0^\infty \langle\tau|P_\lambda (A_a^x)^2 P_\lambda|\tau\rangle\, d\lambda \\ &= \mathop{\mathbb{E}}_{x\sim\mu}\sum_a \int_0^\infty \left(\langle\tau|P_\lambda A_a^x P_\lambda A_a^x|\tau\rangle + \langle\tau|P_\lambda A_a^x (A_a^x P_\lambda - P_\lambda A_a^x)|\tau\rangle\right) d\lambda.\end{aligned} \tag{5.6}$$

To bound the second term of (5.6), by Cauchy-Schwarz's inequality, then by (5.5)

$$\begin{aligned}&\mathop{\mathbb{E}}_{x\sim\mu}\sum_a \int_0^\infty \langle\tau|P_\lambda A_a^x (A_a^x P_\lambda - P_\lambda A_a^x)|\tau\rangle\, d\lambda \\ &\leq \sqrt{\mathop{\mathbb{E}}_{x\sim\mu}\sum_a \int_0^\infty \langle\tau|P_\lambda (A_a^x)^2 P_\lambda|\tau\rangle\, d\lambda} \cdot \sqrt{\mathop{\mathbb{E}}_{x\sim\mu}\sum_a \int_0^\infty ||A_a^x P_\lambda - P_\lambda A_a^x||_2^2 d\lambda} \leq 1\cdot\sqrt{2}\delta^{\frac{1}{4}}.\end{aligned} \tag{5.7}$$

By rewriting the first term of (5.6) and bounding the second term using (5.7),

$$\begin{aligned}&\mathop{\mathbb{E}}_{x\sim\mu}\sum_a \int_0^\infty (\langle\tau|P_\lambda A_a^x P_\lambda A_a^x|\tau\rangle + \langle\tau|P_\lambda A_a^x (A_a^x P_\lambda - P_\lambda A_a^x)|\tau\rangle) \\ &\leq \mathop{\mathbb{E}}_{x\sim\mu}\sum_a \int_0^\infty \langle\tau|P_\lambda (P_\lambda A_a^x P_\lambda)^2 P_\lambda|\tau\rangle + \sqrt{2}\delta^{\frac{1}{4}}.\end{aligned}$$

Solving for $\mathbb{E}_{x\sim\mu}\sum_a \int_0^\infty \langle\tau|P_\lambda (P_\lambda A_a^x P_\lambda)^2 P_\lambda|\tau\rangle$, we have

$$1 - 2\delta - \sqrt{2}\delta^{\frac{1}{4}} \leq \mathop{\mathbb{E}}_{x\sim\mu}\sum_a \int_0^\infty \langle\tau|P_\lambda (P_\lambda A_a^x P_\lambda)^2 P_\lambda|\tau\rangle\,. \tag{5.8}$$

Hence, by Lemma 3.3.1, for each $\lambda \in \mathbb{R}$ and $x \in \mathscr{X}$, there exists a set of PVMs $\{A_a^{\lambda,x}\} \subseteq P_\lambda \mathscr{A} P_\lambda$ such that

$$\int_0^\infty \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|P_\lambda (P_\lambda A_a^x P_\lambda - A_a^{\lambda,x})^2 P_\lambda|\tau\rangle \, d\lambda \leq 2\delta + \sqrt{2}\delta^{\frac{1}{4}} = O(\delta^{\frac{1}{4}}). \tag{5.9}$$

Combining (5.4) and (5.9) via the triangle inequality

$$\int_0^\infty \mathop{\mathbb{E}}_{x\sim\mu} \sum_a ||(A_a^x - A_a^{\lambda,x})P_\lambda||_2^2 d\lambda \leq O(\delta^{\frac{1}{4}}). \qquad \square$$

We are now ready to give the proof of Theorem 5.1.1.

Proof. Let $\{C_{x,y,a,b}\} \in \mathscr{C}_{qc}(\mathscr{X},\mathscr{A})$ be a δ-synchronous correlation for the game $\mathscr{G}$. By Theorem 4.1.3, we can, by perturbing $C_{x,y,a,b}$ up to arbitrary $\varepsilon > 0$ accuracy, assume the existence of some tracially embeddable strategy $(\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{A_a^x\}_{a\in\mathscr{A}}, \{B_b^y\}_{b\in\mathscr{A}})$ on a tracial von Neumann algebra $(\mathscr{A},\tau)$ such that

$$C_{x,y,a,b} = \langle\tau|\sigma A_a^x (B_b^y)^{op} \sigma|\tau\rangle.$$

By Corollary 4.3.6, there exist a set of PVMs $\{Q_a^x\}$ and a $O(\delta^{\frac{1}{4}})$-synchronous, symmetric and projective strategy $\mathscr{S}' = (\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{Q_a^x\}_{a\in\mathscr{A}})$ such that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} |\langle\tau|\sigma A_a^x (B_y^b)^{op}\sigma|\tau\rangle - \langle\tau|\sigma Q_a^x (Q_b^y)^{op}\sigma|\tau\rangle| \leq O(\delta^{\frac{1}{4}}), \tag{5.10}$$

with

$$\mathop{\mathbb{E}}_{x\sim\mu} \| \left(A_a^x - Q_a^x\right)\sigma|\tau\rangle \|^2 \leq O(\delta). \tag{5.11}$$

Applying Proposition 5.2.2 to $\mathscr{S}'$, we get a set of projectors $\{P_\lambda\}_{[0,\infty]} \subseteq \mathscr{A}$ with $\int_0^\infty P_\lambda d\lambda = (\sigma^+)^2$, and a set of PVMs $\{A_a^{\lambda,x}\}_{\lambda\in[0,\infty],x\in\mathscr{X}}$, with $\{A_a^{\lambda,x}\} \subseteq P_\lambda \mathscr{A} P_\lambda$ for all $\lambda \in \mathbb{R}^+$, such that

$$\int_0^\infty \mathop{\mathbb{E}}_{x\sim\mu} \sum_a ||(Q_a^x - A_a^{\lambda,x})P_\lambda||_2^2 d\lambda \leq O(\delta^{\frac{1}{4}}). \tag{5.12}$$

For $\tau(P_\lambda) \neq 0$, the symmetric strategy $\mathscr{S}^\lambda = (\mathscr{L}^2(P_\lambda \mathscr{A} P_\lambda, P_\lambda|\tau\rangle), \frac{1}{\sqrt{\tau(P_\lambda)}} P_\lambda|\tau\rangle, \{A_a^{\lambda,x}\})$ is synchronous since $\frac{1}{\sqrt{\tau(P_\lambda)}} P_\lambda|\tau\rangle$ is a tracial state for $P_\lambda \mathscr{A} P_\lambda$ (see the remark after Definition 4.3.1). Let $B_a^{\lambda,x}$ be the corresponding opposite algebra element for $A_a^{\lambda,x}$ define on the tracial von Neumann algebra $(P_\lambda \mathscr{A} P_\lambda, \frac{1}{\sqrt{\tau P_\lambda}} P_\lambda)$. In other words, we have $(A_a^{\lambda,x})P_\lambda|\tau\rangle = (B_a^{\lambda,x})P_\lambda|\tau\rangle$ for all $x \in \mathscr{X}$ and $a \in \mathscr{A}$. We claim that for all $x \in \mathscr{X}$ and $a \in \mathscr{A}$

$$\langle\tau|P_\lambda A_a^{\lambda,x} B_a^{\lambda,x} P_\lambda|\tau\rangle = \langle\tau|P_\lambda (A_a^{\lambda,x})(A_a^{\lambda,x})^{op} P_\lambda|\tau\rangle$$

where the op map above corresponds to the opposite algebra map for $(\mathscr{A},|\tau\rangle)$. In other words, we have the synchronous strategy $\mathscr{S}^\lambda$ can be expressed as a symmetric strategy within the tracial von Neumann algebra $(\mathscr{A},\tau)$. Since $A_a^{\lambda,x} \in P_\lambda \mathscr{A} P_\lambda$, we can express $A_a^{\lambda,x}$ as $P_\lambda A_a^{\lambda,x} P_\lambda$ in $\mathscr{A}$, and hence

$$\langle\tau|P_\lambda (A_a^{\lambda,x})(A_a^{\lambda,x})^{op} P_\lambda|\tau\rangle = \langle\tau|P_\lambda (A_a^{\lambda,x}) P_\lambda (B_a^{\lambda,x} P_\lambda)|\tau\rangle = \langle\tau|P_\lambda A_a^{\lambda,x} B_a^{\lambda,x} P_\lambda|\tau\rangle,$$

where the last equation follows from $(P_\lambda \mathscr{A} P_\lambda)' = \mathscr{A}' P_\lambda$ for any von Neumann algebra $\mathscr{A}$ and projector $P_\lambda \in \mathscr{A}$. Hence, showing the above claim.

Since $\int_0^\infty \tau(P_\lambda) d\lambda = \tau\left(\int_0^\infty P_\lambda d\lambda\right) = \tau(\sigma^2) = 1$, we can define the probability distribution P over $[0,\infty)$ by $P(\lambda) = \tau(P_\lambda)$. Let $C_{x,y,a,b}^\lambda$ denote the correlation set generated by the strategy $\mathscr{S}^\lambda$ for λ with $\tau(P_\lambda) \neq 0$ and any arbitrary synchronous correlation otherwise. We claim that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} |\langle\tau|\sigma Q_a^x (Q_b^y)^{op}\sigma|\tau\rangle - \int_0^\infty P(\lambda) \cdot C_{x,y,a,b}^\lambda d\lambda| \leq O(\delta^{\frac{1}{8}}). \tag{5.13}$$

To show this, we prove the following three inequalities,

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma Q_a^x(Q_b^y)^{op}\sigma|\tau\rangle-\int_0^\infty\langle\tau|(Q_a^x)^{op}(Q_b^y)^{op}P_\lambda|\tau\rangle\,d\lambda|\le O(\delta^{\frac{1}{4}}),\tag{5.14}$$

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\int_0^\infty\langle\tau|P_\lambda Q_a^x(Q_b^y)^{op}P_\lambda|\tau\rangle\,d\lambda-\int_0^\infty\langle\tau|(Q_a^x)^{op}(Q_b^y)^{op}P_\lambda|\tau\rangle\,d\lambda|\le O(\delta^{\frac{1}{8}}),\tag{5.15}$$

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\int_0^\infty\langle\tau|P_\lambda Q_a^x(Q_b^y)^{op}P_\lambda|\tau\rangle\,d\lambda-\int_0^\infty P(\lambda)\cdot C^\lambda_{x,y,a,b}d\lambda|d\lambda\le O(\delta^{\frac{1}{8}}).\tag{5.16}$$

For (5.14), since $\mathscr{S}'$ is $O(\delta^{\frac{1}{4}})$-synchronous,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma Q_a^x(Q_b^y)^{op}\sigma|\tau\rangle-\int_0^\infty\langle\tau|(Q_a^x)^{op}(Q_b^y)^{op}P_\lambda|\tau\rangle\,|d\lambda\\
&=\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma Q_a^x(Q_b^y)^{op}\sigma|\tau\rangle-\langle\tau|\sigma(Q_b^y)^{op}(Q_b^y)^{op}\sigma|\tau\rangle\,|d\lambda\\
&\le\sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma(Q_a^x-(Q_a^x)^{op})^2\sigma|\tau\rangle}\cdot\sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma(Q_b^y)^{op})^2\sigma|\tau\rangle}\le O(\delta^{\frac{1}{4}}).
\end{aligned}$$

For (5.15), by a similar calculation as above and Jensen's inequality

$$\begin{aligned}
&\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\int_0^\infty\langle\tau|P_\lambda Q_a^x(Q_b^y)^{op}P_\lambda|\tau\rangle\,d\lambda-\int_0^\infty\langle\tau|(Q_a^x)^{op}(Q_b^y)^{op}P_\lambda|\tau\rangle\,d\lambda\\
&\le\int_0^\infty\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|P_\lambda(Q_a^x-(Q_a^x)^{op})(Q_b^y)^{op}P_\lambda|\tau\rangle\,|d\lambda\\
&\le\sqrt{\int_0^\infty\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|P_\lambda(Q_a^x-(Q_a^x)^{op})^2P_\lambda|\tau\rangle\,|d\lambda}\sqrt{\int_0^\infty\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|P_\lambda((Q_b^y)^{op})^2P_\lambda|\tau\rangle\,|d\lambda}\\
&\le\sqrt{\int_0^\infty\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}\|P_\lambda Q_a^x-Q_a^xP_\lambda\|_2^2d\lambda}\cdot 1\le O(\delta^{\frac{1}{8}}),
\end{aligned}$$

where the last inequality follows from (5.5). For (5.16), by replacing $C^\lambda_{x,y,a,b}$ with $\langle\tau|P_\lambda A_a^\lambda(A_a^\lambda)^{op}P_\lambda|\tau\rangle$ for $P(\lambda)=0$,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\int_0^\infty\langle\tau|P_\lambda Q_a^x(Q_b^y)^{op}P_\lambda|\tau\rangle\,d\lambda-\int_0^\infty P(\lambda)\cdot C^\lambda_{x,y,a,b}d\lambda|d\lambda\\
&\le\int_0^\infty\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|P_\lambda Q_a^x(Q_b^y)^{op}P_\lambda|\tau\rangle\,d\lambda-\langle\tau|P_\lambda A_a^\lambda(A_a^\lambda)^{op}P_\lambda|\tau\rangle\,|d\lambda.
\end{aligned}$$

Hence, by (5.12) and $|\tau\rangle$ being tracial, $\int_0^\infty\mathbb{E}_{x\sim\mu}\sum_a\|((Q_a^x)^{op}-(A_a^{\lambda,x})^{op})P_\lambda\|_2^2d\lambda\le O(\delta^{\frac{1}{4}})$. By applying Lemma 4.3.4 for each λ (along with Lemma 4.3.3 to bound the synchronicity), (5.16) follows. By applying triangle inequality on (5.14), (5.15) and (5.15), the inequality (5.13) follows. By combining (5.10) and (5.13), we have the equation below, which concludes the proof for Theorem 5.1.1,

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma A_a^x(B_y^b)^{op}\sigma|\tau\rangle-\int_0^\infty P(\lambda)\cdot\langle\tau|P_\lambda A_a^{\lambda,x}(A_b^{\lambda,y})^{op}P_\lambda|\tau\rangle\,d\lambda|\le O(\delta^{\frac{1}{8}}).$$

For the "furthermore" part of the theorem, by combining (5.11) with the definition of P_λ

$$\begin{aligned}
&\int_0^\infty \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \|(A_a^x - A_a^{\lambda,x})P_\lambda\|_2^2 d\lambda \\
&\leq \int_0^\infty \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \|(Q_a^x - A_a^{\lambda,x})P_\lambda\|_2^2 d\lambda + \int_0^\infty \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|(Q_a^x - A_a^x)^2 P_\lambda^2|\tau\rangle\, d\lambda \\
&\leq O(\delta^{\frac{1}{4}}) + \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \langle\tau|(Q_a^x - A_a^x)^2 \left(\int_0^\infty P_\lambda d\lambda\right)|\tau\rangle \\
&= O(\delta^{\frac{1}{4}}) + \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \|(Q_a^x - A_a^x)\sigma|\tau\rangle\|_2^2 \\
&= O(\delta^{\frac{1}{4}} + \delta),
\end{aligned}$$

where the second line follows from triangle inequality, the third line follows from (5.12) and the last line follows from (5.11). □

Now we are ready to give a proof for Theorem 5.1.1. We remark that the proof follows a similar structure as [Vid22, Corollary 4.1].

Proof. We first assume the strategy $\mathscr{S} = (\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{A_a^x\})$ is symmetric, projective with $\sigma \in \mathscr{A}^+$. Let $\{P_\lambda\}$ and $\{A_a^{\lambda,x}\}$ be the set of projectors and PVMs promised by Proposition 5.2.2. Let $\mathscr{S}^\lambda = (\mathscr{L}^2(P_\lambda \mathscr{A} P_\lambda, P_\lambda\tau), \frac{1}{\sqrt{\tau(P_\lambda)}} P_\lambda|\tau\rangle, \{A_a^{\lambda,x}\})$ be the synchronous strategy from $\{A_a^{\lambda,x}\}$. Given $\lambda \in [0,\infty)$ with $\tau(P_\lambda) \neq 0$. By the "soundness property" from the assumption, for each $\lambda \in [0,\infty)$, since each $\mathscr{S}^\lambda$ is synchronous, and $(P_\lambda \mathscr{A} P_\lambda)' = \mathscr{A}' P_\lambda$, there exists a family of POVM $\{G_b^{\lambda,y}\}_{(y,b)\in\mathscr{Y}\times\mathscr{B}} \subseteq P_\lambda \mathscr{A} P_\lambda$ such that

$$\frac{1}{\tau(P_\lambda)} \mathop{\mathbb{E}}_{(x,y)\sim\rho} \sum_{a\in\mathscr{A}} \langle\tau|P_\lambda A_a^{x,\lambda} (G_{g_{xy}(a)}^{\lambda,y})^{op} P_\lambda|\tau\rangle \geq \kappa(\omega^{co}(\mathscr{G},\mathscr{S}^\lambda)),$$

where op from above is defined on the tracial von Neumann algebra $(\mathscr{A},\tau)$. For each $(y,b) \in \mathscr{B}\times\mathscr{Y}$, since $\int_0^\infty P_\lambda G_b^{\lambda,y} P_\lambda d\lambda \leq \sigma^2$, by Lemma 3.2.1, there exist $H_b^y \leq \mathbb{I}$ such that $\sigma H_b^y \sigma = \int_0^\infty P_\lambda G_b^{\lambda,y} P_\lambda d\lambda$. We see that

$$\sum_b \int_0^\infty P_\lambda G_b^{\lambda,y} P_\lambda d\lambda = \int_0^\infty P_\lambda \left(\sum_b G_b^{\lambda,y}\right) P_\lambda d\lambda = \sigma^2,$$

and hence, by a similar continuity argument as in Lemma 3.2.1, $\{H_b^y\}$ forms a valid POVM in $\mathscr{A}$. We argue that $\{H_b^y\}$ is the desired family of measurements for the corollary. We see that

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\rho} \sum_{a\in\mathscr{A}} \langle\tau|\sigma A_a^x (H_{g_{xy}(a)}^y)^{op}\sigma|\tau\rangle &= \mathop{\mathbb{E}}_{(x,y)\sim\rho} \sum_{a\in\mathscr{A}} \langle\tau|A_a^x \sigma H_{g_{xy}(a)}^y \sigma|\tau\rangle \\
&= \int_0^\infty P(\lambda) \left(\frac{1}{\tau(P_\lambda)} \mathop{\mathbb{E}}_{(x,y)\sim\rho} \sum_{a\in\mathscr{A}} \langle\tau|P_\lambda A_a^x (G_{g_{xy}(a)}^{\lambda,y})^{op} P_\lambda|\tau\rangle\right) d\lambda
\end{aligned} \tag{5.17}$$

where $P = \tau(P_\lambda)$. We claim that

$$\int_0^\infty P(\lambda) \left(\frac{1}{\tau(P_\lambda)} \mathop{\mathbb{E}}_{(x,y)\sim\rho} \sum_{a\in\mathscr{A}} \langle\tau|P_\lambda A_a^x (G_{g_{xy}(a)}^{\lambda,y})^{op} P_\lambda|\tau\rangle\right) d\lambda \geq O(\kappa(\varepsilon - \mathrm{poly}(\delta)) - \delta^{\frac{1}{8}}). \tag{5.18}$$

To show the above inequality, we prove the following two inequalities:

$$\int_0^\infty P(\lambda) \cdot \left|\frac{1}{\tau(P_\lambda)} \mathop{\mathbb{E}}_{(x,y)\sim\rho} \sum_{a\in\mathscr{A}} \langle\tau|P_\lambda (A_a^{\lambda,x} - A_a^x)(G_{g_{xy}(a)}^{\lambda,y})^{op} P_\lambda|\tau\rangle\right| d\lambda \leq O(\delta^{\frac{1}{8}}), \tag{5.19}$$

$$\int_0^\infty P(\lambda) \cdot \left(\frac{1}{\tau(P_\lambda)} \mathop{\mathbb{E}}_{(x,y)\sim\rho} \sum_{a\in\mathscr{A}} \langle\tau|P_\lambda A_a^{\lambda,x} (G_{g_{xy}(a)}^{\lambda,y})^{op} P_\lambda|\tau\rangle\right) d\lambda \geq \kappa\left(\omega^{co}(\mathscr{G},\mathscr{S}) - \mathrm{poly}(\delta)\right). \tag{5.20}$$

For (5.19)

$$\begin{aligned}
&\int_0^\infty P(\lambda)\cdot\left|\frac{1}{\tau(P_\lambda)}\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|P_\lambda(A_a^{\lambda,x}-A_a^x)(G^{\lambda,y}_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right|d\lambda\\
&\le\int_0^\infty\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\left|\langle\tau|P_\lambda(A_a^{\lambda,x}-A_a^x)(G^{\lambda,y}_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right|d\lambda\\
&\le\sqrt{\int_0^\infty\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\left|\langle\tau|P_\lambda(A_a^{\lambda,x}-A_a^x)^2P_\lambda|\tau\rangle\right|d\lambda}\cdot\sqrt{\int_0^\infty\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\left|\langle\tau|P_\lambda(G^{\lambda,y}_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right|d\lambda}\\
&\le\sqrt{\int_0^\infty\mathop{\mathbb{E}}_{x\sim\mu}\sum_{a\in\mathscr{A}}\|(A_a^{\lambda,x}-A_a^x)P_\lambda\|_2^2d\lambda}\cdot 1\le O(\delta^{\frac{1}{8}}).
\end{aligned}$$

For (5.20), by Theorem 5.1.1, $\int_0^\infty P(\lambda)\cdot\omega^{co}(\mathscr{G},\mathscr{S}^\lambda)d(\lambda)\ge\omega^{co}(\mathscr{G},\mathscr{S})-\text{poly}(\delta)$, since κ is convex

$$\begin{aligned}
\int_0^\infty P(\lambda)\cdot\left(\frac{1}{\tau(P_\lambda)}\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|P_\lambda A_a^{\lambda,x}(H^y_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right)d\lambda&\ge\kappa\left(\int_0^\infty\omega^{co}(\mathscr{G},\mathscr{S}^\lambda)\right)\\
&\ge\kappa\left(\omega^{co}(\mathscr{G},\mathscr{S})-\text{poly}(\delta)\right).
\end{aligned}$$

Hence, by the triangle inequality, (5.18) follows from (5.19) and (5.20). Combining (5.17) and (5.18),

$$\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|\sigma A_a^x(H^y_{g_{xy}(a)})^{op}\sigma|\tau\rangle\ge O\left(\kappa(\varepsilon-\text{poly}(\delta))-\delta^{\frac{1}{8}}\right),$$

which concludes the corollary for the special case claimed above.

For the general case where $\mathscr{S}=(\mathscr{L}^2(\mathscr{A},\tau),\sigma|\tau\rangle,\{A_a^x\},\{(B_a^x)^{op}\})$ is a tracially embeddable strategy. By Lemma 4.3.3, the symmetric strategy $\mathscr{S}'=(\mathscr{L}^2(\mathscr{A},\tau),\sigma|\tau\rangle,\{A_a^x\})$ is an $O(\delta)$-synchronous strategy, with

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma A_a^x(B_b^y)^{op}\sigma|\tau\rangle-\langle\tau|\sigma A_a^x(A_b^y)^{op}\sigma|\tau\rangle|\le O(\sqrt{\delta}),$$

which implies $\omega^{co}(\mathscr{G},\mathscr{S}')\ge\omega^{co}(\mathscr{G},\mathscr{S})-O(\text{poly}(\delta))$. By the special case proven above, there exists a family of POVMs $\{H_{g^y_{xy}(a)}\}$ such that

$$\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|\sigma A_a^x(H^y_{g_{xy}(a)})^{op}\sigma|\tau\rangle\ge O(\kappa(\omega^{co}(\mathscr{G},\mathscr{S})-\text{poly}(\delta))-\delta^{\frac{1}{8}}),\tag{5.21}$$

which concludes the proof for the general case. □

5.3 Consequences of the rounding theorem on synchronous games

The rounding theorem leads to several important corollaries in this thesis. We discuss one such corollary related to synchronous games in this section. A synchronous game $\mathscr{G}$ is called *c-balanced* if there exists some constant $c\in[0,1]$ such that $c\cdot\mu_x(x)\le\mu(x,x)$ and $c\cdot\mu_y(x)\le\mu(x,x)$ for all $x\in\mathscr{X}$. In other words, the synchronous question pair will always appear with at least probability c on the marginal distribution. As seen later on in this thesis, this condition can always be enforced when considering non-local games in the interactive proof setting.

We have the following lemma about any correlations which are near perfect for any balanced game.

Lemma 5.3.1 (Almost perfect correlation implies almost synchronous correlation)**.** *Let $\mathscr{G}=(\mathscr{X},\mathscr{A},\mu,D)$ be a c-balance synchronous game, and for model $t\in\{*,qc\}$, let $\mathscr{S}=(\mathscr{L}^2(\mathscr{A},\tau),|\tau\rangle,\{A_a^x\},\{(B_b^y)^{op}\})$ be a tracially embeddable strategy such that $\omega(\mathscr{G},\mathscr{S})\ge1-\varepsilon$. Then $\delta_{sync}(\mu,C)\le\frac{\varepsilon}{c}$ and there exists a symmetric and projective strategy $\mathscr{S}^{sym}=(\mathscr{L}^2(\mathscr{A},\tau),\sigma|\tau\rangle,\{P_a^x\})$ defined on the same Hilbert space as $\mathscr{S}$ such that*

$$\omega(\mathscr{G},\mathscr{S}^{sym})\ge1-\varepsilon-\left(\frac{\varepsilon}{c}\right)^{\frac{1}{4}}.$$

Proof. For any correlation C, $\mathcal{G}$ being synchronous and $\omega(\mathcal{G},\mathcal{S}) \geq 1-\varepsilon$ implies

$$\sum_{x\in\mathcal{X}} \mu(x,x) \sum_{a\neq b} \langle\tau|\sigma A_a^x B_b^x \sigma|\tau\rangle \leq \varepsilon.$$

By the c-balanced condition,

$$c\cdot\left(\mathop{\mathbb{E}}_{x\sim\mu_x} \sum_{a\neq b} \langle\tau|\sigma A_a^x B_b^x \sigma|\tau\rangle\right) \leq \sum_{x,a\neq b} \mu(x,x)\langle\tau|\sigma A_a^x B_b^x \sigma|\tau\rangle \leq \varepsilon$$
$$c\cdot\left(\mathop{\mathbb{E}}_{x\sim\mu_y} \sum_{a\neq b} \langle\tau|\sigma A_a^x B_b^x \sigma|\tau\rangle\right) \leq \sum_{x,a\neq b} \mu(x,x)\langle\tau|\sigma A_a^x B_b^x \sigma|\tau\rangle \leq \varepsilon.$$

This shows that $\delta_{\text{sync}}(\mu,\mathcal{S}) \leq \frac{\varepsilon}{c}$. For the second part of the lemma, by Corollary 4.3.6, there exist a symmetric strategy $\mathcal{S}^{\text{sym}} = (\mathcal{L}^2(\mathcal{A},\tau),\sigma|\tau\rangle,\{P_a^x\})$ such that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a\in A} |\langle\tau|\sigma A_a^x (B_b^y)^{op}\sigma|\tau\rangle - \langle\tau|\sigma P_a^x (P_b^y)^{op}\sigma|\tau\rangle| \leq O\left(\left(\frac{\varepsilon}{c}\right)^{\frac{1}{4}}\right). \tag{5.22}$$

By the triangle inequality, $|\omega(\mathcal{G},\mathcal{S}^{\text{sym}}) - \omega(\mathcal{G},\mathcal{S})| \leq O((\frac{\varepsilon}{c})^{\frac{1}{4}})$, and hence the lemma follows accordingly. □

When combine with Theorem 5.1.1 applied to the correlations generated by $\mathcal{S}^{\text{sym}}$, we can conclude the following corollary.

Corollary 5.3.2 (Almost perfect correlation implies almost synchronous correlation)**.** *For any c-balanced game $\mathcal{G}$ and for $t \in \{*, co\}$*

$$\omega^t(\mathcal{G}) - \omega_s^t(\mathcal{G}) \leq \text{poly}\left(\varepsilon, \frac{1}{c}\right).$$

Part III

Interactive proof systems and the compression theorem

CHAPTER 6

Interactive proof system and the uncompubility of MIP^* and $\mathsf{MIP}^{\mathrm{co}}$

We start this chapter by introducing the complexity classes MIP^* and $\mathsf{MIP}^{\mathrm{co}}$, or multiplayer interactive proof systems with entanglement, the complexity class which is central to this thesis. We will also introduce the compression map for these interactive proof systems, and combine it with the compression paradigm introduced in Chapter 2 to show the uncomputability of both MIP^* and $\mathsf{MIP}^{\mathrm{co}}$.

6.1 Interactive proof systems

In this section, we define the complexity classes MIP^* and $\mathsf{MIP}^{\mathrm{co}}$ more formally. Recall from the introduction that MIP stands for *multi-prover interactive proof system*, the class of languages L decidable by a probabilistic polynomial-time classical verifier when given (classical) interacting with computationally unbounded and non-communicating provers (i.e. the provers cannot talk to each other). In this model, the verifier can interact with multiple provers and may interact with the provers through multiple rounds of interactions. The verifier might adapt his questions based on the previous round of interactions and may leverage the lack of communication between the provers to "cross-interrogate" them. If $z \in \mathsf{L}$, the verifier can formulate an interaction such that the prover can provide enough evidence to convince the verifier to accept with probability 1. On the other hand, if $z \notin \mathsf{L}$, the verifier can also formulate an interaction which ensures that the provers can only convince the verifier with probability at most $\frac{1}{2}$ to accept the given instance[1]). As shown in [BFL91], the computational power of MIP is equivalent to non determinate exponential time (NEXP), the exponential variant of NP, and this can be achieved with just a one-round interaction with two provers.

In this thesis, we consider two variants of MIP where the provers are still non-communicating, but are now given shared entanglement when communicating with provers. We denote the variant where the provers share the tensor product model of entanglement as MIP^* and the commuting operator model of entanglement as $\mathsf{MIP}^{\mathrm{co}}$. In this thesis, we focus on the variant of MIP^* and $\mathsf{MIP}^{\mathrm{co}}$ with two provers and one-round of interactions since this is sufficient for proving lower bounds. The two provers one-round MIP^* protocol (resp. $\mathsf{MIP}^{\mathrm{co}}$) is denoted as $\mathsf{MIP}^*(2,1)$ (resp. $\mathsf{MIP}^{\mathrm{co}}(2,1)$) in the literature, and for simplicity of notation, unless otherwise specified, we drop $(2,1)$ when discussing $\mathsf{MIP}^*(2,1)$ (resp. $\mathsf{MIP}^{\mathrm{co}}(2,1)$). In this thesis, we also work with MIP with completeness 1 and soundness $\frac{1}{2}$, meaning that there exists a verifier behaviour such that if $z \in \mathsf{L}$, then the verifier accepts with probability 1, and if $z \notin \mathsf{L}$, then the verifier accepts with probability at most $\frac{1}{2}$. The completeness 1, soundness $\frac{1}{2}$ MIP^* protocol

[1] The original formulation is $\geq \frac{2}{3}$ if $x \in \mathsf{L}$ and $\leq \frac{1}{3}$ otherwise. However, we remark this is equivalent to the formulation given due to sequential repetition.

(resp. $\mathsf{MIP}^{\mathrm{co}}$)) is denoted as $\mathsf{MIP}^*_{1,\frac{1}{2}}$ (resp. $\mathsf{MIP}^{\mathrm{co}}_{1,\frac{1}{2}}$) in the literature, and similarly, we drop the subscript for the simplicity of notation. For a more general definition on MIP^*, we refer the readers to [VW16, Section 6.1]. We formally give the definitions for MIP^* and MIP^{co} used in this thesis below.

Definition 6.1.1 (Multi-prover proof system with entanglement)**.** *Let $t \in \{*, co\}$. A language $\mathtt{L}$ is in MIP^t if there exist a pair of probabilistic (potentially multi-input) Turing machines $(\mathtt{Q},\mathtt{D})$ such that $\mathsf{TIME}_{\mathtt{Q}}(z) = \mathsf{TIME}_{\mathtt{D}}(z) = O(\mathrm{poly}(|z|))$ for all $z \in \{0,1\}^*$. Furthermore, there exists an infinite sequence of games $\mathcal{G}_z = (\mathcal{X}_z, \mathcal{A}_z, \mu_z, D_z)$ indexed by $z \in \{0,1\}^n$ and two increasing polynomial functions $x(n), a(n) : \mathbb{N} \to \mathbb{N}$ with $\mathcal{X}_z = \{0,1\}^{x(|z|)}$ and $\mathcal{A}_z = \{0,1\}^{a(|z|)}$, such that*

- ***(Uniformity)*** *$\mathtt{Q}(z, sample)$ outputs a sample from the distribution μ_z, and $\mathtt{D}(z,x,y,a,b) = D_z(x,y,a,b)$ for all $(x,y,a,b) \in \mathcal{X}_z^2 \times \mathcal{A}_z^2$.*
- ***(Completeness)*** *If $z \in \mathcal{L}$, then $\omega^t(\mathcal{G}_z) = 1$.*
- ***(Soundness)*** *If $z \notin \mathcal{L}$, then $\omega^t(\mathcal{G}_z) \leq \frac{1}{2}$.*

The above definition is similar to the one given in [JNV+22a, Definition 5.29]. Intuitively, the pair of Turing machines $(\mathtt{Q},\mathtt{D})$ completely specifies the behaviour for the verifier. We remark that given the pair of Turing machines $(\mathtt{Q},\mathtt{D})$, it is hard to extract the exact description of $\mathcal{G}_z = (\mathcal{X}_z, \mathcal{A}_z, \mu_z, D_z)$ for a particular instance $z \in \{0,1\}^*$ by the definition above. However, as seen in the next subsection, we can hardcode $(\mathtt{Q},\mathtt{D})$ to run other computational procedures in a way such that the description for $\mathcal{G}_z = (\mathcal{X}_z, \mathcal{A}_z, \mu_z, D_z)$ can be properly extracted. As observed in [CHT+04], for $t \in \{*, co\}$, the complexity class MIP^t is complete with respect to the following decision problem.

Definition 6.1.2 ($(1,\frac{1}{2})$ non-local game value problem)**.** *For $t \in \{*, co\}$, the $(1,\frac{1}{2})$ t non-local game value problem is a decision problem defined by the following two sets.*

- $\mathsf{L}^{\mathsf{MIP}^t}_{yes} = \{\langle \mathcal{G} \rangle \,|\, \omega^t(\mathcal{G}) = 1\}$.
- $\mathsf{L}^{\mathsf{MIP}^t}_{no} = \{\langle \mathcal{G} \rangle \,|\, \omega^t(\mathcal{G}) \leq \frac{1}{2}\}$.

For clarity, we refer to the $(1,\frac{1}{2})$ $*$ non-local game value problem as the $(1,\frac{1}{2})$ tensor product value problem, and the $(1,\frac{1}{2})$ co non-local game value problem as the $(1,\frac{1}{2})$ commuting operator value problem.

There is a fundamental difference between the complexity class $\mathsf{MIP}^*/\mathsf{MIP}^{\mathrm{co}}$ to non-local games. In essences, these complexity classes are a variant of **succinct representation** of non-local games. Based on this, we give a notion of an "uniform succinct problem instance" for interactive proof systems similarly to Definition 2.3.1.

Definition 6.1.3 (Uniform verifier sequence)**.** *Let $t \in \{*, co\}$, and let $\mathcal{G}_n = (\mathcal{X}_n, \mathcal{A}_n, \mu_n, D_n)$ be a sequence of games. A verifier sequence $\mathcal{V} = (\mathtt{Q},\mathtt{D})$ is a pair of Turing machines such that $\mathtt{Q}(n, sample)$ outputs a sample from the distribution μ_n, and $\mathtt{D}(z,x,y,a,b) = D_n(x,y,a,b)$ for all $(x,y,a,b) \in \mathcal{X}_n^2 \times \mathcal{A}_n^2$. Furthermore, we say that $\mathcal{V}$ runs in $O(\mathbf{f}(n))$ time if*

$$\mathsf{TIME}_{\mathtt{Q}} = \mathsf{TIME}_{\mathtt{D}} = O(\mathbf{f}(n)).$$

In the above definition, although $\mathcal{V}$ is defined using two separate Turing machines, this is equivalent to the single Turing machine succinct description given in Definition 2.3.1 by adding one extra input. In comparison to Definition 2.3.1, the uniform verifier sequence allows the verifier to extra information from both the sampling distribution and the verification function.

6.2 Compression theorem for interactive proof system

We introduce the gap-compression theorem for non-local games in this chapter and show how the theorem can be used to lower-bound the complexity of MIP^* and $\mathsf{MIP}^{\mathrm{co}}$ in this section. Before we continue, we point out that the succinct compression map used below is actually not based on the uniform verifier sequence given in Definition 6.1.3, but rather a more restrictive uniform succinct description known as *conditionally linear verifier*. Intuitively, this is a uniform sequence on a set of games with a specific input distribution (the kth level conditional linear distribution), as well as extra input clauses on $\mathtt{Q}$ for the verifier to extract additional information on said input distribution. Since defining this verifier requires additional background information, and it is not required to understand the uncomputability result, we define conditionally linear verifier in more detail in Chapter 7. For now, one should think of a synchronous conditional verifier that takes an extra "level" parameter" k which dictates the structure of the sampling distribution for the sequences of games, and in addition to it being a uniform verifier sequence, it also has the following properties:

- The runtime of the sampling procedure and the decision procedure are $O(f(n,\mathrm{poly}(k)))$ instead (i.e. the runtime has an extra dependency on $\mathrm{poly}(k)$).
- For any $k' \leq k$, any k'-level conditional linear verifier is also a k-level conditional linear verifier.

In this chapter, one can safely assume that the level parameter k is a constant (however, as seen later in this thesis, the constant parameter becomes important when proving the compression theorem). Furthermore, for $t \in \{*, co\}$, we use $k-\mathsf{CLMIP}^t$ to denote the complexity class described by the $(1, \frac{1}{2})$ non-local game value problem for any games with a constant-level conditional linear distribution. As a reminder, the notion of an oracularizable strategy is defined in Section 4.4.

Theorem 6.2.1 (Gap compression for non-local games)**.** *For all constants $\alpha, k \in \mathbb{N}$, there exists an algorithm $\mathtt{Gapcompress}_{\alpha,k}$ that takes the input a pair of Turing machines $(\mathtt{Q},\mathtt{D})$. $\mathtt{Gapcompress}_{\alpha,k}$ outputs a tuple of Turing machines $\mathscr{V}^{Comp} = (\mathtt{Q}^{Comp}, \mathtt{D}^{Comp})$ such that the following holds:*

There exists an integer $\gamma = O(poly(\alpha, k))$ such that, for models $t \in \{*, co\}$

1. *(Runtime): $\mathsf{TIME}_{\mathtt{Gapcompress}_{\alpha,k}}(\mathtt{Q},\mathtt{D}) = O(\mathrm{poly}(\alpha), |\mathtt{Q}|, |\mathtt{D}|)$.*
2. *(Independence of the sampler): The Turing machine $\mathtt{Q}^{Comp}$ only depends on the parameter α and k and is a sampler for a synchronous 7-th level CL verifier. The Turing machine $\mathtt{D}^{Comp}$ is a decider for a synchronous 7-th level CL verifier. $\mathscr{V}^{Comp}$ is a synchronous 7-th level CL verifier sequence for an infinite sequence of games $\mathscr{G}^{Comp} = \{\mathscr{G}_n^{Comp}\}_{n\in\mathbb{N}}$.*
3. *(Complexity bounds for the output) $\mathscr{V}^{Comp}$ has sample complexity and verification complexity of $O(\log(n)^\gamma)$.*

Furthermore, if the input $\mathscr{V} = (\mathtt{Q},\mathtt{D})$ is a synchronous k'th level CL verifier for the infinite sequence of synchronous games $\mathscr{G} = \{\mathscr{G}_n\}_{n\in\mathbb{N}}$ for some constant $k' < k$, and there exists a constant $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$

$$\max\{\mathsf{TIME}_{\mathtt{Q}}, \mathsf{TIME}_{\mathtt{D}}\} \leq n^\alpha. \tag{6.1}$$

Then there exists a constant $n_0^{Comp} = \mathrm{poly}(\gamma, n_0)$ such that for all $n \geq n_0^{Comp}$

- *(Completeness) If there exists a perfect oracularizable strategy for $\mathscr{G}_n$ in model t, then there exists a perfect oracularizable strategy for $\mathscr{G}_n^{Comp}$ in model t.*
- *(Soundness)*

$$\omega^t(\mathscr{G}_n) \leq \frac{1}{2} \Longrightarrow \omega^t(\mathscr{G}_n^{Comp}) \leq \frac{1}{2}.$$

By clause 2 of the above theorem, for any α, k, the output for the algorithm $\texttt{Gapcompress}_{\alpha,k}$ will always be a synchronous 7-th level CL verifier (even if the input $(\mathtt{Q},\mathtt{D})$ might not be a valid uniform verifier sequences). Hence, as a corollary, Theorem 6.2.1 shows the following

Corollary 6.2.2. *For $t \in \{*, co\}$ and constant $k \in \mathbb{N}$ such that $k \geq 7$, the complexity class k-CLMIP^t is a succinctly weakly compressible problem.*

We remark that because of the parallel repetition theorem introduced later in this thesis (Proposition 7.4.4), the soundness property from Theorem 6.2.1 can be controlled to any constant $c \in (0,1)$. Before we continue, we first give the definition for a trivial synchronous accepting/rejecting game for both MIP^* and $\mathsf{MIP}^{\mathrm{co}}$ below.

Definition 6.2.3 (Accepting/Rejecting game)**.** *We define the synchronous accepting game to be the game $\mathscr{G}^{accept} = (\mathscr{X}, \mathscr{A}, \mu, D^{accept})$, where $\mathscr{X} = \{\star\}$ and $\mathscr{A} = \{0,1\}^*$,*

$$D^{accept}(\star, \star, a, b) = \delta_{a,0}\delta_{b,0}$$

for all $a, b \in \{0,1\}^*$ (i.e. the prover automatically wins if both provers return 0). We define the synchronous rejecting game to be the game $\mathscr{G}^{reject} = (\mathscr{X}, \mathscr{A}, \mu, D^{reject})$, where $\mathscr{X} = \{0,1\}$, $\mu(1,0) = \mu(0,1) = \frac{1}{3}$, $\mu(0,0) = \mu(1,1) = \frac{1}{6}$, and $\mathscr{A} = \{0,1\}^*$,

$$D^{reject}(x, y, 0, 0) = \delta_{x,y},$$

and $D^{reject}(x, y, 0, 0) = 0$ for all other $a, b \in \{0,1\}^*$.

For model $t \in \{*, co\}$, we have $\omega^t(D^{\text{accept}}) = 1$ and $\omega^t(D^{\text{reject}}) = \frac{1}{3}$. Both games are trivially samplable via a CL distribution for any level and always computable in constant time. Intuitively, this is the trivial synchronous game which is in the yes/no case for the $(1, 1/2)$ tensor product/commuting operator value problem defined in Definition 6.1.2 for both $t \in \{*, co\}$. This shows that for $t \in \{*, co\}$, both $\mathsf{L}^{k\text{-}\mathsf{CLMIP}^t}_{\text{yes}}$ and $\mathsf{L}^{k\text{-}\mathsf{CLMIP}^t}_{\text{no}}$ are non-empty. Based on Corollary 6.2.2 we show the main theorem of this thesis.

6.2.1 $\mathsf{MIP}^{\mathrm{co}} = \mathsf{coRE}$

In order to show the main theorem of this thesis, we first show the following lemma.

Lemma 6.2.4. $\mathsf{MIP}^{\mathrm{co}} \subseteq \mathsf{coRE}$.

The above lemma also shows that $\mathsf{L}^{k\text{-}\mathsf{CLMIP}^{co}}_{\text{yes}} \subseteq \mathsf{L}^{\mathsf{MIP}^{\mathrm{co}}}_{\text{yes}} \in \mathsf{coRE}$. To show this lemma, we recall the well-known hierarchy algorithm due to Navascués, Pironion, and Acín [NPA08]. We summarize the functionality of the NPA hierarchy below.

Theorem 6.2.5 (Output for the NPA hierarchy [NPA08])**.** *For any integer $n \in \mathbb{N}$ and $\mathscr{G} = (\mathscr{X}, \mathscr{A}, \mu, D)$, there exists a terminating algorithm $\texttt{NPAhierarchy}$ such that $\texttt{NPAhierarchy}(\mathscr{G}, n) = \varepsilon_n \in \mathbb{R}$ with*

$$\lim_{n\to\infty} \varepsilon_n = \omega^{co}(\mathscr{G})$$

For another reference about the NPA hierarchy, we refer to the excellent lecture notes by John Watrous [Wat22]. Based on Theorem 6.2.5, for $\delta \in [0,1]$ we define the $\texttt{Searchfromabove}_\delta$ algorithm below.

Input: $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$
If the input $\mathcal{G}$ is not a valid description of the game, terminate and **return** "Error";
Set $n = 1$
while *True* **do**
    Compute $\varepsilon_n = \texttt{NPAhierarchy}(\mathcal{G}, n)$
    if $\varepsilon_n < \delta$ **then**
        Return False ;
    n = n+1 ;
end

Pseudocode 5: The description for $\texttt{searchfromabove}_\delta$.

$\texttt{Searchfromabove}_\delta$ gives a systematic way to generate a sequence of upper bounds for $\omega^{co}(\mathcal{G})$ to check whether $\omega^{co}(\mathcal{G}) \leq \delta$. By definition, the algorithm runs forever if $\omega^{co}(\mathcal{G}) \geq \delta$. Hence the algorithm $\texttt{Searchfromabove}_1$ directly implies Lemma 6.2.4.

By combining Corollary 6.2.2, Theorem 2.2.5 and the fact that $\mathsf{L}_{\text{yes}}^{k\text{-CLMIP}^{co}} \in \mathsf{coRE}$ this shows that k-$\mathsf{CLMIP}^{co} = \mathsf{coRE}$ and hence $\mathsf{coRE} \in \mathsf{MIP}^{\text{co}}$. This, along with Lemma 6.2.4, shows the main theorem of this thesis.

Corollary 6.2.6. $\mathsf{MIP}^{\text{co}} = \mathsf{coRE}$

Input: Integer n.
Run 🚢 for n steps. If 🚢 halts in the given steps, return $\mathcal{G}^{\text{reject}}$.
Compute the description of $\mathcal{V}$.
Simulate the following program describe by step 5 and 6 for $\max\{0, n - C_0\}$ steps:
    Compute the description of $\mathcal{G}_{C_0}$ of $\mathcal{G}_{C_0} = \mathcal{V}(C_0)$, the C_0th game of the CL verifier $\mathcal{V}$.
    Run $\texttt{Searchfromabove}_1$ specified in Pseudocode 5 with $\mathcal{G}_{C_0}$ as the input.
If the above subroutine halts in $\max\{0, n - C_0\}$ steps, return $\mathcal{G}^{\text{accept}}$.
Apply $\texttt{Gapcompress}_{\alpha,7}$ on the verifier $\mathcal{V}$ to obtain $\mathcal{V}^{\text{comp}}$.
Compute and **return** $\mathcal{G}_{n+1}^{\text{comp}} = \mathcal{V}^{\text{comp}}(n+1)$.

Pseudocode 6: The description for $\mathcal{V}$ which can be used to show that $\mathsf{coRE} \subseteq k$-$\mathsf{CLMIP}^{co}$. 🚢 is the instance of the halting problem for the reduction.

We remark that if we specifically tailored Pseudocode 3 specifically for k-CLMIP^{co}, we have the following pseudocode, where in the pseudocode below, $\mathcal{G}_{C_0}$ is represented as Pseudocode 6 with C_0 being hard coded in. Pseudocode 6 is also a more refined version of [MNY22, Pseudocode 4].

6.2.2 $\mathsf{MIP}^* = \mathsf{RE}$

Similarly to the previous section, we need to first show the following lemma.

Lemma 6.2.7. $\mathsf{MIP}^* \subseteq \mathsf{RE}$

By a similar reason as above, the above lemma shows that $\mathsf{L}_{\text{no}}^{k\text{-CLMIP}^*} \in \mathsf{coRE}$. We show the above lemma by recalling a well-known fact about the quantum tensor correlations in the literature.

Fact 6.2.8 (Discretization of quantum tensor correlations)**.** *For any integer $n \in \mathbb{N}$ and $\varepsilon > 0$, there exists a terminating algorithm to search over C_q^n to generate a finite subset $C_\varepsilon^n \subseteq C_q^n$ such that for all correlations $C \in C_q^n$, there exists a correlation $C' \in C_\varepsilon^n$ such that*

$$\sum_{x,y,a,b} |C_{x,y,a,b} - C'_{x,y,a,b}| \leq \varepsilon.$$

Based on Fact 6.2.8, for $\delta \in [0,1]$, we define the algorithm $\texttt{searchfrombelow}_\delta$ below.

Input: $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$
Set $n = 1$
while *True* **do**
  Compute the finite subset $C^n_{1/n}$ defined by Fact 6.2.8.
  for $C' \in C^n_{\frac{1}{n}}$ **do**
    Compute $\varepsilon_n = \mathbb{E}_{(x,y)\sim\mu} \sum_{(a,b)\in\mathcal{A}^2} C'_{x,y,a,b} D(x,y,a,b)$
    if $\varepsilon_n > \delta$ **then**
      Return True ;
  end
  n = n+1 ;
end

Pseudocode 7: The description for $\texttt{searchfrombelow}_\varepsilon$ for $\varepsilon \in [0,1)$.

Intuitively, $\texttt{searchfrombelow}_\delta$ gives a systematic way to search through the correlation set C_q to find a strategy which validates that $\omega^*(\mathcal{G}) > \delta$ (which might be impossible depending on $\mathcal{G}$).

We are now ready to show Lemma 6.2.7.

Proof. To show that $\mathsf{MIP}^* \subseteq \mathsf{RE}$, we need to show that the algorithm above halts in the Yes case for the non-local game value problem. Hence, let $\mathcal{G}$ be a game such that $\omega^*(\mathcal{G}) = 1$ and consider the algorithm $\texttt{searchfrombelow}_{0.5}$ running on $\mathcal{G}$. Combining the definition ω^* and $C_q = \bigcup_{n\in\mathbb{N}^+} C^n_q$, we see that there must exist some $n \in \mathbb{N}$ and $C \in C^n_q$ such that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{(a,b)\in\mathcal{A}^2} C_{x,y,a,b} D(x,y,a,b) \geq 0.75.$$

Hence by the definition, there exists a constant $C' \in C^n_{\frac{1}{n}}$ such that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{(a,b)\in\mathcal{A}^2} (C_{x,y,a,b} - C'_{x,y,a,b}) D(x,y,a,b) \leq \sum_{x,y,a,b} |C_{x,y,a,b} - C'_{x,y,a,b}| \leq \frac{1}{n},$$

where the inequality follows since $\mu(x,y), D(x,y,a,b) \leq 1$. Combining the two inequalities yields

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{(a,b)\in\mathcal{A}^2} C'_{x,y,a,b} D(x,y,a,b) > 0.5.$$

This completes the proof of the lemma. □

By a similar proof as the above lemma, we see that $\texttt{searchfrombelow}_\varepsilon(\mathcal{G})$ terminates iff $\omega^*(\mathcal{G}) > \varepsilon$. We remark that $\texttt{searchfrombelow}_\varepsilon$ does not work for the set of commuting operator correlations, as there is no way to discretize the set of commuting operator strategies based on the dimension of the Hilbert space.

The algorithm $\texttt{searchfrombelow}_{0.5}$ is precisely the algorithm needed to show Lemma 6.2.4. By combining Corollary 6.2.2, Theorem 2.2.4 and the fact that $\mathsf{L}^{k\text{-}\mathsf{CLMIP}^*}_{\text{no}} \in \mathsf{coRE}$, this shows that k-$\mathsf{CLMIP}^* = \mathsf{RE}$ and hence $\mathsf{RE} \in \mathsf{MIP}^*$. This, along with Lemma 6.2.7, shows the main theorem of this thesis.

Corollary 6.2.9. $\mathsf{MIP}^* = \mathsf{RE}$.

In a similar vein as Corollary 6.2.6, we remark that Corollary 6.2.9 can be proven using Pseudocode 8 below.

Input: Integer n.
Run $\textcircled{\scriptsize M}$ for n steps. If $\textcircled{\scriptsize M}$ halts in the given steps, **return** $\mathcal{G}^{\text{accept}}$.
Compute the description of $\mathscr{V}$.
Simulate the following program describe by step 5 and 6 for $\max\{0, n - C_0\}$ steps:
    Compute the description of $\mathcal{G}_{C_0}$ of $\mathcal{G}_{C_0} = \mathscr{V}(C_0)$, the C_0th game of the CL verifier $\mathscr{V}$.
    Run $\texttt{Searchfrombelow}_{0.5}$ specified in Pseudocode 7 with $\mathcal{G}_{C_0}$ as the input
If the above subroutine halts in $\max\{0, n - C_0\}$ steps, return $\mathcal{G}^{\text{reject}}$
Apply $\texttt{Gapcompress}_{\alpha,7}$ on the verifier $\mathscr{V}$ to obtain $\mathscr{V}^{\text{comp}}$.
Compute and **return** $\mathcal{G}^{\text{comp}}_{n+1} = \mathscr{V}^{\text{comp}}(n+1)$.

Pseudocode 8: The description for $\mathscr{V}$ which generates the game sequences $\{\mathcal{G}_n\}_{n\in\mathbb{N}}$ for the proof of $\mathsf{RE} \subseteq \mathsf{MIP}^*$. $\textcircled{\scriptsize M}$ is the instance of the halting problem for the reduction.

6.2.3 Finding an explicit separation between C_q and C_{qc} is RE-complete

Finally, we give the last application of Theorem 6.2.1, which is to show that finding a Bell test between the tensor product model and commuting operator model is RE-complete. In order to show this, we first give a formal definition of this problem using a decision problem.

Definition 6.2.10 (The δ-Bell test separation decision problem)**.** *Given a constant $\delta \in (0,1)$, the δ-Bell test separation decision problem $\mathsf{D}_{\delta\text{-bell}}$ is defined by the following two sets.*

- $\mathsf{L}^{\mathsf{D}_{\delta\text{-bell}}}_{yes} = \{\langle\mathcal{G}\rangle \,|\, \omega^*(\mathcal{G}) = \omega^{co}(\mathcal{G})\}$.
- $\mathsf{L}^{\mathsf{D}_{\delta\text{-bell}}}_{no} = \{\langle\mathcal{G}\rangle \,|\, |\omega^*(\mathcal{G}) - \omega^{co}(\mathcal{G})| \geq \delta\}$.

The above problem is already known to be RE-hard prior to this work using the algorithm $\texttt{Searchsamevalue}_\delta$, which we give in Pseudocode 9 below for completeness. As shown in [JNV+22a, Theorem 12.10] the set $\mathsf{L}^{\mathsf{D}_{\frac{1}{2}\text{-bell}}}_{\text{no}}$ is non-empty, as there exists a (synchronous 12-th level CL samplable game) such that $|\omega^*(\mathcal{G}) - \omega^{co}(\mathcal{G})| \geq \frac{1}{2}$. Given this, we have the following theorem for the complexity of the $\frac{1}{2}$-Bell test separation problem.

Theorem 6.2.11. *The $\frac{1}{2}$-Bell test separation problem is* RE*-complete.*

The above theorem follows from realizing that Theorem 6.2.1 also shows that the $\frac{1}{2}$-Bell test separation problem is weakly compressible for synchronous games which are 7-th level CL samplable with the "yes" case consisting of games $\mathcal{G}$ such that $\omega^*(\mathcal{G}) = \omega^{co}(\mathcal{G}) = 1$, and the "no" case consisting of games $\mathcal{G}$ such that $\omega^{co}(\mathcal{G}) = 1$ and $\omega^*(\mathcal{G}) \leq \frac{1}{2}$ (where the $\mathcal{G}$ in both cases are synchronous 12-th level CL sample games). Since the argument follows similarly as the one given in Section 6.2.2, we do not give the details here. We remark that the above proof can be easily changed to any $\delta \in (0,1)$ since the soundness condition from Theorem 6.2.1 can be changed to hold for any $\delta \in (0,1)$.

We remark that instead of using Theorem 2.2.4, Theorem 6.2.11 can also be proven by considering the following uniform verifier sequence defined in Pseudocode 9. By using a similar argument as the proof of Theorem 2.2.4, one can infer that $\omega^*(\mathcal{G}_{C_0}) < 0.5$ by using line 5 of Pseudocode 10 and $\omega^*(\mathcal{G}_{C_0}) = 1$ by using line 6 of Pseudocode 10.

Input: $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$
Set $n = 1, \varepsilon_1^{\text{lower}} = 0$
while *True* **do**
  Compute the finite subset $C^n_{1/n}$ defined by Fact 6.2.8.
  for $C' \in C^n_{\frac{1}{n}}$ **do**
    Compute $\varepsilon' = \mathbb{E}_{(x,y)\sim\mu} \sum_{(a,b)\in\mathcal{A}^2} C'_{x,y,a,b} D(x,y,a,b)$
    if $\varepsilon' > \varepsilon_n^{lower}$ **then**
      $\varepsilon_n^{\text{lower}} = \varepsilon'$;
  end
  Compute $\varepsilon_n^{\text{upper}} = \texttt{NPAhierarchy}(\mathcal{G}, n)$
  if $|\varepsilon_n^{upper} - \varepsilon_n^{lower}| \leq \min(\delta - \frac{1}{n}, 0)$ **then**
    Return True ;
  $n = n + 1$;
  $\varepsilon_n^{\text{lower}} = \varepsilon_{n-1}^{\text{lower}}$;
end

Pseudocode 9: The description for $\texttt{Searchsamevalue}_\delta$.

Input: Integer n.
Run 🚗 for n steps. If 🚗 halts in the given steps, **return** $\mathcal{G}^{\text{accept}}$.
Compute the description of $\mathcal{V}$.
Simulate the following program describe by step 5 and 6 for $\max\{0, n - C_0\}$ steps:
  Compute the description of $\mathcal{G}_{C_0}$ of $\mathcal{G}_{C_0} = \mathcal{V}(C_0)$, the C_0th game of the CL verifier $\mathcal{V}$.
  Run $\texttt{Searchfrombelow}_{0.5}$ specified in Pseudocode 7 with $\mathcal{G}_{C_0}$ as the input
If the above subroutine halts in $\max\{0, n - C_0\}$ steps, return $\mathcal{G}^{\text{reject}}$
Simulate the following program describe by step 9 and 10 for $\max\{0, n - C_0\}$ steps:
  Compute the description of $\mathcal{G}_{C_0}$ of $\mathcal{G}_{C_0} = \mathcal{V}(C_0)$, the C_0th game of the CL verifier $\mathcal{V}$.
  Run $\texttt{Searchfromabove}_1$ specified in Pseudocode 5 with $\mathcal{G}_{C_0}$ as the input.
If the above subroutine halts in $\max\{0, n - C_0\}$ steps, return $\mathcal{G}^{\text{accept}}$.
Apply $\texttt{Gapcompress}_{\alpha,7}$ on the verifier $\mathcal{V}$ to obtain $\mathcal{V}^{\text{comp}}$.
Compute and **return** $\mathcal{G}^{\text{comp}}_{n+1} = \mathcal{V}^{\text{comp}}(n+1)$.

Pseudocode 10: An alternative game sequence for the proof of Theorem 6.2.11. 🚗 is the instance of the halting problem for the reduction.

Theorem 6.2.11 implies that it is impossible for **any** computer program to systematically find a Bell test to separate the quantum tensor product model from the quantum commuting operator model (as it is equivalently to solving the halting problem)! However, if we have prior knowledge about whether a Turing machine halts (for example, the Turing machine that arises from Pseudocode which contains an infinite loop), we could construct a bell experiment that realizes such a separation using Theorem 6.2.11, giving infinitely many bell experiments to test the separation between the tensor product model and the commuting operator model. Unfortunately, since these constructions rely on complexity techniques, this also implies that any experimental setup generated by using Theorem 6.2.11 would be impractical for experimental usage. For reference, the explicit separation proven by [JNV+22a] has an estimated question size and answer size of about 10^{20}. This is completely impractical experimentally, as each question requires a different measurement configuration, and each answer requires a precise measurement setting. We expect any separation generated by Theorem 6.2.11 to have similar, if not higher, question size and answer size as techniques used are similar to the ones used by [JNV+22a]. Thus, it would be an interesting open question whether we can show, using techniques from the operator algebra community, an example of a bell experiment which separates between the two models of entanglement with a more reasonable size.

CHAPTER 7

Conditionally linear distributions

In this section, we introduce conditionally linear distributions and formally define the compression introduced in the last section. Intuitively, these are precisely the distributions that the provers can sample locally by measuring a polynomial number of EPR pair using generalized Pauli measurements. As will be seen in Chapter 9, the low individual degree test used in the PCP construction due to Babai, Fortnow and Lund also follows the same input distribution [BFL91] which is one of the key reasons why the compression theorem can be proved. In order to define the conditionally linear distribution, we first give some necessary facts about both $\mathbb{F}_2$ and $\mathbb{F}_{2^q}$.

7.1 Preliminaries on finite fields

In this section, we recall some basic properties regarding finite fields of the form $\mathbb{F}_{2^p}$. In this thesis, we always assume that p is odd for finite fields of the form $\mathbb{F}_{2^p}$. $\mathbb{F}_{2^p}$ can always be viewed as a p-dimensional vector space over $\mathbb{F}_2$. Unless otherwise specified, we always assume that $\mathbb{F}_{2^p}$ has its basis defined over $\mathbb{F}_2$ (although the basis specified might be different). Given an element $a \in \mathbb{F}_{2^p}$ and a basis $\{\hat{e}_i\}_{i\in[p]}$ for $\mathbb{F}_{2^p}$, there exists a bijection map from a to $\mathbb{F}_2^p$ by

$$\kappa_{\{\hat{e}_i\}} : a \to (a_0, \cdots, a_{p-1}). \tag{7.1}$$

By using this map, any element of $\mathbb{F}_{2^p}$ can be represented as a bit string in $\{0,1\}^p$ as long as the set of bases is specified. Recall from [MP13, Definition 2.1.80], every finite field $\mathbb{F}_{2^p}$ admits a trace function over $\mathbb{F}_2$, or $\mathrm{Tr}(a) : \mathbb{F}_{2^p} \to \mathbb{F}_2$ defined as

$$\mathrm{Tr}(a) := \sum_{i=0}^{p} a^{2^i}.$$

The trace function has the properties that it is $\mathbb{F}_2$-linear, meaning $\mathrm{Tr}(a+b) = \mathrm{Tr}(a) + \mathrm{Tr}(b)$ and $\mathrm{Tr}(cb) = c \cdot \mathrm{Tr}(b)$ for $a, b \in \mathbb{F}_{2^p}$ and $c \in \mathbb{F}_2$. The field $\mathbb{F}_{2^p}$ is a linear space over $\mathbb{F}_2$ of dimension p. We refer to a set of bases $\{\hat{e}_i\}_{i\in[p]}$ for $\mathbb{F}_{2^p}$ over $\mathbb{F}_2$ to be *self-dual* if for all $i, j \in [p]$

$$\mathrm{Tr}(\hat{e}_i \hat{e}_j) = \delta_{i,j}.$$

Self-dual bases are known to exist for all finite fields of the form $\mathbb{F}_{2^q}$ [BGM+93, Theorem 1.9]. We refer to a set of bases $\{\hat{e}_i\}_{i\in[p]}$ as normal if there exists an element $a \in \mathbb{F}_{2^p}$ such that $\hat{e}_i = a^{2^i}$. The following lemma shows that, for p odd, there exists an efficient deterministic algorithm that computes a self-dual normal basis $\{\hat{e}_i\}_{i\in[p]}$ for $\mathbb{F}_{2^p}$ given p. The deterministic algorithm also outputs a description of an efficient algorithm for computing finite field multiplication when elements of $\mathbb{F}_{2^p}$ are represented under the bijection specified by (7.1) using the basis $\{\hat{e}_i\}_{i\in[p]}$.

Lemma 7.1.1 (Computability of finite fields, Lemma 3.16 of [JNV+22a])**.** *There exists a deterministic algorithm that, given an odd integer $p > 0$, outputs a self-dual normal basis of $\mathbb{F}_{2^p}$ over $\mathbb{F}_2$ and the multiplication tables for the basis in* poly(n) *time.*

In the lemma above, a multiplication table for the set of basis $\{\hat{e}_i\}_{i\in[p]}$ is the unique matrix representation $\{M_{\hat{e}_i}\}_{i\in[p]} \subseteq \mathcal{M}_k(\mathbb{F}_2)$ such that

$$M_{\hat{e}_i}\kappa_{\{\hat{e}_i\}}(a) = \kappa(\hat{e}_i a)$$

for all $i \in [p]$ and $a \in \mathbb{F}_{2^p}$. In this thesis, we refer to the set of self-dual basis generated by Lemma 7.1.1 as the *canonical basis* of $\mathbb{F}_{2^p}$. We use $\kappa(a)$ to denote the bijection given in (7.1) for the canonical basis. We represent elements $x \in \mathbb{F}_{2^p}$ as $\kappa(x) \in \{0,1\}^p$ (where we identify elements of $\mathbb{F}_2$ as $\{0,1\}$) in this thesis and we refer to this as the *canonical representation* for $\mathbb{F}_{2^p}$. In this thesis, we represent elements of any element $x \in \mathbb{F}_{2^p}$ as elements of $\mathbb{F}_{2^p}$ through the bijection map κ. Due to Lemma 7.1.1, for elements represented under the canonical representation, addition, multiplication, inversion and computing the trace can all be computed in poly(p) time (see [JNV+22a, Lemma 3.18] for more details).

Let $m \in \mathbb{N}$, any elements in $\mathbb{F}_{2^p}^m$ can be represented as elements of $\{0,1\}^{mp}$ through the canonical representation of $\mathbb{F}_{2^p}$, and we refer to this as the Canonical representation for $\mathbb{F}_{2^p}^m$. For clarity, we use $\{\hat{e}_i\}_{i\in[p\cdot m]}$ to denote the canonical basis for $\mathbb{F}_{2^p}^m$, and $\{e_i\}_{i\in[m]}$ to denote the element $(0_0,\cdots,1_i,\cdots,0_m)$ in $\mathbb{F}_{2^p}^m$ (where 1 is the identity element in $\mathbb{F}_{2^p}$). To abuse notation, we also use κ to denote the map from $\mathbb{F}_{2^p}^m \to \{0,1\}^{p\cdot m}$ where for $s = (s_0,\cdots,s_m) \in \mathbb{F}_{2^p}^m$

$$\kappa(s) := (\kappa(s_0),\cdots,\kappa(s_m)), \tag{7.2}$$

where κ is the bijection given in (7.1) for the canonical basis $\{\hat{e}_i\}_{i\in[p\cdot m]}$. When describing elements of $\mathbb{F}_{2^p}^m$, we always assume that the element is represented under the canonical representation. We refer to a subspace as a *canonical basis subspace* if it is the span of some subsets of the canonical basis. We remark that this is the same definition as the "register subspaces" in [JNV+22a]. We use $\dim(V)$ to denote the dimension of the subspace $V \subseteq \mathbb{F}_{2^p}^m$. For any subspace $W \subseteq V \subseteq \mathbb{F}_{2^p}^m$, we define the orthogonal subspace of W over V as the space

$$W^\perp := \{u \in V : u \cdot w = 0 \text{ for all } v \in W\}.$$

Unless otherwise stated, $W^\perp$ defaults to the orthogonal subspace over $\mathbb{F}_{2^p}^m$.

For subspaces $V_1, V_2 \subseteq V \subseteq \mathbb{F}_{2^p}^m$, we say that the two subspaces are disjoint if $V_1 \cap V_2 = \{0\}$. For any $k \le l$, we refer to a set of pairwise disjoint partition of subspace $\{V_j\}_{j\in[k]}$ of V as a *disjoint partition* of V if $\oplus_{j\in[k]} V_j = V$. For any disjoint partition of subspaces $\{V_i\}_{i\in[k]}$ of $V \subseteq \mathbb{F}_{2^p}^m$ and $0 \le i < k$ we write

$$V_{<i} := \bigoplus_{j\in[i]} V_j, \qquad V_{>i} := \bigoplus_{i<j<k} V_j$$

and we use the convention that $V_{<i+1} = V_{\le i}$, $V_{>i} = V_{\ge i+1}$ and $V_{<0} = \{0\}$. Given $\{V_j\}_{j\in[k]}$, a disjoint partition of V and $v \in V$, there exists a unique decomposition $s_j \in V_j$ for each $j \in [k]$ such that $s = \sum_{j\in[k]} s_j$.

Given subspace $U, W \subseteq V \subseteq \mathbb{F}_{2^p}^m$, we say (U,W) forms a pair of complementary subspaces over V if U and W form a disjoint partition of V and $U + W = V$, and we say (U,W) forms a pair of complementary subspace if it forms a complementary subspaces over $\mathbb{F}_{2^p}^m$. Give $v \in V$ and a pair of complementary subspace over V, there exists a unique decomposition $v = u + w$ such that $u \in U$ and $v \in W$. There could potentially be multiple subspace of V which can be used to form a pair complementary subspaces with W. For example, for $V = \mathbb{F}_2^2$ and $W = \text{span}(1,1)$, the subspace span$\{(1,0)\}$ and span$\{(0,1)\}$ both forms a pair of complementary subspace of V with W.

We wish to define a notion of a unique "canonical complement" in this thesis. For a canonical basis subspace V and $W \subseteq V$, we define the *canonical complement* as the following: Let $\{\hat{e}_0,\cdots,\hat{e}_{\dim V-1}\}$ be the canonical basis element used to define V, and let $\{w_1,\cdots,w_{\dim(W)}\}$ be a set of linearly independent vectors in W. Write each of the vector as $w_i = \sum_{j\in[\dim(V)]} a_{i,j}\hat{e}_j$ for some $a_{i,j} \in \mathbb{F}_{2^p}^m$ and run the Gaussian elimination on the $\dim(V)$ by $\dim(W)$ matrix defined by

$(a_{i,j})$. Let I be the set of $\dim(V)$ column with leading 1 entries in the resulting matrix, we define the *canonical complement* over V, or W^C to be the subspace $\text{span}\{\hat{e}_j | j \notin I\}$. Unless otherwise stated, W^C defaults to the canonical complement of $\mathbb{F}_{2^p}^m$. The canonical complement is unique and can be computed efficiently in poly time. We remark that this is the same definition for canonical complement given in [JNV+22a, Definition 3.6].

We recall the following lemma regarding the subspaces of $\mathbb{F}_{2^p}^m$.

Lemma 7.1.2 (Lemma 3.14 of [JNV+22a])**.** *Let $\{\hat{e}_i\}$ be the canonical basis for $\mathbb{F}_{2^p}$ over $\mathbb{F}_2$ and let V be a subspace of $\mathbb{F}_{2^p}^m$ with linear independent basis $\{b_1 \cdots, b_t\} \subseteq \mathbb{F}_{q^k}^n$. Then the following holds:*

- *$\kappa(V)$ is a subspace of $\mathbb{F}_2^{mp}$.*
- *$\{\kappa(\hat{e}_i b_j)\}_{(i,j)\in[p]\times[n]}$ forms a set of linearly independent basis of $\kappa(V)$ over $\mathbb{F}_2$.*
- *Let V, W be complementary subspaces of of $\mathbb{F}_{2^p}^m$. Then $\kappa(V)$ and $\kappa(W)$ are complementary subspaces of $\mathbb{F}_2^{pm}$. Furthermore, for all $a \in \mathbb{F}_{2^p}^m$ with $a = a^V + a^W$, $a^V \in V$ and $a^W \in W$, we have $\kappa(a^V) \in \kappa(V)$ and $\kappa(a^W) \in \kappa(W)$.*

Given $(v_0, \cdots, v_{n-1}) \in \mathbb{F}_{2^p}^m$ and $j \in [n]$, we use $\pi_{>j}^m$ to denote the function which zeros out the first j entries of $\mathbb{F}_{2^p}^m$ and we use $\pi_{\leq j}^m$ to denote the function which zeros out everything except for the first j entries of $\mathbb{F}_{2^p}^m$. In other words

$$\begin{aligned} \pi_{>j}^m(v_0, \cdots, v_{m-1}) &= (0, \cdots, 0, v_j, \cdots, v_{m-1}) \\ \pi_{\leq j}^m(v_0, \cdots, v_{m-1}) &= (v_0, \cdots, v_{j-1}, 0, \cdots, 0). \end{aligned} \tag{7.3}$$

We remark that this is a different map define in (3.1), as the entry specified here are over the finite field elements $\mathbb{F}_{2^p}$ instead of the $\{0,1\}$ string. We further remark that $\kappa^{-1} \circ \pi_{\leq j} \circ \kappa$ and $\mathbb{F}_{2^p}^m$ are different maps, where the first map are usually used for treating an element from $\mathbb{F}_{2^p}^m$ as a string and the second one are usually used for treating an element from $\mathbb{F}_{2^p}^m$ as a vector space over $\mathbb{F}_{2^p}$.

In this thesis, we work with linear functions over canonical basis subspace. For a linear function $\mathtt{L}$ mapping from $V \to V$, we use $\ker(\mathtt{L})$ to denote the subspace of V such that $\mathtt{L}(a) = 0$ for all $a \in \ker(\mathtt{L})$.

For a linear function $\mathtt{L} : V \to V$ over a canonical basis subspace $V \subseteq \mathbb{F}_{2^p}^m$, we define the linear function $\mathtt{L}^\perp : V \to V$ as the projection into the subspace of $\left(\ker(\mathtt{L})^\perp\right)^C$. To be more precise, for $v = v_1 + v_2 \in V$ with $v_1 \in \ker(\mathtt{L})^\perp$ and $v_2 \in \left(\ker(\mathtt{L})^\perp\right)^C$, $\mathtt{L}(v) = v_2$. We remark that this is the same as the linear map defined in [JNV+22a, Definition 3.11]. By a simple calculation, we see that

$$\ker(L)^\perp = \ker\left(\mathtt{L}^\perp\right).$$

7.1.1 Affine lines and polynomials over a finite field

In this subsection, we recall some properties related to affine lines and low-degree polynomials over a finite field. In this thesis, we refer to an affine line $\mathbf{l}$ over $\mathbb{F}_{2^p}^m$ as the set of the form

$$\{u + t \cdot v : t \in \mathbb{F}_{2^p}\}$$

for $u, v \in \mathbb{F}_{2^p}^m$. Given a line $\mathbf{l}$, we wish to define a unique $u_\mathbf{l}, v_\mathbf{l} \in \mathbb{F}_{2^p}^m$ which can be used to represent $\mathbf{l}$. Given $u \in V$, we define the linear function $\text{Null}_v^{\text{LN}} : \mathbb{F}_{2^p}^m \to \mathbb{F}_{2^p}^m$ as

$$\text{Null}_v^{\text{LN}}(u) := u_v,$$

where $u = u_v + u_v^C$ is the unique decomposition such that $u_v \in \text{span}(v)$ and $u_{v^C} \in \text{span}(v)^C$. We remark that $\text{Null}_v^{\text{LN}}$ is the same as [JNV+22a, Definition 7.3]. We define the canonical representation of an affine line as the following:

Definition 7.1.3 (Canonical representation of an affine line)**.** *Let $p \in \mathbb{N}$ be an odd integer and $m \in \mathbb{N}$, and let $l = \{u + tv : t \in \mathbb{F}_{2^p}\}$ be an affine line passing through $\mathbb{F}_{2^p}^m$. The canonical representation of l is defined as*

$$Can(l) := (v, Null_v^{LN}(u)) \in \mathbb{F}_{2^p}^{2m}.$$

In the definition above, we see that $\text{Null}_v^{\text{LN}}(u) = \text{Null}_v^{\text{LN}}(u')$ for $u, u' \in l$ as $u' = u + t \cdot v$ any scalar $t \in \mathbb{F}_{2^p}$. Hence, the canonical representation is independent of the initial point chosen. We remark that this is the same definition given in [JNV+22a, Definition 7.3].

Given a function $\mathbf{f} : \mathbb{F}_{2^p}^m \to \mathbb{F}_{2^p}$, we say $\mathbf{f}$ is an *$m-$variant polynomial* over $\mathbb{F}_{2^p}$ if $\mathbf{f}$ is of the form

$$\mathbf{f}(x_1, \cdots, x_m) = \sum_{(i_1,\cdots,i_m)\in[2^p]^m} \alpha_{i_1,\cdots,i_m} x_1^{i_1} \cdots x_m^{i_m},$$

where each of the $\alpha_{i_1,\cdots,i_m}$ are some coefficients in $\mathbb{F}_{2^p}$. Furthermore, we say that $\mathbf{f}$ has an *individual degree* of $d \in \mathbb{N}$ if the sum above is defined over $[d]^m$ instead (in other words, $\alpha_{i_1,\cdots,i_m} = 0$ if there exists a $j \in [m]$ such that $i_j > d$). We use $\text{IdPoly}(p, m, d)$ to denote the set of polynomials $\mathbf{g} : \mathbb{F}_{2^p}^m \to \mathbb{F}_{2^p}$ with individual degree of at most d. We recall the following lemma regarding the distance between two distinct low-individual degree polynomials.

Lemma 7.1.4 (Schwartz-Zippel [Sch80; Zip79])**.** *Let $\boldsymbol{f}, \boldsymbol{g} \in IdPoly(p, m, d)$ be two different m-variant polynomials with individual degree of at most c, then*

$$\Pr_{u\sim\mathbb{F}_{2^p}^m}[\boldsymbol{f}(u) = \boldsymbol{g}(u)] \leq \frac{md}{2^p}.$$

For a more comprehensive introduction for finite fields, we refer the readers to [MP13].

7.2 Conditionally linear distributions

In order to introduce conditionally linear distributions, we need to first introduce the notion of conditionally linear functions.

Definition 7.2.1 (Conditionally linear functions)**.** *Let $p \in \mathbb{N}$ be an odd integer and $m, k \in \mathbb{N}$, and let V be a canonical basis subspace of $\mathbb{F}_{2^p}^m$. We say that the function $\mathtt{L} : V \to V$ is a k-th level conditionally linear function over V if $\mathtt{L}$ can be defined using the following construction:*

- *There exists a disjoint partition of subspaces $\{V_h\}_{h\in[k]}$ of V, where each V_h is a canonical basis subspace, which we refer to as "registers" for the function $\mathtt{L}$.*
- *There exists a single linear function $\mathtt{L}_{0,0} : V_0 \to V_0$, this is referred to as the zeroth-level linear function of $\mathtt{L}$.*
- *For $0 < j < k$, the function $\mathtt{L}_j : V_{\leq j} \to V_{\leq j}$ is defined recursively as follows:*
  - *There exists a collection of linear functions $\{\mathtt{L}_{j,s} : V_j \to V_j\}_{s\in V_{<j}}$, which is referred to as the j-th level linear functions.*
  - *For every input $s \in V_{\leq j}$, write s as $s = s_j + s_{<j}$ for $s_j \in V_j$ and $s_{<j} \in V_{<j}$, then*

$$\mathtt{L}_j := \mathtt{L}_{j-1}(s_{<j}) + \mathtt{L}_{j,\mathtt{L}_{j-1}(s_{<j})}(s_j)$$

- *Finally, we have $\mathtt{L} = \mathtt{L}_{k-1}$.*

We refer to Figure 7.1 for more intuition about conditionally linear functions. In this thesis, we abbreviate conditionally linear as CL. In this thesis, we also define CL functions $\mathtt{L}$ with p being an even integer. When this occurs, we assume that the range of $\mathtt{L}$, V , contains an additional canonical basis (i.e. $\mathtt{L} : \mathbb{F}_{2^{p+1}}^m \to \mathbb{F}_{2^{p+1}}^m$), and the additional canonical basis created are merged into the register V_0 and lies in the kernel space of $L_{0,0}$. We give a simple example of a CL function below.

Example 7.2.2. Consider the finite field $\mathbb{F}_2^3$ with basis (e_0, e_1, e_2), and the function

$$\mathtt{L}(x_0, x_1, x_2) = (x_0, x_0 \cdot x_1 + (1 + x_0) \cdot x_2, 0) = (x_0, x_0 \cdot x_1 + x_1 \cdot x_2 + x_2, 0).$$

The function $\mathtt{L}$ is a second level CL function with registers $V_0 = \mathrm{span}((1,0,0))$ and $V_1 = \mathrm{span}((0,1,0), (0,0,1))$. $\mathtt{L}$ can be defined as the following: the zeroth level linear function for $\mathtt{L}$ can be taken as

$$\mathtt{L}_{0,0}(x_0, 0, 0) = (x_0, 0, 0),$$

and two first level linear functions for $\mathtt{L}$ can be specified by

$$\mathtt{L}_{1,0}(0, x_1, x_2) = (0, x_1, 0), \quad \text{and} \quad \mathtt{L}_{1,1}(0, x_1, x_2) = (0, x_1, 0).$$

Since $\{0\} \subseteq V \subseteq \mathbb{F}_{2^p}^m$, any k-th level CL function is also a k'-th level CL function for $k' \geq k$. Since the definition for a CL function is recursive, for each $j \in [k]$ and $s \in V_j$, one can define a $k-j$ level CL function from $V_{\geq j} \to L_{\geq j}$ by setting the initial linear function to be $\mathtt{L}_{j,s}$. Similarly, we can combine a k_1-th level CL function and a collection of k_2-th level CL functions with the same register into a $(k_1 + k_2)$-level CL function through the series composition. We make this transformation more precise below. We remark that this is the composition defined in [JNV+22a, Lemma 4.7].

Definition 7.2.3 (Series composition of CL functions)**.** *Let $V \subseteq \mathbb{F}_{2^p}^m$ be a canonical basis subspace and let $V = V^1 \oplus V^2$ be a disjoint partition where both V^1 and V^2 are canonical basis subspaces. Let $\mathtt{M} : V^1 \to V^1$ be a k_1-th level CL function with registers $\{V_j^1\}_{j\in[k_1]}$, and let $\{\mathtt{N}^x : V^2 \to V^2\}_{x\in V^1}$ be a collection of k_2-th level CL functions which share the same registers $\{V_j^2\}_{j\in[k_2]}$. Define the series composition between $\mathtt{M}$ and $\{\mathtt{N}^x\}_{x\in V^1}$ to be a $(k_1 + k_2)$-level CL function $\mathtt{L} : V \to V$, defined as follows:*

- *The function has registers $\{V_0, \cdots V_{n+m}\}$, where $V_j = V_j^1$ for $0 \leq j < k_1$ and $V_j = V_{j-k_1}^2$ for $k_1 \leq j < k_1 + k_2$.*
- *For $0 \leq j < k_1$ and for each $h \in V_{<j}$, we have*

$$\mathtt{L}_{j,h} = \mathtt{M}_{j,h}$$

- *For $k_1 \leq j < k_1 + k_2$, we define*

$$\mathtt{L}_{j,h_v+h_w} = \mathtt{N}^{h_v}_{j-m,h_w},$$

where $h_v \in V$ and $h_w \in W$.

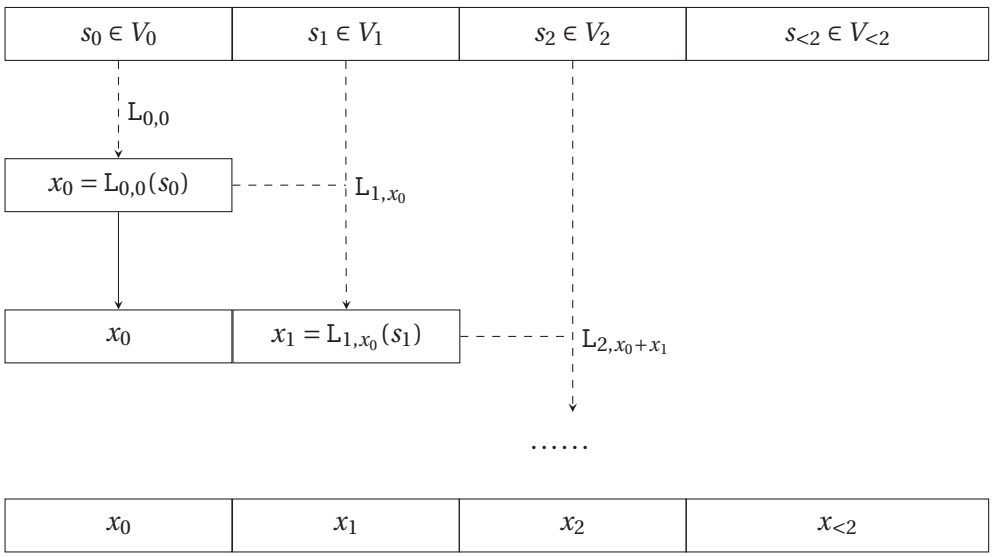

Figure 7.1: A diagram representation of the conditionally linear function $\mathtt{L}$.

In other words, for every input $s \in V$, write $s = v_1 + v_2$, where $v_1 \in V^1$ and $v_2 \in V^2$. $\mathtt{L}$ first applies the k_1-th level CL function $\mathtt{M}$ to v_1, the V^1 component for V. Based on the output of $\mathtt{M}(v_1)$, $\mathtt{L}$ will then apply the k_2-th level CL function $\mathtt{N}^{\mathtt{M}(v_1)}$ to v_2, the V^2 component within V. We can also combine two CL functions in parallel as described below, we remark that this is the CL function constructed in [JNV+22a, Lemma 4.9].

Definition 7.2.4 (Parallel composition of CL functions)**.** *Let $V \subseteq \mathbb{F}_{2^p}^m$ be a canonical basis subspace and let $V = V^1 \oplus V^2$ be a disjoint partition where both V^1 and V^2 are canonical basis subspaces. Let $\mathtt{M} : V^1 \to V^1$ and $\mathtt{N} : V^2 \to V^2$ be a k-th level CL function with registers $\{V_j^1\}_{j\in[k]} \subseteq V^1$ and $\{V_j^2\}_{j\in[k]} \subseteq V^2$ respectively. Define the parallel composition between $\mathtt{M}$ and $\mathtt{N}$ to be a level k CL function $\mathtt{L} : V \to V$, defined as follows:*

- *The function has registers $\{V_j = V_j^1 \oplus V_j^2\}$.*
- *For every $j \in [k]$ and $h_{<j} \in V_{<j}$, write $h_{<j} = h_{<j}^1 + h_{<j}^2$ for $h_{<j}^1 \in V_{<j}^1$ and $h_{<j}^2 \in V_{<j}^2$ (resp. $h_j^2 \in V_j^2$). Define $\mathtt{L}_{j,h_{<j}} : V_j \to V_j$ to be*

$$\mathtt{L}_{j,h_{<j}} = \mathtt{M}_{j,h_{<j}^1} + \mathtt{N}_{j,h_{<j}^2}$$

We introduce the notion of a CL distribution below. In this thesis, we consider mainly games with a CL distribution as the question distribution.

Definition 7.2.5 (Conditionally linear distributions)**.** *A distribution μ is a (k, m, p) CL distribution if there exist two k-th level CL function $\mathtt{L}^P : \mathbb{F}_{2^p}^m \to \mathbb{F}_{2^p}^m$, $P \in \{A, B\}$ with the same register $\{V_j\}_{j\in[k]}$ such that μ can be sampled in the following way.*

1. *Uniformly sample $s \in \mathbb{F}_{2^p}^m$.*
2. *Give $\mathtt{L}^A(s)$ to Alice, and $\mathtt{L}^B(s)$ to Bob.*

Given a sample $(x, y) \sim \mu$, we refer to the $s \in \mathbb{F}_{2^p}^m$ as the "seed" for the given sample if $(x, y) = (\mathtt{L}^A(s), \mathtt{L}^B(s))$.

We refer to a non-local game $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$ as CL samplable if the sampling procedure is a CL distribution, and as a k-th level CL samplable game if furthermore the question distribution is a k-th level CL distribution.

For any game $\mathcal{G}$, if μ is a (k, m, p) CL distribution, then we can write $\mathcal{X} = \mathbb{F}_{2^p}^m = \{0,1\}^{p\cdot k}$ by mapping elements of $\mathbb{F}_{2^p}^m$ into its canonical representation. We remark that in [JNV+22a] each of the $\mathtt{L}$ does not necessarily need to share the same register. However, we choose to present it this way for simplicity of notation, and the introspection procedure in Chapter 9 would still work even if $\mathtt{L}^A$ and $\mathtt{L}^B$ are defined using different registers. Finally, we have the following lemma stating that any r-fold parallel repetition of a CL samplable game is still CL samplable, and the proof follows trivially from applying the parallel composition given in Definition 7.2.4.

Lemma 7.2.6. *Let $r \in \mathbb{N}$, and let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$ be a CL samplable game via a (k, m, p)-CL distribution. Then, $\mathcal{G}^{\otimes r} = (\mathcal{X}^r, \mathcal{A}^r, \mu^{\otimes r}, D^{\otimes r})$ is samplable via a $(k, r \cdot m, p)$ CL distribution.*

7.2.1 Typed conditionally linear distributions

In this subsection, we introduce a more general notion of CL distributions, which we call the *typed* CL distribution in this thesis. We also show that any game with a typed CL distribution as its sampling distribution can be modified into a game which is CL samplable with only a polynomial decay on the success rate. We remark that the typed CL distribution defined in this subsection is essentially the same as [JNV+22a, Section 4, Section 6], but with some minor tweaks to accommodate synchronicity condition in the context of non-local games.

Let $(\mathtt{T},\mathtt{E})$ be an undirected graph with at least one edge in $\mathtt{E}$. We represent the elements of $\mathtt{T}$ as elements of $[|\mathtt{T}|]$ and we represent the edges of the graph as $(v^0,v^1)\in\mathtt{T}^2$ with $v^0=v^1$ representing a self-edge in the graph. We introduce the notion of a *Typed CL distribution below.*

Definition 7.2.7 (Typed conditionally linear distribution)**.** *Let* $(\mathtt{T},\mathtt{E})$ *be a typed graph with* $\mathtt{T}=\{0,1\}^{l^t}$ *and let* $\{\mathtt{L}^v : V\to V\}_{v\in\mathtt{T}}$ *be a collection of k-th level CL functions of size p over a canonical basis subspace* $V\subseteq\mathbb{F}_{2^p}^m$*, where all of the* $\mathtt{L}^v$ *share the same registers* $\{V_j\}_{j\in[k]}$*. A distribution* μ *is a* $(\mathtt{T},\mathtt{E},\{\mathtt{L}^v\})$*-typed distribution if* μ *can be sampled in the following manner:*

1. *Uniformly sample* $(v_0,v_1)\in\mathtt{T}^2$ *and perform rejection sampling until* $(v_0,v_1)\in\mathtt{E}$*.*
2. *Uniformly sample* $s\in\mathbb{F}_{2^p}$ *and* $b\in\{0,1\}$*.*
3. *Compute* $x_0=\mathtt{L}^{v_0}(s)$ *and* $x_1=\mathtt{L}^{v_1}(s)$
4. *Return the pair* $((v_b,x_b),(v_{1-b},x_{1-b}))$*, as the sample outcome.*

We remark in the above sampling procedure, the self-edges are sampled with twice the weight in comparison to other edges. Let $\mathscr{G}=(\mu,V,\mathscr{X},\mathscr{A})$ be a synchronous game such that μ is a $(\mathtt{T},\mathtt{E},\{\mathtt{L}^v\})$-typed distribution. We refer to the set $\mathtt{T}$ as the "question label" and the set $\mathtt{E}$ as the "question pair" for the game $\mathscr{G}$. Since the question label is included as an output in the sampling procedure, the synchronous question pair for $\mathscr{G}$ corresponds to the self-edges in the graph $(\mathtt{T},\mathtt{E})$. For the remainder of thesis, when discussing games with typed CL distribution as the sampling distribution, we also assume that $\frac{|\mathtt{T}|}{p}$ is always an integer in this thesis by implicitly padding $\mathtt{T}$ with extra vertices which only contains self-loops. Since we usually associate $|\mathtt{T}|=O(p)$ in this thesis, this assumption does not change the complexity of $\mathtt{T}$.

Intuitively, typed CL distributions are a generalization of CL distributions, where instead of having two CL functions, there could potentially be $|\mathtt{T}|$ different CL functions used for sampling. We give a method of taking a synchronous game with typed CL distribution and simulating it with a synchronous game using a normal CL distribution. Given a graph $(\mathtt{T},\mathtt{E})$ and a vertex $v\in\mathtt{T}$, we define the *neighbour indicator* for v to be the vector $\mathrm{neigh}_{\mathtt{E}}(v)\in\{0,1\}^{|\mathtt{T}|}$ such that

$$\mathrm{neigh}_{\mathtt{E}}(v_0)_{v_1}:=\begin{cases}1 & \text{if } \{v_0,v_1\}\in\mathtt{E}\\ 0 & \text{otherwise}\end{cases}.$$

The following transformation shows that games with a typed-CL distribution as the sampling procedure can always be simulated with a (normal) CL samplable game. We remark that this construction is similar to the one given in [JNV+22a, Section 6.2].

Definition 7.2.8 (Detyped conditionally linear distribution)**.** *Let* $p,m\in\mathbb{N}$ *with* p *an odd positive integer. Let* $(\mathtt{T},\mathtt{E})$ *be a graph,* $\{\mathtt{L}^v : V\to V\}_{v\in\mathtt{T}}$ *be a collection of k-th CL function with registers* $\{V_j\}_{j\in[k]}$ *for some canonical basis subspace* $V\subseteq\mathbb{F}_{2^p}^m$ *and let* μ *be a* $(\mathtt{T},\mathtt{E},\{\mathtt{L}^v\}_{v\in\mathtt{T}})$*-typed distribution. Let* $b=\frac{|\mathtt{T}|}{p}$ *and* $l^t=\left\lceil\frac{\log_2(|\mathtt{T}|)}{p}\right\rceil$*.*

We define the detyped transformation for μ*, denoted as* μ^D*, to be a* $(k+2,m+4\cdot b+2\cdot l^t,p)$ *CL distribution. Where the second level CL function* $\mathtt{L}^A,\mathtt{L}^B:\mathbb{F}_{2^p}^{4\cdot b+2\cdot l^t}\oplus V\subseteq\mathbb{F}_{2^p}^{4\cdot b+2\cdot l^t+m}$*.*

- $\mathtt{L}^A$ *is defined as a series composition between* $\mathtt{L}^{A,1}$ *and* $\{\mathtt{L}^{A,2,s}\}_{s\in S_1}$ *as per Definition 7.2.3. Where* $\mathtt{L}^{A,1}$ *is a second level CL samplable function acting on* $V^1=\mathbb{F}_{2^p}^{2\cdot l^t+4\cdot b}$*, and* $\{\mathtt{L}^{A,2,s}\}_{s\in S_1}$ *is a collection of k-th level CL functions acting on* $V^2=V$*. We represent elements of* V^1 *under the canonical representation (i.e. as elements of* $\{0,1\}^{2\cdot p(l^t+2\cdot b)}$*) in this definition formulation. We give the details for each level of* $\mathtt{L}^A$ *below:*
  - *The first second level CL function* $\mathtt{L}^{A,1}$ *acts on two registers,* $V_0^1=\{0,1\}^{2\cdot p\cdot(l^t+b)}=\{0,1\}^{2b\cdot p+4\cdot|\mathtt{T}|}$ *and* $V_1^1=\{0,1\}^{2\cdot p\cdot b}=\{0,1\}^{2\cdot|\mathtt{T}|}$*.*

– *We define the zeroth level linear function for* $\mathtt{L}^{A,1}_{0.0}$ *as follows: For all* $x \in V^1_0$ *as* (x_0, x_1, x_2, x_3)*, where* $x_0, x_2 \in \{0,1\}^{p \cdot l^t}$ *and* $x_1, x_3 \in \{0,1\}^{|\mathtt{T}|}$*, then* $\mathtt{L}^{A,1}_{0,0}$

$$\mathtt{L}^{A,1}_{0,0}(x_0, x_1, x_2, x_3) = (x_0, x_1, 0, 0) \tag{7.4}$$

for all elements $x \in V^1$*.*

– *Fix* $s_0 \in \{0,1\}^{p \cdot l^t}$ *and* $s_1 \in \{0,1\}^{|\mathtt{T}|}$*, and let* $v = \textbf{\textit{bininv}}((\pi_{\leq \lceil \log_2(|\mathtt{T}|) \rceil}(s_0))$ *(i.e. extract the first* $\lceil \log_2(|\mathtt{T}|) \rceil$ *bits of* s_0 *and treat it as an integer). For all* $x \in V^1_1$*, parse* x *as* (x_4, x_5)*, where* $x_4, x_5 \in \{0,1\}^{|\mathtt{T}|}$*. Whenever* $s_0 \in [|\mathtt{T}|] = \mathtt{T}$ *and* $s_1 = \mathit{neigh}_{\mathtt{E}}(v)$*, we define the first level linear function for* $\mathtt{L}^{A,1}$ *as*

$$\mathtt{L}^{A,1}_{1,(s_0,s_1,s_2,s_3)}(x_4, x_5) = (x_4, (x_5)_v), \tag{7.5}$$

where $(x_5)_v$ *zeros out all entries of* x_5 *except for the* v*th entry. Otherwise*

$$\mathtt{L}^{A,1}_{1,(s_0,s_1,s_2,s_3)}(x_4, x_5) = 0.$$

– *The collection of* k*-th level CL functions* $\{\mathtt{L}^{A,2,s}\}_{s \in S_1}$ *is defined to be the following: Parse* $s \in V^1$ *as* $(s_0, s_1, s_2, s_3, s_4, s_5)$*, and define* v *the same way as the above clause. We define*

$$\mathtt{L}^{A,2,s}(x) = \mathtt{L}^v \tag{7.6}$$

whenever $v \in \mathtt{T}$*,* $s_1 = s_4 = \mathit{neigh}_{\mathtt{E}}(v)$ *and* $(s_5)_v = 1$*. Otherwise we set* $\mathtt{L}^{A,2,s}(x) = 0$

• $\mathtt{L}^B$ *is defined mostly similar to* $\mathtt{L}^A$*, except we replace the equation* (7.4)*,* (7.5) *and* (7.6) *by*

$$\mathtt{L}^{B,1}_{0,0}(x_0, x_1, x_2, x_3) = (x_2, x_3, 0, 0),$$
$$\mathtt{L}^{B,1}_{1,(s_0,s_1,s_2,s_3)}(x_4, x_5) = (x_5, (x_4)_v),$$
$$\mathtt{L}^{B,2,s}(x) = \mathtt{L}^v$$

We also give the notion of a "non-trivial seed" for a detyped CL distribution below. Using the same notation as Definition 7.2.8, for $s = s^1 \oplus s^2 \in V^1 \oplus V^2$. Parse $s^1 = (s_0, s_1, s_2, s_3, s_4, s_5)$, where $s_0, s_2 \in \{0,1\}^{p \cdot l^t}$ and $s_1, s_3, s_4, s_5 \in \{0,1\}^{|\mathtt{T}|}$. Furthermore, parse s_0 into an element v_0 as per described in the definition of $\mathtt{L}^{A,1}$, and parse s_2 into v_1 in the same way. We call s a non-trivial seed for the CL distribution μ^{D} if the following holds

1. $v_0, v_1 \in \mathtt{T}$ and $(v_0, v_1) \in \mathtt{E}$
2. $s_1 = s_4 = \mathrm{neigh}_{\mathtt{E}}(v_0)$ and $s_3 = s_5 = \mathrm{neigh}_{\mathtt{E}}(v_1)$

otherwise, we refer to s as a trivial seed.

The detyping procedure might seem convoluted at first. Intuitively, (s_0, s_1, s_4) in s^1 as given in the above definition dictates which vertices $\mathtt{L}^{A,1}$ samples and (s_2, s_3, s_5) dictates which vertices $\mathtt{L}^{B,1}$ samples. If we perform a rejection sampling on the set of non-trivial seeds, we see that this is equivalent to the original typed CL distribution since both $s_1 = s_4 = \mathrm{neigh}_{\mathtt{E}}(v_0)$ and $s_3 = s_5 = \mathrm{neigh}_{\mathtt{E}}(v_1)$ occurs with the same probability given a fixed $(v_0, v_1) \in \mathtt{E}$. For a detyped sampler μ^{D}, in the event that a non-trivial seed is being chosen, the resulting output (x, y) has the following form:

$$x = (v, \mathrm{neigh}_{\mathtt{E}}(v), 0, 0, \mathrm{neigh}_{\mathtt{E}}(v), (\mathrm{neigh}_{\mathtt{E}}(u))_v, \mathtt{L}^v(s^2))$$
$$y = (u, \mathrm{neigh}_{\mathtt{E}}(u), 0, 0, \mathrm{neigh}_{\mathtt{E}}(u), (\mathrm{neigh}_{\mathtt{E}}(v))_u, \mathtt{L}^u(s^2)) \tag{7.7}$$

for some $u, v \in \mathtt{T}^2$. Since the seed is non-trivial, $(u, v) \in \mathtt{E}$ and hence $(\mathrm{neigh}_{\mathtt{E}}(u))_v$ and $(\mathrm{neigh}_{\mathtt{E}}(v))_u$ will always be 1. By Definition 7.2.8, the output from μ^{D} can only be parsed in the above form iff s is initially chosen to be a non-trivial seed. We have the following lemma lower bounding the set of non-trivial seeds in a detyped transformation.

Lemma 7.2.9 (Percentage of non-trivial seeds in a detyped CL distribution)**.** *Let μ be a $(\mathtt{T},\mathtt{E},\{\mathtt{L}^v\}_{v\in\mathtt{T}})$-typed distribution with $|\mathtt{T}| = t$ and $\{\mathtt{L}^v : V \to V\}_{v\in\mathtt{T}}$ for some subspace $V = \mathbb{F}_{2^p}^m$ and let μ^D be the corresponding $(k+2, m+4\cdot b+2\cdot l^t, p)$ detyped CL distribution as defined in Definition 7.2.8. For every s sampled uniformly random from $\mathbb{F}_{2^p}^{4\cdot b+2\cdot l^t} \oplus V$, s is a non-trivial seed for μ^D with probability at least*

$$\frac{1}{4t^2\cdot 16^t}.$$

Proof. We use the same notation as the one given in Definition 7.2.8. We first point out that for $s = s^1 \oplus s^2 \in V^1 \oplus V^2$, only s^1 dictates whether s is a non-trivial seed. Hence consider s^1 uniformly sampled from V^1 and write $s^1 = (s_0, s_1, s_2, s_3, s_4, s_5)$ into the parsing as defined as in Definition 7.2.8.

We first lower bound the probability that s^1 satisfies the first clause of being a non-trivial seed. If we take the first $\lceil \log_2(t) \rceil$ bit $s_0 \in \{0,1\}^{l^t\cdot p}$ and convert it back to integer through the map $\mathbf{bininv}(\cdot)$, then with probability at least $\frac{1}{2}$ we obtain some $v_0 \in \mathtt{T}$. Similarly, with probability $\frac{1}{2}$, we can parse s_2 into some $v_1 \in \mathtt{T}$. Since we assume there is at least one edge in $\mathtt{E}$, the probability that $(v_0, v_1) \in \mathtt{E}$ is at least $\frac{1}{n^2}$ given that $(v_0, v_1) \in [|t|] \times [|t|]$. Hence s^1 satisfies the first clause with probability at least $\frac{1}{4t^2}$.

For the second clause, for any given $v \in \mathtt{T}$, since there is only one unique string in $\{0,1\}^{|\mathtt{T}|}$ which is equal to $\mathrm{neigh}_{\mathtt{E}}(v)$. Given $v_0 \in \mathtt{T}$, s_1 and s_3 will have probability $\left(\frac{1}{2^t}\right)^2$ to be equal to $\mathrm{neigh}_{\mathtt{E}}(v_0)$. Hence s^1 satisfies the second clause with probability $\left(\frac{1}{2^t}\right)^4 = \frac{1}{16^t}$ given the first clause. Hence s^1 has a probability $\frac{1}{4t^2}\cdot\frac{1}{16^t}$ of being a non-trivial seed. □

Given a synchronous game with a typed CL distribution as the input distribution, we define the following transformation which replaces the typed CL distribution with its detyped counterpart below.

Definition 7.2.10 (Detyped conditionally linear game)**.** *Let $p, m \in \mathbb{N}$ with p being an odd positive integer. Let $(\mathtt{T},\mathtt{E})$ with $\mathtt{T} = \{0,1\}^{l^t}$ and let $\{\mathtt{L}^v : V \to V\}_{v\in\mathtt{T}}$ be a k-th level CL function for some canonical basis subspace $V \subseteq \mathbb{F}_{2^p}^m$. Let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, V)$ be a synchronous game with μ being a $(\mathtt{T},\mathtt{E},\{\mathtt{L}^v\}_{v\in\mathtt{T}})$-typed distribution as defined in Definition 7.2.7. We define the detyped and synchronization transformation $\mathcal{G}^D = (\mathcal{X}^D, \mathcal{A}, \mu^D, V^D)$ for $\mathcal{G}$ as follows:*

- *The distribution μ^D is the $(k+2, m+4\cdot b+2\cdot l^t, p)$ CL distribution given by Definition 7.2.8.*
- *For the question pair $(x,y) \in \mathbb{F}_{2^p}^{m+4\cdot b+2\cdot l^t}$, the verification V^D is given as follows:*
  - *If s is a non-trivial seed, parse (x,y) as per Equation (7.7), then*

$$V^D(x,y,a,b) = V\left((v^0, L^{v^0}(s)), (v^1, L^{v^1}(s)), a, b\right)$$

    in other words, the same as the original game.
  - *Otherwise, $V^D(x,y,a,b) = \delta_{a,b}$.*

We see that the above transformation also preserves synchronicity for a given game. One might wonder the reason for such a roundabout way to defining the detyping procedure, since instead, one can sample two arbitrary vertices from $\mathtt{T}$ and perform rejection sampling on the case where these two vertices are connected in $\mathtt{E}$. As seen in the lemma below, the main reason for the roundabout way is that it allows the transformation to also preserves any perfect oracularizable strategies after the transformation.

To be more precise, let $\mathcal{G}$ be a non-local game which uses some typed CL sampler μ. If we were to replace μ with the "sampling two vertices, and perform rejection sampling" approach for $\mathcal{G}$, any oracularizable strategy might no longer be oracularizable because the measurement operator used between "two non-connected vertices" might not commute. Using the detyping procedure listed above, at least one of the provers can deduce whether s is the trivial seed,

and thus, can adjust their measurement operator accordingly. To this end, we have the following lemma which shows how much the completeness/soundness condition changes for a detyped and synchronization transformation. We remark that the proof of this lemma is similar to [JNV+22a, Lemma 6.18].

Lemma 7.2.11 (Preservation of completeness/soundness of the detyped and synchronization game)**.** *Let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, V)$ be a synchronous game with μ being a $(\mathtt{T}, \mathtt{E}, \{\mathtt{L}^v\}_{v\in\mathtt{T}})$-typed distribution. For model $t \in \{*, co\}$, then*

- ***(Completeness)** If there exists a perfect oracularizable synchronous strategies for $\mathcal{G}$ in model t, then there exist a perfect oracularizable synchronous strategies for $\mathcal{G}^D$ in model t.*
- ***(Soundness)** If $\omega^t(\mathcal{G}) \geq 1-\varepsilon$, then*

$$\omega^t(\mathcal{G}^D) \geq 1 - \frac{\varepsilon}{\left(4|\mathtt{T}|^2 \cdot 16^{|\mathtt{T}|}\right)}.$$

Proof. Fix $t \in \{*, co\}$, and assume that the quantum strategy performed below is defined using model t. In the proof below, for $v \in \mathtt{T}$, we define

$$\text{view}(v) = (\mathbf{bin}(v), \text{neigh}_{\mathtt{E}}(v), 0, 0, \text{neigh}_{\mathtt{E}}(v), e_v), \tag{7.8}$$

where $\mathbf{bin}(v)$ above is only taken over the first $\lceil \log_2(|\mathtt{T}|) \rceil$ bits. By definition, for $(x, y) \sim \mu^{\mathrm{D}}$, $(x, y) = \left((\text{view}(v_0), \mathtt{L}^{v_0}(s^2), (\text{view}(v_1), \mathtt{L}^{v_1}(s^2))\right)$ for some $(v_0, v_1) \in \mathtt{E}$ and $s^2 \in \mathbb{F}_{2^p}^m$ iff μ^{D} is sampled using a non-trivial seed in the sampling procedure.

Completeness. Let $\mathcal{S} = (\mathcal{A}, \{\tilde{A}_a^x\}, \mathcal{H}, |\tau\rangle)$ be a perfect oracularizable synchronous strategy for $\mathcal{G}$. We define an oracularizable synchronous strategy for $\mathcal{G}^{\mathrm{D}}$ as follows: Given a question label $x \in \mathcal{X}$, if there exists some $v \in \mathtt{T}$ and $s^2 \in \mathbb{F}_{2^p}^m$ such that $x = (\text{view}(v), \mathtt{L}^v(s^2))$, then set $A_a^x = \tilde{A}_a^{(v, \mathtt{L}^v(s^2))}$. If x cannot be parsed in the above format, the prover always returns some predetermined fixed element $\star \in \mathcal{A}$. This ensures that $A_\star^x = \mathbb{I}$ and $A_a^x = 0$ for all $a \in \mathcal{A} \setminus \{\star\}$).

Restricted to the question set where (x, y) are parsed correctly, we see that A_a^x is the same as $\tilde{A}_a^{(v, \mathtt{L}^v(s^2))}$ with the same decider function. This implies that A_a^x is a perfect oracularizable synchronous strategy when restricted to this case. Otherwise, since A_a^x returns the same answer, and the only non-trivial question pair in this scenario is the synchronicity question pair. This strategy passes with perfect accuracy. This means that A_a^x remains a perfect oracularizable synchronous strategy in the case where at least one of x, y cannot be parsed correctly, hence showing that the strategy given above is a perfect oracularizable strategy for $\mathcal{G}^{\mathrm{D}}$.

Soundness. Let $\mathcal{S} = (\mathcal{A}, \{A_a^x\}, \{(B_b^y)^{op}\}, |\tau\rangle)$ be a tracially embeddable strategy for $\mathcal{G}^{\mathrm{D}}$ with a success rate of $1-\varepsilon$. By Lemma 7.2.11, with probability $\frac{1}{4|\mathtt{T}|^2 \cdot 16^{|\mathtt{T}|}}$, the output from μ^{D} would be from a non-trivial seed. Conditioning on this case, the distribution μ^{D} is exactly the same as μ. We define the strategy for $\mathcal{G}$ with the same measurement state/algebra as $\mathcal{S}$, but with the measurement operator replaced by $\hat{A}_a^{(v, \mathtt{L}^v(s^2))} = A^{(\text{view}(v), \mathtt{L}^v(s^2))}$. This strategy will succeed at $\mathcal{G}$ with probability at least $1 - 4|\mathtt{T}|^2 \cdot 16^{|\mathtt{T}|} \cdot \varepsilon$. This implies if $\omega^t(\mathcal{G}) \geq 1-\varepsilon$, then

$$\omega^t(\mathcal{G}^{\mathrm{D}}) \geq 1 - \frac{\varepsilon}{\left(4|\mathtt{T}|^2 \cdot 16^{|\mathtt{T}|}\right)},$$

which completes the claim. □

In this thesis, $|\mathtt{T}|$ is typically taken to be the complexity bound of $O(\log(n))$ for some integer n or some constant. Hence the increase in soundness shown in the above lemma can still be "reset" using the parallel repetition presented in Chapter 12.

7.3 Conditionally linear verifier

We are now ready to define a conditional linear verifier, for $t \in \{*, co\}$ and constant $k \in \mathbb{N}$, we define the complexity class synchronous k-th level CL samplable MIP^t (k-CLMIP^t) as the complexity class MIP^t except restricted to synchronous games which are also k-th level CL samplable. This complexity class is complete with respect to the following decision problem.

- $\mathsf{L}_{\text{yes}}^{k\text{-}\mathsf{CLMIP}^t} = \{\langle \mathcal{G} \rangle | \omega^t(\mathcal{G}) = 1, \mathcal{G} \text{ is a synchronous } k\text{-th level CL samplable game}\}$,
- $\mathsf{L}_{\text{no}}^{k\text{-}\mathsf{CLMIP}^t} = \{\langle \mathcal{G} \rangle | \omega^t(\mathcal{G}) \leq \frac{1}{2}, \mathcal{G} \text{ is a synchronous } k\text{-th level CL samplable game}\}$.

Showing that k-CLMIP^{co} (resp. k-CLMIP^*) being coRE-complete (resp. RE-complete) implies that $\mathsf{coRE} \subseteq \mathsf{MIP}^{\text{co}}$ (resp. $\mathsf{RE} \subseteq \mathsf{MIP}^*$).

We formally introduce the notion of a CL verifier below. As mentioned back in Section 6.2, one should think of the CL verifier as a more structured version of a uniform verifier for synchronous k-th level CL samplable games.

Definition 7.3.1 (Conditionally linear verifier)**.** *Let $\mathbf{k}(n), \mathbf{m}(n), \mathbf{p}(n) : \mathbb{N} \to \mathbb{N}$, where the range of $\mathbf{p}(n)$ maps integers to odd integers. Let $\mathcal{G} = \{\mathcal{G}_n = (\mathcal{X}_n, \mathcal{A}_n, \mu_n, D_n)\}_{n\in\mathbb{N}}$ be an infinite sequence of games indexed by $n \in \mathbb{N}$. Each μ_n is a $(\mathbf{k}(n), \mathbf{m}(n), \mathbf{p}(n))$ CL distribution defined over two $\mathbf{k}(n)$-level CL functions $\mathsf{L}^{0,n}, \mathsf{L}^{1,n}$ with registers $\{V_j^n\}_{j\in[\mathbf{k}(n)]}$ as defined in Definition 7.2.5, and $\mathcal{A}_n = \{0,1\}^*$ (in this case, $\mathcal{X}_n = \mathbb{F}_{2^{\mathbf{p}(n)}}^{\mathbf{m}(n)} = \{0,1\}^{\mathbf{p}(n)\cdot\mathbf{m}(n)}$ by definition).*

A $\mathbf{k}(n)$ level CL verifier $\mathcal{V}$ is a tuple $(\mathsf{Q}_\mathcal{V}, \mathsf{D}_\mathcal{V})$, where $\mathsf{Q}_\mathcal{V}$ is a five-input Turing machine, and $\mathsf{D}_\mathcal{V}$ is a six-input Turing machine, such that, for all $n \in \mathbb{N}$

- $\mathsf{Q}_\mathcal{V}(n, \textit{Parameter}) = (\mathbf{k}(n), \mathbf{m}(n), \mathbf{p}(n))$.
- $\mathsf{Q}_\mathcal{V}(n, \textit{Divide}, s) = (s_0, \cdots s_{\mathbf{k}(n)-1})$, *for all $s \in V^n$, where $s_j \in V_j^n$ and $\sum_{j\in[\mathbf{k}(n)]} s_j = s$.*
- $\mathsf{Q}_\mathcal{V}(n, \textit{Function}, p, j, s, x) = \mathsf{L}_{j,s}^{p,n}(x)$, *for all $j \in [\mathbf{k}(n)]$, $s \in V_{<j}^n$, $x \in V_j^n$, and $p \in \{0,1\}$. Where recall $\{\mathsf{L}_{j,s}^{p,n}\}_{s\in V_{<j}^n}$ are the jth level linear function for $\mathsf{L}^{p,n}$ (where we associate $\mathsf{L}^{0,n} = \mathsf{L}^{A,n}$ and $\mathsf{L}^{1,n} = \mathsf{L}^{B,n}$).*
- $\mathsf{D}_\mathcal{V}(n, x, y, a, b) = D_n(x, y, a, b)$ *for all $(x, y) \in \mathcal{X}_n^2$ and $(a, b) \in \mathcal{A}_n^2$*

If $\mathbf{k}(n) = k$ for some constant $k \in \mathbb{N}$, then we simply call $\mathcal{V}$ a k-th level CL verifier. We say that $\mathcal{V}$ has a sampling complexity of $O(\mathbf{f}(n))$ if

$$\mathbf{k}(n) \cdot \mathbf{m}(n) \cdot \mathbf{p}(n) = \mathsf{TIME}_{\mathsf{P}_\mathcal{V}}(n) = \mathsf{TIME}_{\mathsf{Q}_\mathcal{V}}(n) = O(\mathbf{f}(n)),$$

and we say that $\mathcal{V}$ has a verification complexity of $O(\mathbf{g}(n))$ if $\mathsf{TIME}_{\mathsf{D}_\mathcal{V}} = O(\mathbf{g}(n))$.

We use the term CL verifier as a shorthand for any $\mathbf{k}(n)$ verifier. We refer to a CL verifier $\mathcal{V}$ as a *k-th level synchronous Conditionally linear verifier* if every game $\mathcal{G}_n = \{\mathcal{X}_n, \mathcal{A}_n, \mu_n, D_n\}$ generated by $\mathcal{V}$ is a k-th level CL samplable synchronous game. To abuse notation, we write $\mathsf{Q}_\mathcal{V}(n, \text{Parameter}) \leq O(\mathbf{f}(n))$ as a shorthand for $\mathbf{m}(n)\cdot\mathbf{p}(n) \leq O(\mathbf{f}(n))$, where $\mathbf{m}(n)$ and $\mathbf{p}(n)$ are the output for $\mathsf{Q}_\mathcal{V}(n, \text{Parameter})$. In the above definition, we refer to $\mathsf{Q}_\mathcal{V}$ as a CL sampler and $\mathsf{D}_\mathcal{V}$ as a CL decider, and we drop the subscript $\mathcal{V}$ if the underlying game sequence is clear from context. We remark that although in the above definition, the answer set $\mathcal{A}_n = \{0,1\}^*$ is technically an infinite set. However, since the decider D is always assumed to be time-bounded, D can read at most $\mathsf{TIME}_\mathsf{D}(n)$ bits; thus $\mathcal{A}_n$ is, in practice, always a finite set.

Although the above definition does not explicitly give the computational procedure $\mathsf{Q}_\mathcal{V}(n, \text{sample})$ akin to Definition 6.1.3, the computation procedure can be implemented in the following manner:

1. The verifier first runs $\mathtt{Q}_{\mathscr{V}}(n, \text{Parameter})$ to obtain $\mathbf{k}(n)$ and $\mathbf{m}(n)\cdot\mathbf{p}(n)$, then the verifier samples a random seed $s \in \{0,1\}^{\mathbf{m}(n)\cdot\mathbf{p}(n)}$.

2. The verifier then runs $\mathtt{Q}_{\mathscr{V}}(n, \text{Divide}, s) = (s_1 \cdots, s_{\mathbf{k}(n)})$ to partition s into linear components within $\{V_j^n\}_{j\in[\mathbf{k}(n)]}$.

3. For $p \in \{0,1\}$, the verifier performs the following:

   (a) The verifier first computes $x_0^p = \mathtt{Q}_{\mathtt{L}}(n, \text{Function}, p, 0, 0, s_0)$, the output for the 0-th linear function for $\mathtt{L}^{p,n}$.

   (b) For $1 \le j < \mathbf{k}(n)$, the verifier computes $x_j^p = \mathtt{Q}_{\mathtt{L}}(n, \text{Function}, p, j, x_{j-1}^p, s_j)$, the output for the jth linear function for $\mathtt{L}^{p,n}$.

   (c) Finally, the verifier computes $x^p = \sum_j x_j^i$ by adding up all the components together.

4. The verifier returns (x^0, x^1), as the question pair.

Step 1 takes $O(\log(\mathbf{p}(n)) + \log(\mathbf{m}(n))) = O(\mathbf{f}(n))$ time, Step 2 and 3 take $O(\mathbf{k}(n)\cdot\mathbf{m}(n)\cdot\mathbf{p}(n)) = O(\mathbf{f}(n))$ time by Lemma 7.1.1. This implies that $\mathtt{Q}_{\mathscr{V}}(n, \text{sample})$ runs in time $O(\mathbf{f}(n))$, consistent with the definition given in Definition 6.1.3. This shows that a Conditionally linear verifier is a more "refine" version of a uniform verifier sequence given in Definition 6.1.3.

7.4 Proof outline for the compression theorem

Given the formal definition of the conditional linear verifier, we are ready to give a proof for Theorem 6.2.1. For convenience, we restate the theorem below.

Theorem 7.4.1 (Gap compression for non-local games)**.** *For all constants $\alpha, k \in \mathbb{N}$, there exists an algorithm $\mathtt{Gap compress}_{\alpha,k}$ that takes as input a pair of Turing machines $(\mathtt{Q}, \mathtt{D})$. $\mathtt{Gap compress}_{\alpha,k}$ outputs a tuple of Turing machines $\mathscr{V}^{Comp} = (\mathtt{Q}^{Comp}, \mathtt{D}^{Comp})$ such that the following holds:*

There exists an integer $\gamma = O(poly(\alpha, k))$ such that, for models $t \in \{*, co\}$

1. *(Runtime):* $\mathsf{TIME}_{\mathtt{Gapcompress}_{\alpha,k}}(\mathtt{Q}, \mathtt{D}) = O(\text{poly}(\alpha), |\mathtt{Q}|, |\mathtt{D}|)$.

2. *(Independence of the sampler): The Turing machine $\mathtt{Q}^{Comp}$ only depends on the parameter α and k and is a sampler for a synchronous 7-th level CL verifier. The Turing machine $\mathtt{D}^{Comp}$ is a decider for a synchronous 7-th level CL verifier. $\mathscr{V}^{Comp}$ is a synchronous 7-th level CL verifier sequence for an infinite sequence of games $\mathscr{G}^{Comp} = \{\mathscr{G}_n^{Comp}\}_{n\in\mathbb{N}}$.*

3. *(Complexity bounds for the output) $\mathscr{V}^{Comp}$ has sample complexity and verification complexity of $O(\log(n)^{\gamma})$.*

Furthermore, if the input $\mathscr{V} = (\mathtt{Q}, \mathtt{D})$ is a synchronous k'th level CL verifier for the infinite sequence of synchronous games $\mathscr{G} = \{\mathscr{G}_n\}_{n\in\mathbb{N}}$ for some constant $k' < k$, and there exists a constant $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$

$$\max\{\mathsf{TIME}_{\mathtt{Q}}, \mathsf{TIME}_{\mathtt{D}}\} \le n^{\alpha}. \tag{7.9}$$

Then there exists a constant $n_0^{Comp} = \text{poly}(\gamma, n_0)$ such that for all $n \ge n_0^{Comp}$

- *(Completeness) If there exists a perfect oracularizable strategy for $\mathscr{G}_n$ in model t, then there exists a perfect oracularizable strategy for $\mathscr{G}_n^{Comp}$ in model t.*

- *(Soundness)*

$$\omega^t(\mathscr{G}_n) \le \frac{1}{2} \Longrightarrow \omega^t(\mathscr{G}_n^{Comp}) \le \frac{1}{2}$$

Similar to [JNV+22a, Theorem 12.1], the proof of Theorem 6.2.1 relies on three components: Question Reduction, Answer reduction and Parallel repetition, which we state below. The first proposition is question reduction. This proposition states the existence of an algorithm which takes a kth CL verifier and outputs a fourth level CL verifier with $O(\text{polylog}(n))$ sample complexity without drastically increasing the verification complexity and the soundness condition.

Proposition 7.4.2 (Question Reduction)**.** *For all constants $\alpha, k \in \mathbb{N}$, there exists a polynomial time algorithm $\texttt{QuestionReduction}_{\alpha,k}$ that takes, as input, a pair of Turing machines $(\mathtt{Q},\mathtt{D})$ and outputs a tuple of Turing machine machines $(\mathtt{Q}^{QR},\mathtt{D}^{QR})$ such that the following holds:*

For models $t \in \{*, co\}$, $\texttt{QuestionReduction}_{\alpha,k}$ outputs a pair of Turing machine $(\mathtt{Q}^{QR},\mathtt{D}^{QR})$ which is a fourth-level CL-verifier $\mathscr{V}^{QR}$ for a infinite sequence of game $\mathscr{G}^{QR} = \{\mathscr{G}_n^{QR}\}_{n\in\mathbb{N}}$ with the following properties: There exist an integer $\gamma^{QR} = O(poly(\alpha))$ such that

1. *(Computation time):* $\mathsf{TIME}_{\texttt{QuestionReduction}_{\alpha,k}}(\text{poly}(\alpha), |\mathtt{Q}|, |\mathtt{D}|)$.
2. *(Synchronicity) The game sequence $\mathscr{V}^{QR}$ is a synchronous game sequences.*
3. *(Complexity bounds for the output):*
   - $\mathtt{Q}^{QR}(n, Parameter) \leq \max\left\{\log^{\gamma^{QR}}(n), C^{trivial}\right\}$,
   - $\mathsf{TIME}_{\mathtt{Q}^{QR}}(n) \leq \max\left\{\log^{\gamma^{QR}}(n), C^{trivial}\right\}$,
   - $\mathsf{TIME}_{\mathtt{D}^{QR}} \leq \max\left\{n^{\gamma^{QR}}, C^{trivial}\right\}$,

   for some universal constant $C^{trivial}$.
4. *(Independency) $\mathtt{Q}^{QR}$ is a 3-th level CL sampler which only depends on the parameter α and k and does not depend on $|\mathtt{D}|$, $|\mathtt{D}^{QR}| = O(\text{poly}(\alpha, k))$.*

Furthermore if the input $\mathscr{V} = (\mathtt{Q},\mathtt{D})$ is a k'th CL verifier for the infinite sequence of synchronous game $\mathscr{G} = \{\mathscr{G}_n\}_{n\in\mathbb{N}}$ for some constant $k' \in \mathbb{N}$ such that $k' < k$, and there exist some constant $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$

$$\max\left\{\mathsf{TIME}_{\mathtt{Q}}, \mathsf{TIME}_{\mathtt{D}}\right\} \leq n^{\alpha}.$$

Then there exist some $n_0^{QR} = \text{poly}(\gamma^{QR}, n_0)$ with $n_0 \leq n_0^{QR}$ such that for all $n \geq n_0^{QR}$

1. *(Completeness) If there exists a perfect oracularizable synchronous strategy for $\mathscr{G}_n$ in model t, then there exists a perfect oracularizable synchronous strategy for $\mathscr{G}_n^{QR}$ in model t.*
2. *(Soundness) There exist some universal function $\mathbf{s}_\alpha^{QR}$ which depends on k and ε with $\mathbf{s}_\alpha^{QR} = O(\exp(k), \text{poly}(\varepsilon))$ such that, for any polynomial $\varepsilon : \mathbb{N} \to [0,1]$*

$$\omega^t(\mathscr{G}_n) \leq 1 - \varepsilon(n) \Longrightarrow \omega^t(\mathscr{G}_n^{QR}) \leq 1 - \mathbf{s}_\alpha^{QR}(k, \varepsilon(n)).$$

Question Reduction relies on self-testing techniques used in [NV18; Gri20], which are unique to non-local games with entangled provers. The question reduction procedure is a combination between two protocols, the Pauli basis test, which we introduce in Chapter 8 and the Introspection protocol, which we introduce in Chapter 9. We give a proof to the above proposition in Section 9.3. We remark that in comparison to the question procedure given in [JNV+22a], there is no dependency on the parameter n and λ for soundness since the variant of Pauli basis test used in this thesis does not require the low-degree test.

The second proposition is Answer reduction, which gives a algorithm which takes synchronous CL verifier with $O(\text{poly}(n))$ verification complexity and outputs a balanced synchronous CL verifier with $O(\text{polylog}(n))$ verification complexity which does not increase the sample complexity and only increase soundness by another $\text{polylog}(n)$.

Proposition 7.4.3 (Answer Reduction)**.** *For all constant $(\alpha, k) \in \mathbb{N}$ there exists a polynomial time algorithm $\texttt{AnswerReduction}_{\alpha,k}$ that takes, as input, a pair of Turing machines $(\mathtt{Q}, \mathtt{D})$ and outputs a tuple of Turing machine machines $(\mathtt{Q}^{AR}, \mathtt{D}^{AR})$ such that the following holds:*

There exist a polynomial For models $t \in \{*, co\}$, $\texttt{AnswerReduction}_{\alpha}$ outputs a pair of Turing machines $(\mathtt{Q}^{AR}, \mathtt{D}^{AR})$, which defines a synchronous CL-verifier $\mathscr{V}^{AR}$ for an infinite sequence of game $\mathscr{G}^{AR} = \{\mathscr{G}_n^{AR}\}_{n\in\mathbb{N}}$ with the following properties: There exists an integer $\gamma^{AR} = O(poly(\alpha))$ such that

1. *(Runtime): $\texttt{AnswerReduction}_{\alpha}$ have runtime*

$$\mathsf{TIME}_{\texttt{AnswerReduction}_{\alpha}}(\mathrm{poly}(\alpha), |\mathtt{Q}|, |\mathtt{D}|).$$

2. *(Dependency for $\mathtt{Q}^{AR}$) The Turing machine $\mathtt{Q}^{AR}$ only depends on the input $\mathtt{Q}$ and α.*

Furthermore, if the input $\mathscr{V} = (\mathtt{Q}, \mathtt{D})$ is a $\max\{4, k+2\}$ level synchronous CL verifier for the infinite sequence of synchronous game $\mathscr{G} = \{\mathscr{G}_n\}_{n\in\mathbb{N}}$ for some constant $k \in \mathbb{N}$, and there exists some constant $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, we have

- $|\mathtt{D}^{QR}| = O(\mathrm{poly}(\alpha, k))$.
- $\mathtt{Q}(n, Parameter) \leq \log^{\alpha}(n)$,
- $\mathsf{TIME}_{\mathtt{Q}}(n) \leq \log^{\alpha}(n)$,
- $\mathsf{TIME}_{\mathtt{D}} \leq n^{\alpha}$.

Then there exists some integer $n_0^{AR} = \mathrm{poly}(\gamma^{AR}, n_0)$ with $n_0 \leq n_0^{AR}$ such that for all $n \geq n_0^{AR}$

1. *(Complexity bounds for the output):*
   - $\mathsf{TIME}_{\mathtt{Q}^{AR}}(n) \leq \max\left\{\log^{\gamma^{AR}}(n), C^{trivial}\right\}$,
   - $\mathsf{TIME}_{\mathtt{D}^{AR}}(n) \leq \max\left\{\log^{\gamma^{AR}}(n), C^{trivial}\right\}$,

   for some universal constant $C^{trivial}$.
2. *(Level for the CL sampler) The Turing machine $\mathtt{Q}^{AR}$ is a level $\max\{k+2, 5\}$ CL sampler.*
3. *(Completeness) If there exists a perfect oracularizable strategy for $\mathscr{G}_n$ in model t, then there exists a perfect oracularizable strategy for $\mathscr{G}_n^{AR}$ in model t.*
4. *(Soundness) There exist a universal function $\mathbf{s}_{\alpha}^{AR}$ which depends on n and ε with $O(\mathbf{s}_{\alpha}^{AR}) = O(\mathrm{polylog}(n), \mathrm{poly}(\varepsilon))$ such that, for any polynomial $\varepsilon : \mathbb{N} \to [0,1]$*

$$\omega^t(\mathscr{G}_n) \leq 1 - \varepsilon(n) \Longrightarrow \omega^t(\mathscr{G}_n^{QR}) \leq 1 - \mathbf{s}_{\alpha}^{AR}(\varepsilon(n), n)$$

We remark that the universal constant C^{trivial} comes from $\mathscr{G}^{\text{reject}}$. The answer reduction procedure we use is identical to the one used in [JNV+22a, Chapter 10], which is a modification of the PCP of proximity based on techniques from [BFL91]. We further remark that due to the technique used, $\texttt{AnswerReduction}_{\alpha}$ actually works for **all** MIP, MIP* and MIP^{co}. We give a more in depth description of the answer reduction protocol in Chapter 10 and Chapter 11. The third step is the parallel repetition theorem for anchored games.

Proposition 7.4.4 (Parallel repetition)**.** *For all constants $\alpha \in \mathbb{N}$, function $\mathbf{s}(n) : \mathbb{N} \to [0,1]$ with $O(\mathbf{s}(n)) = O(\text{polylog}(n))$. Then there exists a function $\mathbf{r}(n) : \mathbb{N} \to \mathbb{N}$ with $\mathbf{r}(n) = O(\alpha, \mathbf{s}(n))$ and a polynomial time algorithm $\mathtt{Parallelrep}_{\alpha,\mathbf{s}(n)}$ that takes, as input, a pair of Turing machines $(\mathtt{Q},\mathtt{D})$ and outputs a tuple of Turing machine machines $(\mathtt{Q}^{Pararep},\mathtt{D}^{Pararep})$ with*

$$\mathsf{TIME}_{\mathtt{Parallelrep}_{\alpha,\mathbf{s}(n)}}(|\mathtt{Q}|,|\mathtt{D}|) = O(\text{polylog}(n))$$

such that the following holds:

1. *(Independency) $\mathtt{Q}^{Pararep}$ only dependent on $\mathtt{Q}$ and the polynomial functions $\mathbf{s}(n)$.*
2. *(Complexity bounds for the output):*
   - $\mathsf{TIME}_{\mathtt{Q}^{Pararep}}(n) \leq O(\text{poly}(\mathbf{r}(n), log^{\alpha}(n)))$,
   - $\mathsf{TIME}_{\mathtt{D}^{Pararep}}(n) \leq O(\text{poly}(\mathbf{r}(n), log^{\alpha}(n)))$,

Furthermore, if the input $\mathscr{V} = (\mathtt{Q},\mathtt{D})$ is a k level synchronous CL verifier for the infinite sequence of synchronous game $\mathscr{G} = \{\mathscr{G}_n\}_{n\in\mathbb{N}}$ for some constant $k \in \mathbb{N}$, and there exists some constant $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, we have

$$\mathtt{Q}(n, Parameter), \mathsf{TIME}_{\mathtt{Q}}(n), \mathsf{TIME}_{\mathtt{D}} \leq \log^{\alpha}(n).$$

Then for all $n \geq n_0$

1. *(Parameter) $\mathtt{Q}^{AR}$ is a level $k+1$ CL sampler.*
2. *(Completeness) If there exists a perfect oracularizable strategy for $\mathscr{G}_n$ in model t, then there exists a perfect oracularizable strategy for $\mathscr{G}_n^{Pararep}$ in model t.*
3. *(Soundness)*
$$\omega^t(\mathscr{G}_n) \leq 1 - \mathbf{s}(n) \Longrightarrow \omega^t(\mathscr{G}_n^{Pararep}) \leq \frac{1}{2}$$

The anchored parallel repetition theorem is proven for the tensor product model in [BVY21] and we given a proof for the commuting operator model in Chapter 12. We remark that the theorem above actually works for all verifier as well. We show that the anchor transformation and parallel repetition of a kth level synchronous CL verifier becomes a $k+1$th level synchronous CL verifier in Chapter 12.

We are now ready to give a proof for Theorem 6.2.1 below, which is just applying the three propositions above in sequence.

Proof. Given constant $(\alpha, k) \in \mathbb{N}$, we specify the pseudocode for $\mathtt{Gapcompress}_{\alpha,k}$ as the following

Input: Turing Machine $(\mathtt{Q},\mathtt{D})$.
Compute $\mathscr{V}^{\text{QR}} = (\mathtt{Q}^{\text{QR}},\mathtt{D}^{\text{QR}}) = \mathtt{QuestionReduction}_{\alpha,k}(\mathtt{Q},\mathtt{D})$.
Compute $\alpha^{\text{QR}} = \mathbf{s}^{\text{QR}}(\alpha)$, where $\mathbf{s}^{\text{QR}}$ is the function used to define γ^{QR} in Proposition 7.4.2.
Compute $\mathscr{V}^{\text{AR}} = (\mathtt{Q}^{\text{AR}},\mathtt{D}^{\text{AR}}) = \mathtt{AnswerReduction}_{\alpha^{\text{QR}}}(\mathtt{Q}^{\text{QR}},\mathtt{D}^{\text{QR}})$.
Compute $\alpha^{\text{AR}} = \mathbf{s}^{\text{AR}}(\alpha^{\text{QR}})$, where $\mathbf{s}^{\text{AR}}$ is the function used to define γ^{AR} in Proposition 7.4.3.
Compute the description of the function $\mathbf{s}(n) = \mathbf{s}^{\text{AR}}_{\alpha^{\text{AR}}}(\mathbf{s}^{\text{QR}}_{\alpha}(\frac{1}{2}, n), n)$. Where $\mathbf{s}^{\text{AR}}_{\alpha^{\text{AR}}}$ (resp. $\mathbf{s}^{\text{QR}}_{\alpha}$) is the function used in the soundness condition in Proposition 7.4.3 (resp. Proposition 7.4.2).
Return $\mathscr{V}^{\text{Comp}} = (\mathtt{Q}^{\text{Comp}},\mathtt{D}^{\text{Comp}}) = \mathtt{Parallelrep}_{\alpha^{\text{AR}},\mathbf{s}(n)}(\mathtt{Q}^{\text{AR}},\mathtt{D}^{\text{AR}})$.

Pseudocode 11: The description for $\mathtt{Gapcompress}_{\alpha,k}$.

	Time Complexity			
Verifier	Sampler	Decider	Level	Soundness
$\mathscr{V}^{\mathrm{QR}}$	$\leq \log^{O(\mathrm{poly}(\alpha))}(n)$	$\leq n^{O(\mathrm{poly}(\alpha))}$	3	$\omega^t(\mathscr{G}_n) \leq \frac{1}{2} \to \omega^t(\mathscr{G}_n^{\mathrm{QR}}) \leq 1 - \mathbf{s}_\alpha^{\mathrm{QR}}(\frac{1}{2}, n)$
$\mathscr{V}^{\mathrm{AR}}$	$\leq \log^{O(\mathrm{poly}(\alpha))}(n)$	$\leq \log^{O(\mathrm{poly}(\alpha))}(n)$	5	$\to \omega^t(\mathscr{G}_n^{\mathrm{AR}}) \leq 1 - \mathbf{s}_{\alpha^{\mathrm{AR}}}^{\mathrm{AR}}(\mathbf{s}_\alpha^{\mathrm{QR}}(\frac{1}{2}, n), n)$
$\mathscr{V}^{\mathrm{Comp}}$	$\leq \log^{O(\mathrm{poly}(\alpha))}(n)$	$\leq \log^{O(\mathrm{poly}(\alpha))}(n)$	6	$\to \omega^t(\mathscr{G}_n^{\mathrm{Comp}}) \leq \frac{1}{2}$

Table 7.1: The time complexity for the 3 CL verifiers listed in Pseudocode 11 and the soundness statement for the verifier sequences assuming the input $(\mathtt{Q},\mathtt{D})$ is a synchronous k'th-level CL verifier for $k' \leq k$. We remark that only the last column is dependent upon the input $(\mathtt{Q},\mathtt{D})$ being a synchronous CL verifier with the appropriate runtime condition.

For simplicity, we listed how the parameter changes throughout Pseudocode 11 in Table 7.1. We verify each of the clause in Theorem 6.2.1 below:

- (Level of the CL verifier and the output being a synchronous game): Since the output for $\mathtt{AnswerReduction}_{\alpha^{\mathrm{QR}}}$ is always a synchronous CL verifier, and $\mathtt{Parallelrep}_{\alpha^{\mathrm{AR}},\mathbf{s}(n)}$ retains synchronicity for a CL verifier. The output for $\mathtt{Gapcompress}_{\alpha,k}$ will always be a synchronous game sequence. The level for the output of $\mathtt{Gapcompress}_{\alpha,k}$ are tracked in Table 7.1.

- (Computation time) We first see that both α^{AR} and $\mathbf{s}(n)$ in Pseudocode 11 are polynomial functions of α, independent from the input $(\mathtt{Q},\mathtt{D})$. Hence, both steps 2 and 4 can be computed in $O(\mathrm{poly}(\alpha))$ time (or hardcoded into the description of $\mathtt{Gapcompress}_{\alpha,k}$). By the similar reasoning, we have $\mathsf{TIME}_{\mathtt{QuestionReduction}_{\alpha,k}} = \mathsf{TIME}_{\mathtt{AnswerReduction}_{\alpha^{\mathrm{QR}}}} = \mathsf{TIME}_{\mathtt{Parallelrep}_{\alpha^{\mathrm{AR}},\mathbf{s}(n)}} = O(\mathrm{poly}(\alpha), |\mathtt{Q}|, |\mathtt{D}|)$. Hence, $\mathsf{TIME}_{\mathtt{Gapcompress}_{\alpha,k}}(\mathrm{poly}(\alpha), |\mathtt{Q}|, |\mathtt{D}|)$.

- (Independency of the sampler) By Proposition 7.4.4, $\mathtt{Q}^{\mathrm{Comp}}$ only depends on the polynomial $s(n)$ (which itself only depends on α) and $\mathtt{Q}^{\mathrm{AR}}$. $\mathtt{Q}^{\mathrm{AR}}$ depends, in addition to the parameter α, on both $\mathtt{Q}^{\mathrm{QR}}$, and only on $|\mathtt{D}^{\mathrm{QR}}|$ by definition. Since $|\mathtt{D}^{\mathrm{QR}}| = O(\mathrm{poly}(\alpha, k))$ given any $\mathtt{D}$ as input for $\mathtt{QuestionReduction}_{\alpha,k}$. $\mathtt{Q}^{\mathrm{Comp}}$ only depends on the parameters α and k, as claimed.

- (Complexity bounds for the output) This follows from the complexity parameter, which we kept track of in Table 7.1, and the complexity bound for the outputs of $\mathtt{AnswerReduction}_{\alpha^{\mathrm{QR}}}$, which does not depend on the input.

Now, assume the input $\mathscr{V} = (\mathtt{Q},\mathtt{D})$ is a synchronous k'th level CL verifier for the infinite sequence of synchronous game $\mathscr{G} = \{\mathscr{G}_n\}_{n\in\mathbb{N}}$ for some constant $k' \in \mathbb{N}$ with $k' < k$, and some constant $n_0 \in \mathbb{N}$ which satisfies (6.1). Take n_0^{QR} be the constant guarantee by Proposition 7.4.2 and take n_0^{Comp} to be the constant n_0^{AR} guarantee by Proposition 7.4.3 (where in this case, $n_0^{\mathrm{AR}} = \mathrm{poly}(n_0^{\mathrm{QR}}, \alpha^{\mathrm{QR}})$). Fix $n > n_0^{\mathrm{Comp}}$ and model $t \in \{*, co\}$.

- (Completeness) Since the game sequence $\mathscr{V}$ is synchronous. Hence, any perfect strategy for $\mathscr{G}_n$ is also a synchronous strategy. Since $\mathtt{QuestionReduction}_{\alpha,k}$, $\mathtt{AnswerReduction}_{\alpha^{\mathrm{QR}}}$ and $\mathtt{Parallelrep}_{\alpha^{\mathrm{AR}},\mathbf{s}(n)}$ preserve the existence of a perfect oracularizable strategy for $\mathscr{G}_n$ in model t, $\mathtt{Gapcompress}_{\alpha,k}$ also preserves the existence of a perfect oracularizable strategy for $\mathscr{G}_n$ in model t.

- (Soundness) Assume that $\omega^t(\mathscr{G}_n) \leq \frac{1}{2}$; By the soundness property of $\mathtt{QuestionReduction}_{\alpha,k}$, the previous condition implies that $\omega^t(\mathscr{G}_n^{\mathrm{QR}}) \leq 1 - \mathbf{s}_\alpha^{\mathrm{QR}}(\frac{1}{2}, n)$, which, by the soundness property of $\mathtt{AnswerReduction}_{\alpha^{\mathrm{QR}}}$, implies that $\omega^t(\mathscr{G}_n^{\mathrm{QR}}) \leq 1 - \mathbf{s}_{\alpha^{\mathrm{AR}}}^{\mathrm{AR}}(\mathbf{s}_\alpha^{\mathrm{QR}}(\frac{1}{2}, n), n)$. The completeness condition follows from the soundness condition of $\mathtt{Parallelrep}_{\alpha^{\mathrm{AR}},\mathbf{s}(n)}$.

□

CHAPTER 8

Question Reduction part 1: The μ-dependent Pauli Basis test

In the next two chapters, we give a proof for Proposition 7.4.2 by showing an algorithm that takes as input a synchronous CL verifier and transforms it into another synchronous CL verifier with a lower sampling complexity. This chapter focuses on the n-qubit Pauli Basis test, which we introduce in Section 8.3.2. The n^α-qubit Pauli Basis test is a subroutine that forces the provers to perform either all X or all Z Pauli measurements on n^α EPR pairs. A key property is that these tests are "rigid", meaning that if the provers perform this test with near-optimal success rate, any strategy they perform must be "close" to the "measure all X or all Z on n^α EPR pairs strategy". One key observation is that forcing the provers to perform Pauli measurements on EPR pairs allows them to sample n^α independent random bits, which the verifier can take advantage of during the introspection protocol.

There are several protocols which certify EPR pairs in the tensor product model (see, for example, [NV17; JNV+22a]). These tests typically require the verifier to perform cross-verification with some subset drawn uniformly from the EPR pairs. The protocol introduced in this chapter is based on Mikael de la Salle's elegant simplification of the "answer reduction" protocol from [JNV+22a]. We expand on his idea and introduce the μ-dependent Pauli Basis test (Figure 8.3), whereby the referee can, instead, choose an arbitrary distribution μ over the subsets of qubits to be used for cross-checking. Before we introduce the protocol, we first introduce some additional preliminaries needed for this chapter.

8.1 Preliminaries on representation theory

Before introducing this protocol, we introduce two tools from trace norm representation theory which will be important for the analysis of the μ-dependent Pauli Basis test, the state-dependent Gowers-Hatami theorem and the spectral gap inequality based on the well known well-known Poincaré inequality.

8.1.1 State-dependent Gowers-Hatami theorem

In this section, we introduce an important tool for proving rigidity theorem, the state dependent variant of the Gowerss-Hatami theorem [GH17] by Vidick [Vid18, Theorem 4], and present a generalization to the tracial von Neumann algebra setting. For $\sigma \in \mathscr{A}^+$ such that $\tau(\sigma) = 1$, we define the σ-seminorm on $\mathscr{A}$ as $||A||_\sigma = \sqrt{\tau(A^* A\sigma)}$.

This can be seen as a tracial von Neumann algebra analogue of the state-dependent norm defined within [Vid18, Section 1.3]. By a quick calculation, we see that $\|UA\|_\sigma = \|A\|_\sigma$ for all $U \in \mathscr{U}(\mathscr{A})$. For any finite group G and the map $\phi : G \to \mathscr{U}(\mathscr{A})$, we say ϕ is a *unitary representation* if

$\phi(g)\phi(h) = \phi(gh)$ for all $g, h \in \mathcal{G}$. We say the map ϕ is an $(\varepsilon, \sigma)-$representation for some $\sigma \in \mathcal{A}^+$ with $\tau(\sigma) = 1$ if $\phi(g^{-1}) = \phi(g)^*$ and $\mathbb{E}_{g,h\in G} \|\phi(g)\phi(h) - \phi(gh)\|_\sigma \leq \varepsilon$.

Theorem 8.1.1 (Gowerss-Hatami for tracial von Neumann algebras)**.** *Let G be a finite group, $\mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$ a tracial von Neumann algebra in standard form, $\varepsilon \geq 0$ and $\phi : G \to \mathcal{U}(\mathcal{A})$ be an (ε, ρ)-representation for some $\rho \in \mathcal{A}^+$ such that $\tau(\rho) = 1$. Then there exists some isometry $V : \mathcal{H} \to \mathcal{H} \otimes \mathbb{C}^{|G|}$ such that for all $B \in \mathcal{A}'$, $VB = (B \otimes I_{|\mathcal{G}|})V$ and a unitary representation $\phi' : G \to \mathcal{A} \otimes \mathcal{B}(\mathbb{C}^{|G|})$ such that*

$$\mathop{\mathbb{E}}_{g\in G} \|\phi(g) - V^*\phi'(g)V\|_\rho^2 \leq \varepsilon.$$

Proof. Define the isometry $V : \mathcal{H} \to \mathcal{H} \otimes \mathbb{C}^{|G|}$ as $V|\psi\rangle = \frac{1}{\sqrt{|G|}} \bigoplus_{g\in G} \left(\phi(g^{-1})|\psi\rangle\right)$, and note

$$V^*\left(\frac{1}{\sqrt{|G|}} \bigoplus_{g\in G} |\psi_g\rangle\right) = \frac{1}{|G|} \sum_{g\in G} \phi(g^{-1})^* |\psi_g\rangle.$$

We see that for all $|\psi\rangle \in \mathcal{H}$

$$V^*V|\psi\rangle = V^*\left(\frac{1}{\sqrt{|G|}} \bigoplus_{g\in G} \phi(g^{-1})|\psi\rangle\right) = \frac{1}{|G|} \sum_g \phi(g^{-1})^*\phi(g^{-1})|\psi\rangle = |\psi\rangle,$$

showing that $V^*V = \mathbb{1}_{\mathcal{H}}$ and hence V is an isometry. Furthermore for all $B \in \mathcal{A}'$

$$VB|\psi\rangle = \frac{1}{\sqrt{|G|}} \bigoplus_{g\in G} \left(\phi(g^{-1})B|\psi\rangle\right) = (B \otimes I_{|\mathcal{G}|})V|\psi\rangle, \tag{8.1}$$

implying $VB = (B \otimes I_{|\mathcal{G}|})V$. Define $\phi'(g) = \mathbb{1}_{\mathcal{H}} \otimes L_g$, where L_g is the left regular representation for the group G. Note for all $a \in G$

$$V^*\phi'(a)V = \frac{1}{\sqrt{|G|}} V^*(\mathbb{1}_{\mathcal{H}} \otimes L_a)\left(\bigoplus_{g\in G} \phi(g^{-1})\right) \quad = \frac{1}{\sqrt{|G|}} V^* \bigoplus_{ag\in G} \phi(g^{-1}) = \frac{1}{|G|} \sum_g \phi(ag)\phi^*(g).$$

Hence since ϕ is an $(\varepsilon, \rho)-$representation

$$\mathop{\mathbb{E}}_{g\in G} \|\phi(g) - V^*\phi'(g)V\|_\rho^2 = \mathop{\mathbb{E}}_{g\in G} \|\phi(g) - \sum_h \phi(gh)\phi^*(h)\|_\rho^2 \leq \mathop{\mathbb{E}}_{g,h\in G} \|\phi(g)\phi(h) - \phi(gh)\|_\rho^2 \leq \varepsilon$$

completing the claim. □

Note that since each irreducible representation ϕ for a finite group appears $\dim(\phi)$ times in the left regular representation, the proof above follows the same structure as [Vid18, Theorem 4]. We remark that the condition where $VB = (I_{|\mathcal{G}|} \otimes B)V$ for all $B \in \mathcal{A}'$ intuitively corresponds to V only acting on Alice's algebra in the context of non-local games. Using the example given in Section 4.1, the above theorem can be seen as [Vid18, Theorem 4] applied on the first register of $\mathbb{C}^n \otimes \mathbb{C}^n$.

8.1.2 Spectral gap

In this section, we recall the definition of the spectral gap for a probability distribution over a group and introduce a key lemma used to analyze the μ-dependent Pauli Basis test. Many of these definitions and theorems are from [dlS22b] and we refer to the original reference for a more comprehensive introduction to this topic. Let G be a finite group and μ be a probability measure on G; we use $\mathrm{supp}(\mu)$ to denote the support of the distribution, or $\mathrm{supp}(\mu) = \{g \in G | \mu(g) > 0\}$. We say μ is symmetric if $\mu(g^{-1}) = \mu(g)$ for all $g \in G$, and generating if $\mathrm{supp}(\mu)$

generates the group G. We define $\kappa(\mu)$ as the smallest number such that for all unitary representation π, we have

$$sp\left(\sum_g \mu(g)\pi(g)\right) \subseteq \left[-1, 1-\frac{1}{\kappa(\mu)}\right] \cup \{1\},$$

and we note that $\kappa(\mu) = \infty$ in certain cases. We note that $\kappa(\mu)$ is the inverse of the *spectral gap* in the literature and could potentially be infinite given the probability distribution. For convenience of notation, we let $\kappa(\mu) = \infty$ if the probability distribution is neither generating nor symmetric, and justify this in Section 8.3 in the context of the protocol. We recall the following theorem about $\kappa(\mu)$ and note that the below theorem is based on the well-known Poincaré inequality.

Theorem 8.1.2 ([dlS22b], Corollary 2.5)**.** *Let $q = p^n$ for some prime power p, and let $\omega_p = e^{\frac{2\pi i}{p}}$ be the root of unity of order p. Let μ and υ be two symmetric and generating probability measures on $\mathbb{F}_q$. Then, for any unitary representation $\pi_1 : \mathbb{Z}_q \to \mathscr{A}$ and $\pi_2 : \mathbb{Z}_q \to \mathscr{A}$, we have*

$$\mathop{\mathbb{E}}_{(g,h)\in\mathbb{Z}_q^2} ||\pi_1(g)\pi_2(h) - \omega_p^{gh}\pi_2(h)\pi_1(g)||_2^2 \leq \kappa(\mu)\kappa(\upsilon) \sum_{(g,h)\in\mathbb{F}_q^2} \mu(g)\upsilon(h)||\pi_1(g)\pi_2(h) - \pi_2(h)\pi_1(g)||_2^2$$

Note that the above theorem only applies when π_1 and π_2 are both *exact* unitary representations for the group $\mathbb{F}_q$. We also recall the following fact about the spectral gap of $\mathbb{F}_q^n$.

Theorem 8.1.3 ([AR94], Proposition 4.6)**.** *For prime p and $n \in \mathbb{N}$, there exists a subset $S \subseteq \mathbb{F}_p^n$ with $|S| = O(n)$ such that $\kappa(\mu) \leq 2$ for μ, the uniform distribution on S.*

In this thesis, we only consider the above theorem for $q = 2$, and recall from preliminary that we identify $\mathbb{F}_2$ with elements of $\{0, 1\}$.

8.2 Generalized Pauli measurements

We first introduce the notion of a generalized Pauli measurement. Let $p \in \mathbb{N}$ be an odd integer. For $W \in \{X, Z\}$, the generalized Pauli measurement over $\mathbb{F}_{2^p}$ are the sets of PVM $\left\{\rho_a^{W,p} = |a^{W,p}\rangle\langle a^{W,p}|\right\}_{a\in\mathbb{F}_{2^p}}$ where

$$|a^{X,p}\rangle := \frac{1}{\sqrt{2^p}} \sum_{b\in\mathbb{F}_{2^p}} (-1)^{\mathrm{Tr}(ab)} |b\rangle, \qquad |a^{Z,p}\rangle := |a\rangle,$$

for all $a \in \mathbb{F}_{2^p}$. In the case where $p = 1$, the generalized Pauli measurements corresponds to the eigenspace of the Pauli X and Z matrices.

In this thesis, we often associate PVMs with binary observables to better analyze the commutation properties of these measurements. A binary observable is a unitary matrix which squares to the identity. For an odd integer $p \in \mathbb{N}$, $W \in \{X, Z\}$ and $a \in \mathbb{F}_{2^p}$, we define the generalized Pauli matrices $\rho^{W,p}(a)$ as

$$\rho^{X,p}(a) := \sum_{b\in\mathbb{F}_{2^p}} |b + a\rangle\langle b|, \qquad \rho^{Z,p}(a) := \sum_{b\in\mathbb{F}_{2^p}} (-1)^{\mathrm{Tr}(a\cdot b)} |b\rangle\langle b|$$

where the addition and multiplication above are over $\mathbb{F}_{2^p}$. In the case where $p = 1$, we drop the superscript and simply write ρ_i^W to denote the qubit Pauli W-measurement for $i \in \{0, 1\}$ and ρ^W for the Pauli W-matrix. We see that for $W \in \{X, Z\}$ the eigenspace for $\rho^{W,p}(a)$ is precisely the PVM measurement for the generalized Pauli measurement, and we can write

$$\rho^{W,p}(a) := \sum_{b\in\mathbb{F}_{2^p}} (-1)^{\mathrm{Tr}(ab)} \rho_a^{W,p}. \tag{8.2}$$

This also implies that $\rho^{W,p}(a)$ commutes with $\rho^{W,p}(b)$ for any $a, b \in \mathbb{F}_{2^p}$. As shown in [JNV+22a, equation 19], the generalized Pauli measurements can also be written as

$$\rho_a^{W,p} = \mathop{\mathbb{E}}_{b\in\mathbb{F}_{2^p}} (-1)^{-\mathrm{Tr}(ab)} \rho^{W,p}(b). \tag{8.3}$$

By a simple calculation, we see that for all $W \in \{X, Z\}$ and $a, b \in \mathbb{F}_{2^p}$, the generalized Pauli observables obey the following relationships

$$\rho^{W,p}(a) \cdot \rho^{W,p}(b) = \rho^{W,p}(a+b). \tag{8.4}$$

The generalized Pauli observables also follow the "twisted commutation" relations, whereby

$$\rho^{X,p}(a) \cdot \rho^{Z,p}(b) = (-1)^{\mathrm{Tr}(ab)} \rho^{Z,p}(b) \cdot \rho^{X,p}(a). \tag{8.5}$$

For $W \in \{X, Z\}$ and $s \in \mathbb{F}_{2^p}^m$, we define

$$\rho^{W,p}(s) = \bigotimes_{i\in[m]} \rho^{W,p}(s_i) \qquad \text{and} \qquad \rho_s^{W,p} = \bigotimes_{i\in[m]} \rho_{s_i}^{W,p} \tag{8.6}$$

where each $s_i \in \mathbb{F}_{2^p}$. We recall the following lemma which shows the existence of a unitary which converts between the generalized Pauli measurement to the one qubit Pauli measurement.

Lemma 8.2.1 (Lemma 3.26 of [JNV+22a])**.** *Let $p, m \in \mathbb{N}$ where p is an odd integer, there exists a unitary $U_{2\to p} : (\mathbb{C}^2)^{\otimes p\cdot m} \to (\mathbb{C}^{2^p})^{\otimes m}$ such that for all $W \in \{X, Z\}$ and for all $s \in \mathbb{F}_{2^p}^{\otimes n}$, we have*

$$\rho_s^{p,W} = U_{2\to p}^* \left(\bigotimes_{i=0}^{m} \bigotimes_{j=0}^{p} \rho^W_{\kappa(s_i)_j} \right) U_{2\to p}$$
$$\left(U_{2\to p}^* \otimes U_{2\to p}^*\right) |ME_{2^p}\rangle^{\otimes m} = |ME_2\rangle^{\otimes p\cdot m}.$$

To abuse notation, for register subspace $V \subseteq \mathbb{F}_2^m$ and $W \in \{X, Z\}$, we use $\{\rho_s^X\}_{s\in V}$ to denote the measurement

$$\rho_s^W := \sum_{a|a_V=s} \rho_s^W$$

where for $a \in \mathbb{F}_2^m$, $a = a_V + a_V^C$ is the unique decomposition such that $a_V \in V$ and $a_V^C \in V^C$. For register subspace $V \subseteq \mathbb{F}_{2^p}^m$, we use $\{\rho_s^W\}_{s\in V}$ to denote the measurement $\{\rho_s^V\}_{s\in\kappa(W)}$. We recall the following lemma from [JNV+22a].

Lemma 8.2.2 (Lemma 8.5 of [JNV+22a])**.** *Let $\mathrm{L}_1, \mathrm{L}_2 : \mathbb{F}_{2^p}^m \to \mathbb{F}_{2^p}^m$ be two linear map. Then*

$$\ker(\mathrm{L}_2)^\perp \subseteq \ker(\mathrm{L}_1),$$

implies that the measurement operators $\left\{\rho^{Z,p}_{[\mathrm{L}_1|s]}\right\}_{s\in\mathbb{F}_{2^p}^m}$ and $\left\{\rho^{X,p}_{[\mathrm{L}_2|s]}\right\}_{s\in\mathbb{F}_{2^p}^m}$ pairwise commute, as well as the measurement operators $\left\{\rho^{Z,p}_{[\mathrm{L}_2|s]}\right\}_{s\in\mathbb{F}_{2^p}^m}$ and $\left\{\rho^{X,p}_{[\mathrm{L}_1|s]}\right\}_{s\in\mathbb{F}_{2^p}^m}$ pairwise commute

8.3 The Pauli Basis test

8.3.1 The magic square game

We first introduce a key subroutine for the Pauli Basis test, the Mermin-Peres magic square game [Mer90; Per90], in this subsection. We use the BCS formulation of this game as presented in [CM14], where the game is defined by six equations and nine variables over $\mathbb{F}_2$, where the

x_1	x_2	x_3
x_4	x_5	x_6
x_7	x_8	x_9

$\mathbb{I}_2 \otimes \rho^Z$	$\rho^Z \otimes \mathbb{I}_2$	$\rho^Z \otimes \rho^Z$
$\rho^X \otimes \mathbb{I}_2$	$\mathbb{I}_2 \otimes \rho^Z$	$\rho^X \otimes \rho^X$
$\rho^X \otimes \rho^Z$	$\rho^Z \otimes \rho^X$	$-\left(\rho^X \rho^Z\right) \otimes \left(\rho^X \rho^Z\right)$

Figure 8.1: Left: The description for the magic square game, where each row and column corresponds to an equation. Right: A oracularizable perfect strategy for the magic square game.

variables are arranged on a three-by-three grid as presented in Figure 8.1. Every row and column in Figure 8.1 corresponds to a constraint that multiplies to 1, except for the last column, where the constraint multiplies to -1 instead. In this game, the referee randomly samples a constraint and a variable in the constraint and sends the constraint to one of the provers and the variable to the other prover. The prover must then respond with an assignment for their given constraint or variable. The provers win the game if their assignments are consistent with each other. If one of the provers is given an equation as the question, their assignment must also satisfy the constraint for the given equation. We also modify the magic square to be synchronous, meaning the verifier additionally samples a constraint or a variable with constant probability and sends it to both provers and expects the same answer in return.

The magic square game admits a perfect synchronous oracularizable strategy for both quantum models by using the measurement operator defined in Figure 8.1. In this strategy, the provers initially prepare two copies $|\mathrm{ME}_2\rangle$. In the event that the prover receives a variable, he measures the observable as displayed in the grid and returns the resulting eigenvalue (which is either 1 or -1) as the assignment to the variable. If the prover receives an equation, he measures the observables for all three variables, returning the eigenvalue for each of the observables as the corresponding assignment for each of the variables. Since each observable on the grid commutes with all other observables that share a row or column with it, the order of measurement does not matter for the constraint question, and their measurement commutes on all possible question pairs, which implies that this perfect strategy is oracularizable.

Combining [MNY22], Theorem 3.1 part iii) (rigidity for the magic square game under synchronous correlations) and Corollary 8.4.1, we recover the same anti-commutation result for the commuting operator model.

Theorem 8.3.1 (Approximate anti-commutation test for the magic square game in the commuting operator model)**.** *Let $\mathscr{S} = (\mathscr{L}^2(\mathscr{A}, \tau), \sigma|\tau\rangle, \{A_a^x\}, \{B_a^x\})$ be a tracially embeddable strategy for the magic square game such that $\omega^{co}(\text{Magic square}, \mathscr{S}) \geq 1 - \varepsilon$.*

For $i \in \{1 \cdots 9\}$, let $\{A_0^{(\text{variable, } i)}, A_1^{(\text{variable, } i)}\} \subseteq \mathscr{A}$ be Alice's measurement operator for variable i, and likewise with $\{B_0^{\text{variable, } i}, B_1^{\text{variable, } i}\} \subseteq \mathscr{A}'$ with Bob's measurement operator. For $P \in \{A, B\}$

$$P^{(\text{variable, } i)} = P_0^{(\text{variable, } i)} - P_1^{(\text{variable, } i)}.$$

Then

$$||(P^{(\text{variable, } 1)} P^{(\text{variable, } 5)} + P^{(\text{variable, } 5)} P^{(\text{variable, } 1)})||_{\sigma^2}^2 \leq O(\text{poly}(\delta)).$$

8.3.2 The Pauli Basis test

In this section, we give the formal definition for the μ-dependent Pauli Basis test. We describe the sampling and decision procedure in Figure 8.3 and include a diagram for the sampling procedure in Figure 8.2 for clarity. Intuitively, the goal of the Pauli Basis test is to force two honest provers to prepare n copies of EPR pairs between them and measure either $(\rho^X)^{\otimes n}$ or $(\rho^Z)^{\otimes n}$ on their half of the EPR pair.

First, we will give an informal overview of the test. As mentioned at the beginning of this section, the goal of this test is to allow two players to certify that they share a state of the

form $|EPR\rangle^{\otimes n}$. For $W \in \{X,Z\}$, once the players receive the question label (Pauli, W), they are expected to perform the measurement $\{\rho_W^{\otimes n}\}$ on their half of their respective states. The referee achieves this by first sampling some u from the distribution μ and sending the label (Coordinates, W) along with the coordinates u, and expects the player receiving this label to measure according to $\rho_W(u)$. In other words, the referee expects the player to perform a W measurement on the coordinate $i \in [n]$ if $u_i = 1$. The other player will receive (Pauli, W). Since the entangled state expected from the players is EPR pairs, the player returns the same measurement output on the sub-coordinates where they both made the measurement.

The commutation/anti-commutation test is used to enforce consistency on the question label (Coordinates, W). To be more precise, the referee first samples two coordinates (u,v) and determines whether the observable $\rho_X(u)$ and $\rho_Z(v)$ commute or anti-commute based on the value of $u \cdot v \bmod 2$. In the scenario where they do commute, i.e. $u \cdot v \bmod 2 = 0$, the "commutation test" or one of the red edges of Figure 8.3 is performed. In this case, the referee asks one of the players to simultaneously measure both the $\rho_X(u)$ and $\rho_Z(v)$ (which would be consistent since they commute) and compare the measurement output to the other player performing one of the two measurements. In the case where $u \cdot v \bmod 2 = 1$, the "anti-commutation test", or one of the blue edges of Figure 8.3 is performed, where the referee forces both players to play the magic square game from Section 8.3 with the observables $\rho_X(u)$ and $\rho_Z(v)$ for questions 1 and 5 of the magic square game respectively. In this case, Theorem 8.3.1 guarantees that any perfect or near-perfect strategy ensures that the two aforementioned observables anti-commute with each other.

In order to state the main rigidity theorem, let $\mathscr{S} = (\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{A_a^x\}, \{B_a^x\})$ be a projective, tracially embeddable strategy for the μ-dependent Pauli Basis test. For $W \in \{X,Z\}$, and $P \in \{A,B\}$ let $\{P_a^{(\text{Pauli},W)}\}$ be the measurement operator for the question label (Pauli, W) for player P and define $W^P(u)$ as the following unitary observable:

$$W^P(u) = \sum_{q \in \{0,1\}^n} (-1)^{q \cdot u} P_q^{(\text{Pauli},W)}, \tag{8.7}$$

where, by definition, $W^A(u) \subseteq \mathscr{A}$ and $W^B \subseteq \mathscr{A}'$. We state the following rigidity theorem about the μ-dependent Pauli Basis test.

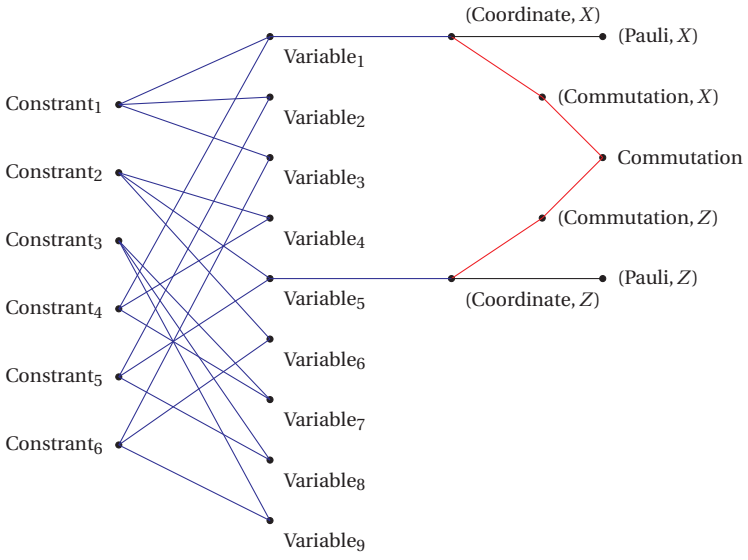

Figure 8.2: The typed graph $(\mathrm{T}^{\mathrm{PB}}, \mathrm{E}^{\mathrm{PB}})$ for the μ-dependent Pauli Basis test, where each of the vertices above also contains a self-loop (this is a black edge). Similar to the definition of the CL distribution, each vertices in the graph represents a potential question label and each edge represents a potential question pair. Each of the edges are colour coded in order to better explain the game procedure and we refer to Figure 8.3 for more details.

Question label	Question content	Answer format
(Pauli,W)		$t_W \in \{0,1\}^n$
(Coordinate, W)	$u_W \in \mu$	$t_{(\text{Pauli, W})} \in \{0,1\}^n$
(Commutation, W)	$u_W \in \mu$	$t_W \in \{0,1\}$
Commutation	$(u_X, u_Z) \in \mu \times \mu$	$(t'_X, t'_Z) \in \{0,1\}$
Variable$_i$	$(u_X, u_Z) \in \mu \times \mu$	$t_{\text{var}} \in \{0,1\}$
Constraint$_i$	$(u_X, u_Z) \in \mu \times \mu$	$t_{\text{cons}} \in \{0,1\}^3$

Figure: Q and A format for the n-qubit Pauli Basis test, where $W \in \{X, Z\}$.

Sampling procedure

1. Sample $(u_X, u_Z) \in \mu \times \mu$ uniformly at random.
2. Uniformly samples $(n_0, n_1) \in \mathrm{T}^{\mathrm{PB}} \times \mathrm{T}^{\mathrm{Paulibasis}}$, where $(\mathrm{T}^{\mathrm{PB}}, \mathrm{E}^{\mathrm{PB}})$ is the graph in Figure 8.2, and perform rejection sampling until $(n_0, n_1) \in \mathrm{E}^{\mathrm{PB}}$.
3. Send the question label and question content corresponding to n_0 to one of the provers, and send the question content corresponding to the n_1 to the other prover.

Verification procedure

- (Self-loop): The provers win iff they output the same answer.
- (Pauli, W) – (Coordinate, W): Alice and Bob win iff $t_W|_{u_W} = t_{(\text{Pauli, W})}|_{u_W}$.
- If (n_0, n_1) are a red edge and $u_x \cdot u_z = 1$, the provers win if for the question label (Commutation, W) and (Commutation), the prover answers 0, otherwise
  - (Coordinate, W) – (Commutation, W): The provers win iff $u_W \cdot t_{(\text{Pauli},W)} = t_W$.
  - (Commutation) – (Commutation, W): The provers win iff $t_W = t'_W$.
- If (n_0, n_1) are a blue edge and $u_x \cdot u_z = 0$, the provers wins if for the question label (Constraint, i) and (Variable, j), the prover answers 0, otherwise:
  - (Variable 1) – (Coordinate, X): The provers iff $u_X \cdot t_X = t_{\text{var}}$.
  - (Variable 5) – (Coordinate, Z): The provers iff $u_Z \cdot t_Z = t_{\text{var}}$.
  - (Constraint) – (Variable): The provers win iff t_{cos} is consistent with the constraint in the magic square game, and t_{var} is consistent with the assignment of v_i within t_{cos}.

Figure 8.3: The description for the n-qubit Pauli Basis test. Where $W \in \{X, Z\}$ in the decision procedure.

Theorem 8.3.2 (Robustness of the μ-dependent Pauli Basis test under the commuting operator model)**.** *Let $\mathcal{G}(\mu)$ be the μ-dependent Pauli Basis test and let $\mathcal{S} = (\mathcal{L}^2(\mathcal{A},\tau),\sigma|\tau\rangle,\{A_a^x\},\{B_a^x\})$ be a tracially embeddable strategy for $\mathcal{G}(\mu)$ with $\omega^{co}(\mathcal{G}(\mu),\mathcal{S}) \geq 1-\varepsilon$. There exist two isometries $V_A : \mathcal{H} \to \mathcal{H}\otimes\mathbb{C}^{2^{2n}}$ and $V_B : \mathcal{H} \to \mathcal{H}\otimes\mathbb{C}^{2^{2n}}$ with $(V_B\otimes\mathbb{I}_{2^{2n}})V_A = (V_A\otimes\mathbb{I}_{2^{2n}})V_B$, and a state $|Aux\rangle \in \mathcal{H}\otimes\mathbb{I}_{2^{2n}}$ such that*

$$\left\|\left(V_B\otimes\mathbb{I}_{2^{2n}}\right)V_A|\psi\rangle - |Aux\rangle|EPR\rangle^{\otimes n}\right\|^2 \leq O\left(\text{poly}(\kappa(\mu),\varepsilon)\right), \tag{8.8}$$

and for all $W \in \{X,Z\}$ and $u \in \{0,1\}^n$

$$\left|\left(V_A W^A(u) V_A^*\otimes\mathbb{I}_{2^{2n}} - \mathbb{I}_{\mathcal{H}}\otimes\mathbb{I}_{2^{2n}}\otimes\rho^W(u)\otimes\mathbb{I}_{2^n}\right)|Aux\rangle|EPR\rangle^{\otimes n}\right|^2 \leq O\left(\text{poly}(\kappa(\mu),\varepsilon)\right), \tag{8.9}$$

$$\left|\left(V_B W^B(u) V_B^*\otimes\mathbb{I}_{2^{2n}} - \mathbb{I}_{\mathcal{H}}\otimes\mathbb{I}_{2^{2n}}\otimes\mathbb{I}_{2^n}\otimes\rho^W(u)\right)|Aux\rangle|EPR\rangle^{\otimes n}\right|^2 \leq O\left(\text{poly}(\kappa(\mu),\varepsilon)\right). \tag{8.10}$$

We provide a proof of Theorem 8.3.2 in Section 8.4. Intuitively, if μ is a uniform distribution on some subset $S \subseteq \{0,1\}^n$, the larger S is, the harder it is for dishonest players to cheat, but this also gives a larger question set (or sampling complexity) for the μ-dependent Pauli Basis test. This naturally gives a trade-off between the sampling complexity and the accuracy of the test. Hence, $\kappa(\mu)$ can intuitively be interpreted as "how good the distribution is" for the theorem above. We also remark that, if μ is not generating, then this implies that there exists some generator $i \in [n]$ such that $s_i = 0$ for all $s \in \text{supp}(\mu)$. Translating this to the context of the μ-dependent Pauli Basis test would imply that there exists an EPR pair which would not get cross-checked in any of the subsets sampled under μ; hence, the robustness theorem should not hold under this instance. Hence, this justifies the assumption we made for $\kappa(\mu) = \infty$ with a distribution that is neither symmetric nor generating earlier in Section 8.1.2. We further remark that the condition $(V_B\otimes\mathbb{I}_{2^{2n}})V_A = (V_A\otimes\mathbb{I}_{2^{2n}})V_B$ is akin to V_A and V_B acting on different Hilbert spaces under the tensor product model, and the extra $\mathbb{C}^{2^n}$ on the auxiliary space results from using the left regular representation of the Weyl-Heisenberg group.

Finally, we remark that the μ-dependent Pauli Basis test given in this thesis is slightly different from the one given in [Lin24, Section 6], namely there are two extra types (the commutation questions in Figure 8.2). This is because in order to use this protocol as a part of the compression theorem, the perfect strategy that succeeds the game needs to have a perfect **oracularizable** strategy, and the extra question test is a "hack" to get around that. This is the same reason why the simplification given in [dlS23] cannot instead be used as part of the question reduction protocol.

8.3.3 The n-qubit Pauli Basis test, a question efficient EPR-tester

In order to proof Proposition 7.4.2, one needs to, in addition, make sure that the verifier can sample the question distribution in $O(\log^{O(\text{poly}(\alpha))}(n))$ time. To this end, recall that a linear $[n,c,d]_{\mathbb{F}_q}$ code is a set $\mathfrak{C}$ of functions $g : [p] \to \mathbb{F}_q$ with size $|\mathfrak{C}| = q^c$ that is closed under linear combination, such that for any two distinct $g \neq g'$, the number of coordinates $i \in [n]$ such that $g(i) \neq g'(i)$ is at least d. In [dlS22b, Example 1.2], given any $[k,n,d]$ binary linear code C, the code space $S_C \subset \mathbb{F}_2^n$ with $|S_C| = \{0,1\}^k$ is a subset such that the uniform distribution of S_C has a spectral gap of $\frac{2k}{d}$. Intuitively, any good binary linear code would lead to an efficient EPR tester. For any $n \in \mathbb{N}$, consider the Justesen code [Jus72] with $R = \frac{\log(n)}{n}$, by definition this is a code with dimension $k = \lfloor\log(n)\rfloor$, length n and distance $d \geq 0.11(n-\log(n))$. Let $S_n^{\text{PB}} \subseteq \{0,1\}^n = \mathbb{F}_2^n$ be the code space of this code. Since the encoding map for the Justesen code can be implemented in $O(\text{poly}(n))$ time, there exists an encoding map $\pi^{\text{PB}} : \mathbb{N}\times\{0,1\}^* \to \{0,1\}^*$ such that $\pi^{\text{PB}}(n,s)$ is a bijection map that maps elements from $\{0,1\}^{\lfloor\log(n)\rfloor}$ to S_n^{PB}, with $\mathsf{TIME}_{\pi^{\text{PB}}}(n) = \text{poly}(n)$. Furthermore, the uniform distribution on S_n^{PB} has a spectral gap of $\frac{2k}{d} \leq \frac{\log(n)}{0.11(n-\log(n))} \leq 30$ for all n.

For $n \in \mathbb{N}$, we define the *n-qubit Pauli Basis test* to be μ dependent-Pauli Basis test with the distribution μ being the uniform distribution over S_n^{PB} and we use $\mathcal{G}_n^{\text{PB}}$ to denote this game.

Combining the property for the Justesen code discussed above with Theorem 8.3.2, we prove following properties about the n-qubit Pauli Basis test.

Theorem 8.3.3 (Properties of the n-qubit Pauli Basis test)**.** *Let $\mathcal{G}_n^{PB}$ be the n-qubit Pauli Basis test*

1. *(Computation time): $\mathcal{G}_n^{PB}$ is samplable via a $(\mathtt{T}^{PB}, \mathtt{E}^{PB}, \{\mathtt{L}^{v}\}_{v\in\mathtt{T}^{PB}})$ typed CL distribution, where each $\mathtt{T}^{PB}: \mathbb{F}_2^{2\cdot s_n^{SB}} \to \mathbb{F}_2^{2\cdot s_n^{SB}}$ is a first level CL function, where $s_n^{SB} = \lfloor \log(n) \rfloor$. $\mathcal{G}_n^{PB}$ has a decision complexity of* $\mathrm{poly}(n)$.

2. *(Completeness): There exists a perfect finite-dimensional symmetric oracularizable strategy $\mathscr{S}^{PB} = (\mathbb{C}^{2^n} \otimes \mathbb{C}^{2^n}, |ME_2\rangle^{\otimes n} \otimes |ME_2\rangle, \{P_a^x\})$ such that for all $s \in \{0,1\}^n$ and $W \in \{X,Z\}$*

$$P_s^{(Pauli,W)} = \rho_s^W \otimes \mathbb{I}_2.$$

3. *(Soundness): Let $\mathscr{S} = (\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{P_a^x\}_{x\in\mathscr{X}})$ be a projective, tracially embeddable strategy such that $\omega(\mathcal{G}_n^{PB}, \mathscr{S}) \geq 1-\varepsilon$. Then there exist two isometries $V_A : \mathscr{L}^2(\mathscr{A},\tau) \to \mathscr{L}^2(\mathscr{A},\tau) \otimes \mathbb{C}^{2^{2n}}$ and $V_B : \mathscr{L}^2(\mathscr{A},\tau) \to \mathscr{L}^2(\mathscr{A},\tau) \otimes \mathbb{C}^{2^{2n}}$ with $(V_B \otimes \mathbb{I}_{2^{2n}})V_A = V_A(V_B \otimes \mathbb{I}_{2^{2n}})$ and a state $|Aux\rangle \in \mathscr{L}^2(\mathscr{A},\tau) \otimes \mathbb{C}^{2^{2n}}$ such that*

$$\left\| \left(V_B \otimes \mathbb{I}_{2^{2n}}\right) V_A(\sigma|\tau\rangle) - |Aux\rangle\,|ME_2\rangle^{\otimes n} \right\|^2 \leq O\left(\mathrm{poly}(\varepsilon)\right),$$

and for all $W \in \{X,Z\}$ and $u \in \mathbb{F}_{2^n}$

$$\begin{aligned}
&|((V_A P_s^{(Pauli,W)} V_A^*)_{\mathscr{A}A_1A_2} \otimes (\mathbb{I}_{2^{2n}})_{B_1B_2} - \\
&\quad (\mathbb{I}_{\mathscr{H}})_{\mathscr{A}} \otimes (\rho_s^W)_{A_1} \otimes (\mathbb{I}_{2^n})_{A_2} \otimes (\mathbb{I}_{2^{2n}})_{B_1B_2})\,|Aux\rangle_{\mathscr{A}A_2B_2}\,|ME_2\rangle_{A_1B_1}^{\otimes n}|^2 \leq O\left(\mathrm{poly}(\varepsilon)\right). \\
&|((V_B (P_s^{(Pauli,W)})^{op} V_B^*)_{\mathscr{A}B_1B_2} \otimes (\mathbb{I}_{2^{2n}})_{A_1A_2} - \\
&\quad (\mathbb{I}_{\mathscr{H}})_{\mathscr{A}} \otimes (\rho_s^W)_{B_1} \otimes (\mathbb{I}_{2^n})_{B_2} \otimes (\mathbb{I}_{2^{2n}})_{A_1A_2})\,|Aux\rangle_{\mathscr{A}A_2B_2}\,|ME_2\rangle_{A_1B_1}^{\otimes n}|^2 \leq O\left(\mathrm{poly}(\varepsilon)\right).
\end{aligned}$$

where the subscript $\mathscr{A}$, A_1, B_1, A_2, B_2 and $\mathscr{A}$ denotes the registers on which each operator acts (listed for clarity).

Proof. For the remainder of this proof, let $W \in \{X,Z\}$. Fix an integer $n \in \mathbb{N}$. We first show that the $\mathcal{G}_n^{\mathrm{PB}}$ is samplable via a typed CL distribution. By the definition given in Figure 8.2, one can define the sampling procedure for $\mathcal{G}_n^{\mathrm{PB}}$ as the following:

1. Sample $(u_X, u_Z) \in S_n^{\mathrm{PB}} \times S_n^{\mathrm{PB}}$ uniformly at random.
2. Uniformly samples $(n_0, n_1) \in \mathtt{T}^{\mathrm{PB}} \times \mathtt{T}^{\mathrm{Paulibasis}}$, where $(\mathtt{T}^{\mathrm{PB}}, \mathtt{E}^{\mathrm{PB}})$ is the graph in Figure 8.2, and perform rejection sampling until $(n_0, n_1) \in \mathtt{E}^{\mathrm{PB}}$.
3. Send the question label and question content corresponding to n_0 to one of the provers, and send the question content corresponding to the n_1 to the other prover.

We wish to show that the above sampling procedure can be done via a typed CL samplable function. Identify $\mathbb{F}_2$ with $\{0,1\}$, and define each $\mathtt{L}^v$ as follows: For every input $(s_X, s_Z) \in \{0,1\}^{s_n^{\mathrm{SB}}} \times \{0,1\}^{s_n^{\mathrm{SB}}}$, we define each of the linear function accordingly:

- For $v \in \{(\mathrm{Pauli}, W), (\mathrm{Coordinate}, W), (\mathrm{Commutation}, W)\}$,

$$\begin{aligned}
\mathtt{L}^{(\mathrm{Pauli},W)}(s_X, s_Z) &= (0,0), \\
\mathtt{L}^{(\mathrm{Coordinate},X)}(s_X, s_Z) = \mathtt{L}^{(\mathrm{Commutation},X)}(s_X, s_Z) &= (s_X, 0), \\
\mathtt{L}^{(\mathrm{Coordinate},Z)}(s_X, s_Z) = \mathtt{L}^{(\mathrm{Commutation},Z)}(s_X, s_Z) &= (0, s_Z).
\end{aligned}$$

- Otherwise, $\mathtt{L}^v$ is the identity function, or

$$\mathtt{L}^v(s_X, s_Z) = (s_X, s_Z).$$

For the decision process, given input $(u_X, u_Z) \in S_n^{\mathrm{PB}} \times S_n^{\mathrm{PB}}$ and the output listed in the "Answer format" from Figure 8.3, the decision process for all the edges can be decided in $O(\mathrm{poly}(n))$ time since it involves either computing an inner product between elements of $\{0,1\}^n$ or some form of consistency test. Furthermore, since the map π^{PB} is computable uniformly in $O(\mathrm{poly}(n))$ time, the decider can compute each of the (u_X, u_Z) from (s_X, s_Z) sampled above. For the "completeness" condition, we define the synchronous strategy $\mathscr{S}^{\mathrm{PB}}$ over the Hilbert space $\mathbb{C}^{2^n} \otimes \mathbb{C}^{2^n}$ as follows: The joint state used between the two provers is n EPR pairs, or $|\mathrm{ME}_2\rangle^{\otimes n} \otimes |\mathrm{ME}_2\rangle$. For $(u_X, u_Z) \in S_n^{\mathrm{PB}} \times S_n^{\mathrm{PB}}$, define

$$\rho^W(u_W)_0 = \sum_{b \cdot u_W = 0} \rho_b^W, \qquad \rho^W(u_W)_0 = \sum_{b \cdot u_W = 1} \rho_b^W,$$

and we see that

$$\rho^W(u_W) = \rho^W(u_W)_0 - \rho^W(u_W)_1.$$

We define the measurement operator $\{P_a^x\}$ for the symmetric strategy $\mathscr{S}^{\mathrm{PB}}$ as

$$\begin{aligned}
&P^{(\mathrm{Pauli},W)} = \rho_{t_W}^W \otimes \mathbb{I}_2 && \text{for all } t_W \in \{0,1\}^n,\\
&P^{(\mathrm{Coordinate},W),u_W}_{t_{(\mathrm{Pauli,\,W})}} = \sum_{t_{(\mathrm{Pauli,\,W})}|_{u_W} = b} \rho_b^W \otimes \mathbb{I}_2 && \text{for all } t_{(\mathrm{Pauli,\,W})} \in \{0,1\}^n.\\
&P^{(\mathrm{Commutation,\,W}),(u_W)}_{t_W} = \rho^W(u_W)_{t_W} \otimes \mathbb{I}_2 && \text{for } u_X \cdot u_Z = 0,\ t_W \in \{0,1\}\\
&P^{(\mathrm{Commutation}),(u_X,u_Z)}_{(t'_W,t'_Z)} = \left(\rho^X(u_X)_{t'_X}\right)\left(\rho^Z(u_Z)_{t'_Z}\right) \otimes \mathbb{I}_2 && \text{for } u_X \cdot u_Z = 0,\ t'_W \in \{0,1\}\\
&P_a^{(\mathrm{Commutation}),(u_X,u_Z)} = \begin{cases} \mathbb{I}_{2^{n+1}} & \text{if } a = 0 \\ 0 & \text{otherwise} \end{cases} && \text{for } u_X \cdot u_Z = 1.\\
&P_a^{(\mathrm{Commutation,\,W}),(u_W)} = \begin{cases} \mathbb{I}_{2^{n+1}} & \text{if } a = 0 \\ 0 & \text{otherwise} \end{cases} && \text{for } u_X \cdot u_Z = 1.
\end{aligned}$$

For the magic square game, given (u_X, u_Z) such that $u_x \cdot u_z = 1$, by (8.5), the observables $\rho^X(u_X)$ and $\rho^Z(u_Z)$ anti-commutes. By applying [JNV+22a, Theorem 7.11], there exists a symmetric projective oracularizable strategy $\mathscr{S}^{\mathrm{MS}} = (\mathbb{C}^{2^{n+1}} \otimes \mathbb{C}^{2^{n+1}}, |\mathrm{ME}_2\rangle^{\otimes n} \otimes |\mathrm{ME}_2\rangle, \{M_a^x\})$ such that for $b \in \{0,1\}$,

$$M_b^{(\mathrm{Variable},1)} = \rho^W(u_W)_n \otimes \mathbb{I}_2, \qquad M_b^{(\mathrm{Variable},5)} = \rho^W(u_W)_n \otimes \mathbb{I}_2.$$

For $i \in \{2,3,4,6,7,8,9\}$, and $j \in \{1, \cdots, 6\}$,set

$$P_b^{(\mathrm{Variable},i),(u_x,u_z)} = M_b^{(\mathrm{Variable},i)}, \qquad P_b^{(\mathrm{Constraint},j),(u_x,u_z)} = M_b^{(\mathrm{Constraint},j)}.$$

In the event that $u_x \cdot u_z = 0$, set

$$P_b^{(\mathrm{Variable},i),(u_x,u_z)} = P_b^{(\mathrm{Constraint},j),(u_x,u_z)} = \begin{cases} \mathbb{I}_{2^{n+1}} & \text{if } b = 0 \\ 0 & \text{otherwise} \end{cases}.$$

Since all measurement operators $\mathscr{S}^{\mathrm{PB}}$ are defined within $\bigotimes_{i=0}^{n+1} \mathscr{M}_2(\mathbb{C})$, for all $a \in \mathscr{A}$

$$(P_a \otimes \mathbb{I}_{\mathbf{B}})\, |\mathrm{ME}_2\rangle_{\mathbf{AB}}^{\otimes n} \otimes |\mathrm{ME}_2\rangle_{\mathbf{AB}} = \left(\mathbb{I} \otimes P_a^T\right) |\mathrm{ME}_2\rangle_{\mathbf{AB}}^{\otimes n} \otimes |\mathrm{ME}_2\rangle_{\mathbf{AB}}$$

for all measurements within $\mathscr{S}^{\mathrm{PB}}$. We now verify that $\mathscr{S}^{\mathrm{PB}}$ is indeed an oracularizable and perfect strategy by considering all possible question pairs in $\mathscr{G}_n^{\mathrm{PB}}$.

- (Self-loop) Since all the measurements are projective, this is trivially true.
- (Pauli, W) – (Coordinate, W) This follows since both question labels require a measurement in the W basis.
- If $u_x \cdot u_z = 0$
  - The red-edge question pairs are trivially perfect/oracularizable, since one of the measurement operators is always the identity.
  - (Coordinate, W) – (Commutation, W): This follows since both question labels require a measurement in the W basis.
  - (Commutation) – (Commutation, W): By (8.5), $\rho^X(u_X)$ commutes with $\rho^Z(u_Z)$. Hence, $\{\rho^X(u_X)_i\}_{i\in\{0,1\}}$ pairwise commute with $\{\rho^Z(u_Z)_i\}_{i\in\{0,1\}}$, and the statement follows accordingly.
- If $u_x \cdot u_z = 1$
  - The blue edge question pairs are trivially perfect/oracularizable, since one of the measurement operators is always the identity.
  - (Variable 1) – (Coordinate, X): This follows since both question labels require a measurement in the X basis.
  - (Variable 5) – (Coordinate, Z): This follows since both question labels require a measurement in the Z basis.
  - (Constraint) – (Variable): This follows by [JNV+22a, Theorem 7.11].

This shows that $\mathscr{S}^{\mathrm{PB}}$ is a symmetric oracularizable strategy.

For soundness, let $\mathscr{S} = (\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{P_a^x\}_{x\in\mathscr{X}})$ be a projective, tracially embeddable strategy such that $\omega(\mathscr{G}_n^{\mathrm{PB}}, \mathscr{S}) \geq 1-\varepsilon$. By Theorem 8.3.3 and the spectral gap property for the Justesen code, we have

$$\left\| \left(V_B \otimes \mathbb{I}_{2^{2n}}\right) V_A |\psi\rangle - |\mathrm{Aux}\rangle |\mathrm{EPR}\rangle^{\otimes n} \right\|^2 \leq O\left(\mathrm{poly}(\varepsilon)\right),$$

and for all $W \in \{X,Z\}$ and $u \in \{0,1\}^n$

$$\left| \left(V_A W^A(u) V_A^* \otimes \mathbb{I}_{2^{2n}} - \mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \rho^W(u) \otimes \mathbb{I}_{2^n}\right) |\mathrm{Aux}\rangle |\mathrm{EPR}\rangle^{\otimes n} \right|^2 \leq O\left(\mathrm{poly}(\varepsilon)\right), \tag{8.11}$$

$$\left| \left(V_B W^B(u) V_B^* \otimes \mathbb{I}_{2^{2n}} - \mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \mathbb{I}_{2^n} \otimes \rho^W(u)\right) |\mathrm{Aux}\rangle |\mathrm{EPR}\rangle^{\otimes n} \right|^2 \leq O\left(\mathrm{poly}(\varepsilon)\right). \tag{8.12}$$

By rewriting (8.11), we see that

$$\begin{aligned}
&\left| \left(V_A W^A(u) V_A^* \otimes \mathbb{I}_{2^{2n}} - \mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \rho^W(u) \otimes \mathbb{I}_{2^n}\right) |\mathrm{Aux}\rangle |\mathrm{EPR}\rangle^{\otimes n} \right|^2 \\
&= \langle \mathrm{Aux}| \langle \mathrm{EPR}|^{\otimes n} \left(\sum_{a\in\{0,1\}} (-1)^{\mathrm{Tr}(a)} (\tilde{P}_a \otimes \mathbb{I}_{2^{2n}}) - \sum_{b\in\{0,1\}} (-1)^{\mathrm{Tr}(b)} \mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \rho_b^W \otimes \mathbb{I}_{2^n} \right)^2 |\mathrm{Aux}\rangle |\mathrm{EPR}\rangle^{\otimes n} \\
&= 2 - \sum_{a\in\{0,1\}} \langle \mathrm{Aux}| \langle \mathrm{EPR}|^{\otimes n} (\tilde{P}_a \otimes \mathbb{I}_{2^{2n}}) (\mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \rho_a^W \otimes \mathbb{I}_{2^n}) |\mathrm{Aux}\rangle |\mathrm{EPR}\rangle^{\otimes n} \\
&\quad - \sum_{a\in\{0,1\}} \sum_{b\in\{0,1\}} (-1)^{\mathrm{Tr}(a+b)} (\langle \mathrm{Aux}| \langle \mathrm{EPR}|^{\otimes n} (\tilde{P}_a \otimes \mathbb{I}_{2^{2n}}) (\mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \rho_b^W \otimes \mathbb{I}_{2^n}) \\
&\qquad - (\mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \rho_b^W \otimes \mathbb{I}_{2^n}) (\tilde{P}_a \otimes \mathbb{I}_{2^{2n}}) |\mathrm{Aux}\rangle |\mathrm{EPR}\rangle^{\otimes n}) \\
&= 2 - \sum_{a\in\{0,1\}} \langle \mathrm{Aux}| \langle \mathrm{EPR}|^{\otimes n} (\tilde{P}_a \otimes \mathbb{I}_{2^{2n}}) (\mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \rho_a^W \otimes \mathbb{I}_{2^n}) |\mathrm{Aux}\rangle |\mathrm{EPR}\rangle^{\otimes n} \\
&\quad - \sum_{a\in\{0,1\}} \sum_{b\in\{0,1\}} (-1)^{\mathrm{Tr}(a+b)} (\langle \mathrm{Aux}| \langle \mathrm{EPR}|^{\otimes n} (\tilde{P}_a \otimes \mathbb{I}_{2^{2n}}) (\mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \rho_b^W \otimes \mathbb{I}_{2^n}) \\
&\qquad - (\tilde{P}_a \otimes \mathbb{I}_{2^{2n}}) (\mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \rho_b^W \otimes \mathbb{I}_{2^n}) |\mathrm{Aux}\rangle |\mathrm{EPR}\rangle^{\otimes n})
\end{aligned}$$

$$= 2 - \sum_{a\in\{0,1\}} \langle \text{Aux}| \langle \text{EPR}|^{\otimes n} (\tilde{P}_a \otimes \mathbb{I}_{2^{2n}})(\mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \rho_a^W \otimes \mathbb{I}_{2^n}) |\text{Aux}\rangle |\text{EPR}\rangle^{\otimes n}$$
$$= \sum_{a\in\{0,1\}^n} |\left(V_A \tilde{P}_a^A V_A^* \otimes \mathbb{I}_{2^{2n}} - \mathbb{I}_{\mathscr{H}} \otimes \mathbb{I}_{2^{2n}} \otimes \rho_a^W \otimes \mathbb{I}_{2^n}\right) |\text{Aux}\rangle |\text{EPR}\rangle^{\otimes n}|^2,$$

where the second to last inequality follows from $\rho_b^W \otimes \mathbb{I}|EPR\rangle = \mathbb{I} \otimes \rho_b^W |EPR\rangle$. (8.12) follows from a similar calculation, this shows the "soundness clause" of Theorem 8.3.3.

□

8.4 Proof of Theorem 8.3.3

In this section, we show the rigidity for the μ-dependent Pauli Basis test. Before we begin, we first show the following theorem which allows us to work with synchronous strategies throughout the proof. We start by proving a corollary based on the rounding theorem (Theorem 5.1.1), which allows us to work using synchronous strategies when working with rigidity statements.

8.4.1 Rounding approximate algebraic relationships

In this subsection, we show an approximation related to almost synchronous correlations related to algebraic relationships. We remark that the below corollary is an analogue of [Vid22, Corollary 4.4] for the commuting operator model.

Corollary 8.4.1. *Let $\mathcal{G} = (\mathcal{X}^2, \mathcal{A}^2, \mu, D)$ be a synchronous game with $\{X, Z\} \subseteq \mathcal{X}$ and $\mathbb{Z}_q \subseteq \mathcal{A}$ for some $q = p^n$ for some prime power p. Furthermore, assume $\mu(X,X) = O(1)$ and $\mu(Z,Z) = O(1)$. Let $\chi : \mathbb{Z}_q^2 \to \mathbb{C}$, $\chi(a,b) = \omega_p^{ab}$, where $\omega_p = e^{\frac{2\pi i}{p}}$. Suppose for every synchronous symmetric projective strategy $\mathcal{S}^{sync} = (\mathscr{L}^2(\mathscr{B}, \tau_{\mathscr{B}}), |\tau_{\mathscr{B}}\rangle, \{A_a^x\})$ follows the "algebraic relations" property:*

> *For $X(b) = \sum_{j\in\mathbb{Z}_q} \omega_p^{bj} A_j^X$, and $Z(b) = \sum_{j\in\mathbb{Z}_q} \omega_p^{bj} A_j^Z$, where $\omega_p = e^{\frac{2\pi i}{p}}$, there exist a convex, monotone, non-decreasing function $\kappa : [0,1] \to \mathbb{R}_{\geq 0}$ such that*
>
> $$\mathop{\mathbb{E}}_{(a,b)\in\mathbb{Z}_q\times\mathbb{Z}_q} ||X(a)Z(b) - \omega_p^{ab} Z(b)X(a)|\tau_{\mathscr{B}}\rangle||^2 \leq \kappa(\omega^{co}(G, \mathcal{S}^{sync})). \tag{8.13}$$

Then the "algebraic relations" property extends to any arbitrary δ-synchronous symmetric, projective strategy $\mathcal{S} = (\mathscr{L}^2(\mathscr{A}, \tau), \sigma|\tau\rangle, \{A_a^x\}, \{B_a^x\})$ on both measurement operator $\{A_a^x\}$ and $\{B_a^x\}$, with (8.13) replaced with

$$\mathop{\mathbb{E}}_{(a,b)\in\mathbb{Z}_q\times\mathbb{Z}_q} ||(X(a)Z(b) - \omega_p^{ab} Z(b)X(a))\sigma|\tau\rangle||^2 \leq O(\eta(\omega^{co}(G, \mathcal{S}) + \text{poly}(\delta)) + O(\delta^{1/8})). \tag{8.14}$$

We remark that (8.14) is equivalent to an approximation in the state-dependent norm, and could potentially used for other rigidity statements. We remark that the proof below follows a similar structure as [Vid22, Corollary 4.4].

Proof. Define $X(a)$ and $Z(b)$ in the same way as the "algebraic relations" property on the measurement A_a^x. Let $\{P_\lambda\}$ and $\{A_a^{\lambda,x}\}$ be the set of projectors and PVMs promised by Proposition 5.2.2, and let $\mathcal{S}^\lambda = (\mathscr{L}^2(P_\lambda \mathscr{A} P_\lambda, P_\lambda \tau), \frac{1}{\sqrt{\tau(P_\lambda)}} P_\lambda |\tau\rangle, \{A_a^{\lambda,x}\})$ be the synchronous strategy constructed from $\{A_a^{\lambda,x}\}$. Let $X^\lambda(a) = \sum_{j\in\mathbb{Z}_q} \omega_p^{bj} A_j^{\lambda,X}$ and similarly $Z^\lambda(a) = \sum_{j\in\mathbb{Z}_q} \omega_p^{bj} A_j^{\lambda,Z}$. For each $a \in \mathbb{Z}_q$

$$\int_0^\infty P(\lambda)||(X(a) - X^\lambda(a))P_\lambda||_2^2 d\lambda \leq \int_0^\infty P(\lambda) \sum_{j\in\mathbb{Z}_q} ||\omega_p^{bj}(A_x^X - A_x^{\lambda,X})P_\lambda||_2^2 d\lambda \leq O(\delta^{\frac{1}{4}}), \tag{8.15}$$

where the last inequality follows from Proposition 5.2.2 and the assumption that $\mu(X,X) = O(1)$. Similarly,

$$\int_0^\infty P(\lambda)||(Z(a) - Z^\lambda(a))P_\lambda||_2^2 \leq O(\delta^{\frac{1}{4}}). \tag{8.16}$$

To show Equation (8.14), we first note that via the triangle inequality and because $|\tau\rangle$ is a tracial state

$$\begin{aligned}\mathop{\mathbb{E}}_{(a,b)} ||(X(a)Z(b)-\omega_p^{ab}Z(b)X(a))\sigma||_2^2 &= \mathop{\mathbb{E}}_{(a,b)} \langle\psi|\sigma(X(a)Z(b)-\omega_p^{ab}Z(b)X(a))\sigma|\psi\rangle \\ &\leq \int_0^\infty P(\lambda) \mathop{\mathbb{E}}_{(a,b)\in\mathbb{Z}_q\times\mathbb{Z}_q} ||(X(a)Z(b)-\omega_p^{ab}Z(b)X(a))P_\lambda||_2^2 d\lambda.\end{aligned}$$

Also, since η is convex,

$$\int_0^\infty P(\lambda) \mathop{\mathbb{E}}_{(a,b)} ||(X^\lambda(a)Z^\lambda(b)-\omega_p^{ab}Z^\lambda(b)X^\lambda(a))P_\lambda||_2^2 \leq \eta(\int_0^\infty \omega^{co}(\mathcal{G},\mathscr{S}^\lambda) \leq \eta(\omega^{co}(\mathcal{G},\mathscr{S})+\text{poly}(\delta)).$$

Hence, it is sufficient to show the following to conclude the lemma:

$$\int_0^\infty P(\lambda) \mathop{\mathbb{E}}_{(a,b)} ||\Big((X(a)Z(b)-\omega_p^{ab}Z(b)X(a))-(X^\lambda(a)Z^\lambda(b)-\omega_p^{ab}Z^\lambda(b)X^\lambda(a))\Big)P_\lambda||_2^2 d\lambda \leq O(\delta^{1/8}). \tag{8.17}$$

Rearranging Equation (8.17) via the triangle inequality

$$\begin{aligned}&\int_0^\infty P(\lambda) \mathop{\mathbb{E}}_{(a,b)} ||\Big((X(a)Z(b)-\omega_p^{ab}Z(b)X(a))-(X^\lambda(a)Z^\lambda(b)-\omega_p^{ab}Z^\lambda(b)X^\lambda(a))\Big)P_\lambda||_2^2 \\ &\leq \int_0^\infty P(\lambda) \mathop{\mathbb{E}}_{(a,b)} \Big(||X(a)Z(b)-X^\lambda(a)Z^\lambda(b)||_2^2+||\omega_p^{ab}(Z^\lambda(b)X^\lambda(a))-Z(b)X(a))||_2^2\Big) d\lambda \\ &= \int_0^\infty P(\lambda) \mathop{\mathbb{E}}_{(a,b)} ||X(a)Z(b)-X(a)Z^\lambda(b)||_2^2+||X(a)Z^\lambda(b)-X^\lambda(a)Z^\lambda(b)||_2^2 \\ &\quad +||(Z^\lambda(b)X^\lambda(a))-Z^\lambda(b)X(a)||_2^2+||(Z^\lambda(b)X(a)-Z(b)X(a))||_2^2 d\lambda.\end{aligned}$$

Since the proof for the above four components is similar, we will only include the proof for the first term below:

$$\int_0^\infty P(\lambda) \mathop{\mathbb{E}}_{(a,b)} ||X(a)Z(b)-X(a)Z^\lambda(b)||_2^2 d\lambda \leq 1\cdot\sqrt{\int_0^\infty P(\lambda) \mathop{\mathbb{E}}_b ||Z(b)-Z^\lambda(b)||_2^2 d\lambda} \leq O(\delta^{1/8}),$$

where the last inequality follows from Equation (8.16), completing the proof of the corollary. □

8.4.2 Commutation/anti-commutation property related to measurement observable

In order to show Theorem 8.3.2, we show the following proposition about the observables defined within (8.7).

Proposition 8.4.2. *Let $\mathcal{G}(\mu)$ be the μ-dependent Pauli Basis test and let $\mathscr{S}=(\mathscr{L}^2(\mathscr{A},\tau),\sigma|\tau\rangle,\{A_a^x\},\{B_a^x\})$ be a projective, tracially embeddable strategy for $\mathcal{G}(\mu)$ with $\omega^{co}(\mathcal{G}(\mu),\mathscr{S})\geq 1-\varepsilon$. Let W^A be the unitary observable define within (8.7), then*

$$\begin{aligned}\mathop{\mathbb{E}}_{(u,v)\in(\mathbb{Z}_{2^n})^2} ||X^A(u)Z^A(v)-(-1)^{u\cdot v}Z^A(v)X^A(u)||_{\sigma^2}^2 &\leq O\big(\text{poly}(\kappa(\mu),\varepsilon)\big), \\ \mathop{\mathbb{E}}_{(u,v)\in(\mathbb{Z}_{2^n})^2} ||X^B(u)Z^B(v)-(-1)^{u\cdot v}Z^B(v)X^B(u)||_{\sigma^2}^2 &\leq O\big(\text{poly}(\eta(\mu),\varepsilon)\big).\end{aligned}$$

Proof. We first assume that the strategy $(\mathscr{L}^2(\mathscr{A},\tau),|\tau\rangle,\{A_a^x\})$ is synchronous (and use $W(a)$ instead to represent the observables defined on (8.7)). This implies that each A_a^x is a projective measurements. We first wish to show that

$$\mathop{\mathbb{E}}_{(u,v)\sim\mu\times\mu} \big\|X(u)Z(v)-(-1)^{u\cdot v}Z(u)X(v)\big\|_2^2 \leq O(\text{poly}(\varepsilon)). \tag{8.18}$$

We remark that (8.18) is proven implicitly in [dlS22b, Proposition 3.8]. For $u, r \in \{0,1\}^n$, we define $A^{(\text{Pauli},W)}_{t|_u=r} = \sum_{t\in\{0,1\}^n, t|_u=r} A^{(\text{Pauli},W)}_t$. Hence, we can rewrite (8.7) as

$$W(u) = \sum_{\substack{r\in\{0,1\}^n \\ |r| \bmod 2=0}} A^{(\text{Pauli},W)}_{t|_u=r} - \sum_{\substack{r\in\{0,1\}^n \\ |r| \bmod 2=1}} A^{(\text{Pauli},W)}_{t|_u=r}$$

Let $\{A^{(\text{Coordinate, X}),u}_{t_{(\text{Pauli, X})}}\}_{((\text{Pauli, X}))\in\{0,1\}^n}$ and $\{A^{(\text{Coordinate, Z}),v}_{t_{(\text{Pauli, Z})}}\}_{((\text{Pauli, Z}))\in\{0,1\}^n}$ be the measurement operators for the question label (Coordinate, X) and (Coordinate, Z) respectively. Furthermore, for $i \in \{0,1\}$ and $W = \{X, Z\}$, define $A^{(\text{Coordinate, W}),u}_{|t_W|=i} = \sum_{\substack{t_W\in\{0,1\}^n \\ |t_W| \bmod 2=i}} A^{(\text{Coordinate, W}),u}_{t_W}$, and define the observable $W^{(\text{Coordinate}),u} = A^{(\text{Coordinate, W}),u}_{|t_W|=0} - A^{(\text{Coordinate, W}),u}_{|t_W|=1}$. Since in $\mathcal{G}(\mu)$, the question pair on the black edge gets picked with a constant probability, the players can fail these questions with at most $O(\varepsilon)$ probability over these two question pairs with strategy $\mathcal{S}$. This implies for $W \in \{X, Z\}$

$$\mathop{\mathbb{E}}_{u\sim\mu} \sum_{t_W\in\{0,1\}^n} \langle\tau| A^{(\text{Pauli},W)}_{t|_u=t_W} A^{(\text{Coordinate, W}),u}_{t_W} |\tau\rangle \le 1 - O(\varepsilon).$$

Rewriting the above equation

$$\begin{aligned}
&\mathop{\mathbb{E}}_{u\sim\mu} \sum_{t_W\in\{0,1\}^n} ||A^{(\text{Pauli},W)}_{t|_u=t_W} - A^{(\text{Coordinate, W}),u}_{t_W}||_2^2 \\
&= 2 - 2\left(\mathop{\mathbb{E}}_{u\sim\mu} \sum_{t_W\in\{0,1\}^n} \langle\tau| A^{(\text{Pauli},W)}_{t|_u=t_W} A^{(\text{Coordinate, W}),u}_{t_W}\right) \le O(\varepsilon),
\end{aligned}$$

and hence, by the triangle inequality

$$\mathop{\mathbb{E}}_{u\sim\mu} ||W(u) - W^{(\text{Coordinate}),u}||_2^2 = \mathop{\mathbb{E}}_{u\sim\mu} || \sum_{t_X\in\{0,1\}^n} (-1)^{|t_X|} \left(A^{(\text{Pauli},W)}_{t|_u=t_W} - A^{(\text{Coordinate, W}),u}_{t_W}\right) ||_2^2 \le O(\varepsilon). \tag{8.19}$$

In order to show the lemma, we first show

$$\mathop{\mathbb{E}}_{(u,v)\sim\mu\times\mu} ||X^{(\text{Coordinate}),u} Z^{(\text{Coordinate}),v} - (-1)^{u\cdot v} Z^{(\text{Coordinate}),v} X^{(\text{Coordinate}),u}||_2^2 \le O(\text{poly}(\varepsilon)), \tag{8.20}$$

then use Equation (8.19) to conclude the lemma. We consider the case where $u \cdot v = 0$ and $u \cdot v = 1$ for (8.20) separately. First, assume that $u \cdot v = 0$ in step 1 of the sampling procedure for the μ dependent Pauli Basis test. Let μ_0^2 denote the distribution $\mu \times \mu$ conditioned on the output $u \cdot v = 0$, and let $\{A^{(\text{Commutation}),u,v}_{(t_X,t_Z)}\}_{(t_X,t_Z)\in\{0,1\}^{n+n}}$ be the measurement operators corresponds to the "commutation" question label with answer pair (t_X, t_Z), and for $W \in \{X, Z\}$, let $\{A^{(\text{Commutation}),u}_{(t_W)}\}_{(t_W)\in\{0,1\}^n}$ be the measurement operators corresponds to the "(Commutation, W)" question label with answer (t_W). Since $\mathcal{S}$ succeeds with probability $1-\varepsilon$, the players can fail the (Commutation) – (Commutation, W) with probability at most $O(\varepsilon)$, for $W \in \{X, Z\}$, this implies

$$\mathop{\mathbb{E}}_{(u,v)\sim\mu_0^2} \sum_{t_X\in\{0,1\}} \langle\tau| A^{(\text{Commutation}),u,v}_{(t_X,t_Z)} A^{(\text{Commutation}),u}_{t_W} |\tau\rangle \le 1 - O(\varepsilon). \tag{8.21}$$

Rewriting the above inequality the same way as (8.19), we have

$$\mathop{\mathbb{E}}_{(u,v)\sim\mu_0^2} \sum_{(t_W,t_Z)\in\{0,1\}^2} ||A^{(\text{Commutation}),u,v}_{(t_X,t_Z)} - A^{(\text{Commutation},W),u}_{t_W}||_2^2 \le O(\varepsilon).$$

By a similar argument using the (Coordinates, W) – (Commutation, W) test,

$$\mathop{\mathbb{E}}_{(u,v)\sim\mu_0^2} \sum_{t_W\in\{0,1\}} || \sum_{t\in\{0,1\}^n, u\cdot t=t_W} A^{(\text{Coordinates},W),u}_t - A^{(\text{Commutation},W),u}_{t_W}||_2^2 \le O(\varepsilon),$$

and hence by triangle inequality

$$\mathop{\mathbb{E}}_{(u,v)\sim\mu_0^2}\sum_{(t_W,t_Z)\in\{0,1\}^2}||\sum_{t\in\{0,1\}^n,u\cdot t=t_W}A_t^{(\text{Coordinates},W),u}-A_{(t_X,t_Z)}^{(\text{Commutation}),u,v}||_2^2\leq O(\varepsilon),$$

Hence, by the triangle inequality, Jensen's inequality and Cauchy-Schwarz inequality,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{(u,v)\sim\mu_0^2}\sum_{\substack{t_X\in\{0,1\}^n\\t_Z\in\{0,1\}^n}}||A_{t_X}^{(\text{Coordinate,X}),u}A_{t_Z}^{(\text{Coordinate, Z}),v}-A_{t_Z}^{(\text{Coordinate, Z}),v}A_{t_X}^{(\text{Coordinate,X}),u}||_2^2\\
&\leq\mathop{\mathbb{E}}_{(u,v)\sim\mu_0^2}\sum_{\substack{t_X\in\{0,1\}^n\\t_Z\in\{0,1\}^n}}||(A_{t_X}^{(\text{Coordinate,X}),u}-A_{(t_X,t_Z)}^{(\text{Commutation}),u,v})A_{t_Z}^{(\text{Coordinate, Z})v}\\
&\quad+A_{(t_X,t_Z)}^{(\text{Commutation}),u,v}(A_{t_Z}^{(\text{Coordinate, Z}),v}-A_{(t_X,t_Z)}^{(\text{Commutation}),u,v})\\
&\quad+A_{(t_X,t_Z)}^{(\text{Commutation}),u,v}(A_{(t_X,t_Z)}^{(\text{Commutation}),u,v}-A_{t_X}^{(\text{Coordinate,X}),u})\\
&\quad+(A_{(t_X,t_Z)}^{(\text{Commutation}),u,v}-A_{t_Z}^{(\text{Coordinate, Z}),v})A_{t_X}^{(\text{Coordinate,X}),u}||_2^2\leq O(\sqrt{\varepsilon}).
\end{aligned}\tag{8.22}$$

Returning to (8.20), for $W\in\{X,Z\}$ and $u\in\{0,1\}^n$, since $W^{(\text{Coordinate}),u}$ is a binary observable, we can rewrite it as

$$W^{(\text{Coordinate}),u}=\mathbb{I}-2\cdot A_{|t_W|=1}^{(\text{Coordinate, W}),u}.$$

Hence, we can rewrite (8.20) as

$$\begin{aligned}
&\mathop{\mathbb{E}}_{(u,v)\sim\mu_0}||X^{(\text{Coordinate}),u}Z^{(\text{Coordinate}),v}-Z^{(\text{Coordinate}),v}X^{(\text{Coordinate}),u}||_2^2\\
&=2\mathop{\mathbb{E}}_{(u,v)\sim\mu_0}||A_{|t_X|=1}^{(\text{Coordinate, X}),u}A_{|t_Z|=1}^{(\text{Coordinate, Z}),v}-A_{|t_Z|=1}^{(\text{Coordinate, Z}),v}A_{|t_X|=1}^{(\text{Coordinate, X}),u}||_2^2\\
&\leq 2\mathop{\mathbb{E}}_{(u,v)\sim\mu_0}\sum_{\substack{t_X\in\{0,1\}^n\\|t_X|\bmod 2=1\\t_Z\in\{0,1\}^n\\|t_Z|\bmod 2=1}}||A_{t_X}^{(\text{Coordinate,X}),u}A_{t_Z}^{(\text{Coordinate, Z}),v}-A_{t_Z}^{(\text{Coordinate, Z}),v}A_{t_X}^{(\text{Coordinate,X}),u}||_2^2\\
&\leq 2\mathop{\mathbb{E}}_{(u,v)\sim\mu_0}\sum_{\substack{t_X\in\{0,1\}^n\\t_Z\in\{0,1\}^n}}||A_{t_X}^{(\text{Coordinate,X}),u}A_{t_Z}^{(\text{Coordinate, Z}),v}-A_{t_Z}^{(\text{Coordinate, Z}),v}A_{t_X}^{(\text{Coordinate,X}),u}||_2^2\leq O(\sqrt{\varepsilon}),
\end{aligned}\tag{8.23}$$

where the third line follows from $||\cdot||_2\geq 0$, and the last line follows from (8.22). This concludes the case for $u\cdot v=0$. Now we consider the case where $u\cdot v=1$. Let μ_1^2 denote the distribution $\mu\times\mu$ conditioned on the output $u\cdot v=1$. Since $\mathscr{S}$ succeeds with probability $1-\varepsilon$, the players can fail the "anti-commutation test" with probability at most $O(\varepsilon)$. Since the anti-commutation test is a synchronous magic square game with the question labels "variable 1" and "variable 5" are replaced by the question labels $(\text{Coordinate},X),u$ and $(\text{Coordinate},X),v$ respectively. By Theorem 8.3.1

$$\mathop{\mathbb{E}}_{(u,v)\sim\mu_1^2}||X^{(\text{Coordinate}),u}Z^{(\text{Coordinate}),v}+Z^{(\text{Coordinate}),v}X^{(\text{Coordinate}),u}||_2^2\leq O(\text{poly}(\varepsilon)),\tag{8.24}$$

which concludes the case for $u\cdot v=1$. Combining (8.23) and (8.24) shows (8.20). By combining (8.20) and (8.19) via the triangle inequality, we have

$$\mathop{\mathbb{E}}_{(u,v)\sim\mu\times\mu}||X(u)Z(v)-(-1)^{u\cdot v}Z(u)X(v)||_2^2\leq O(\text{poly}(\varepsilon)),$$

showing (8.7). Given that each $\{A_q^{\text{Pauli},W}\}$ is a projective measurement for $W\in\{X,Z\}$, the observable $\{W(u)\}_{u\in\{0,1\}}$ is an (exact) unitary representation for the group $\mathbb{Z}_{2^n}$. Hence, we apply Theorem 8.1.2 on (8.18) to obtain

$$\mathop{\mathbb{E}}_{(u,v)}\left\|X(u)Z(v)-(-1)^{u\cdot v}Z(u)X(v)\right\|_2^2\leq\kappa(\mu)^2\cdot O(\text{poly}(\varepsilon)),\tag{8.25}$$

where the expectation is over the uniform distribution over $\{0,1\}^{2n}$. This concludes the lemma for the case where $\mathscr{S}$ is synchronous.

For the general case, note both the labels (Pauli, X) and (Pauli, Z) appear as a synchronicity question pair with probability $O(1)$ (i.e. independent of n) within $\mathscr{G}(\mu)$. Thus, we can use Corollary 8.4.1 to extend this result to the general case, concluding the proof for this proposition. □

8.4.3 Proof of the rigidity statement

Using Proposition 8.4.2, we are ready to give a proof for Theorem 8.3.2. We remark that this proof follows a similar structure as [JNV+22a, Lemma A.24].

Proof. By using Lemma 3.3.1, we can assume that the measurement operators for $\mathscr{S}$ are PVMs with $O(\text{poly}(\varepsilon))$ overhead. With this assumption, for $W \in \{X,Z\}$ and $P \in \{A,B\}$, $\{W^P(u)\}$ forms an unitary representation for the group $\mathbb{Z}_{2^n}$. This implies for all $u,v \in \mathbb{Z}_{2^n}$

$$W^P(u)W^P(v) = W^P(u+v).$$

Combine the above equation with Proposition 8.4.2, we see that $\{X^P(u)\} \cup \{Z^P(u)\}$ forms an $(O\big(\text{poly}(\kappa(\mu),\varepsilon)\big),\sigma^2)$-representation for the group $H^{(n)}$. Hence, by Theorem 8.1.1 (and the structure of the left regular representation for the group $H^{(n)}$), there exist two isometries $V_A : \mathscr{H} \to \mathscr{H} \otimes \mathbb{C}^{2^{2n}}$ and $V_B : \mathscr{H} \to \mathscr{H} \otimes \mathbb{C}^{2^{2n}}$ such that for all $(a,b) \in \mathbb{Z}_{2^n}$

$$||X^A(a)Z^A(b) - V_A^*(\mathbb{I}_{\mathscr{H}} \otimes \rho^X(a)\rho^Z(b) \otimes \mathbb{I}_{2^{2n}})V_A||_{\sigma^2} \leq O(\text{poly}(\kappa(\mu),\varepsilon)), \tag{8.26}$$

$$||X^B(a)Z^B(b) - V_B^*(\mathbb{I}_{\mathscr{H}} \otimes \rho^X(a)\rho^Z(b) \otimes \mathbb{I}_{2^{2n}})V_B||_{\sigma^2} \leq O(\text{poly}(\kappa(\mu),\varepsilon)), \tag{8.27}$$

We note that on the second item of the above equation, Theorem 8.1.1 is applied on the *commutant* $\mathscr{A}'$ rather than $\mathscr{A}$. We wish to first show (8.8) by using the synchronous condition to connect the two exact representations. By the self-consistency test (self-loop), which appears with $O(1)$ probability, for all $W \in \{X,Z\}$, we have

$$\sum_{t\in\{0,1\}} \langle\tau|\sigma A_t^{(\text{Pauli},W)} B_t^{(\text{Pauli},W)}\sigma|\tau\rangle \geq 1 - O(\varepsilon).$$

Hence, via the triangle inequality, for all $a \in \{0,1\}^n$

$$\langle\tau|\sigma W^A(a)W^B(a)\sigma|\tau\rangle \geq 1 - O(\varepsilon), \tag{8.28}$$

where we see that $X^B(a) \in \mathscr{A}$ in this formulation. We will first consider (8.8), we wish to first show

$$\begin{aligned}&\langle\tau|\sigma V_A^*(V_B^* \otimes \mathbb{I}_{2^{2n}}^A)(\mathbb{I}_{\mathscr{H}} \otimes \big(|EPR\rangle\langle EPR|^{\otimes n}\big)^{A_1B_1} \otimes \mathbb{I}_{2^{2n}}^{A_2B_2})(V_B \otimes \mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle \\ &\quad \geq 1 - O(\text{poly}(\kappa(\mu),\varepsilon)).\end{aligned} \tag{8.29}$$

Where, $\mathscr{H}^A = \mathbb{C}^{2^{2n}}$ is the Hilbert space created by the isometry V^A and $\mathscr{H}^B = \mathbb{C}^{2^{2n}}$ are the space created by the isometry V^B. For $W \in \{X,Z\}$, denote $H_W = \mathbb{E}_{a\sim\{0,1\}^n} \rho^W(a) \otimes \rho^W(a)$. We see that

$$\begin{aligned}
&1 - \langle\tau|\sigma V_A^*(V_B^* \otimes \mathbb{I}_{2^{2n}}^A)(\mathbb{I}_{\mathscr{H}} \otimes H_W^{A_1B_1} \otimes \mathbb{I}_{2^n}^{A_2} \otimes \mathbb{I}_{2^n}^{B_2})(V_B \otimes \mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle \\
&= 1 - \mathop{\mathbb{E}}_{a\sim\{0,1\}^n} \langle\tau|\sigma V_A^*(V_B^* \otimes \mathbb{I}_{2^{2n}}^A)\mathbb{I}_{\mathscr{H}} \otimes \big(\rho^W(a) \otimes \mathbb{I}_{2^n}\big)^A \otimes \big(\rho^W(a) \otimes \mathbb{I}_{2^n}\big)^B (V_B \otimes \mathbb{I}_{2^{2n}})V_A\sigma|\tau\rangle \\
&= 1 - \mathop{\mathbb{E}}_{a\sim\{0,1\}^n} \langle\tau|\sigma V_A^* \left(\mathbb{I}_{\mathscr{H}} \otimes \rho^W(a) \otimes \mathbb{I}_{2^n}\right) V_A V_B^* \left(\mathbb{I}_{\mathscr{H}} \otimes \rho^W(a) \otimes \mathbb{I}_{2^n}\right) V_B\sigma|\tau\rangle \\
&= \left(1 - \mathop{\mathbb{E}}_{a\sim\{0,1\}^n} \langle\tau|\sigma W^A(a)W^B(a)\sigma|\tau\rangle\right) \\
&\quad + \mathop{\mathbb{E}}_{a\sim\{0,1\}^n} \langle\tau|\sigma \left(W^A(a) - V_A^* \left(\mathbb{I}_{\mathscr{H}} \otimes \rho^W(a) \otimes \mathbb{I}_{2^n}\right) V_A\right) V_B^* \left(\mathbb{I}_{\mathscr{H}} \otimes \rho^W(a) \otimes \mathbb{I}_{2^n}\right) V_B\sigma|\tau\rangle \\
&\quad + \mathop{\mathbb{E}}_{a\sim\{0,1\}^n} \langle\tau|\sigma W^A(a) \left(W^B(a) - V_B^* \left(\mathbb{I}_{\mathscr{H}} \otimes \rho^W(a) \otimes \mathbb{I}_{2^n}\right) V_B\right)\sigma|\tau\rangle \\
&\leq O(\text{poly}(\kappa(\mu),\varepsilon)),
\end{aligned}$$

where the last line follows from (8.27) and (8.28). We note that the braket around the tensor product in the above equation helps keep track of the register of each space. Hence, we have

$$\langle\tau|\sigma V_A^*(V_B^*\otimes\mathbb{I}_{2^n}^A)(\mathbb{I}_{\mathscr{H}}\otimes H_W^{A_1B_1}\otimes\mathbb{I}_{2^n}^{A_2}\otimes\mathbb{I}_{2^n}^{B_2})(V_B\otimes\mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle\geq 1-O(\mathrm{poly}(\kappa(\mu),\varepsilon)). \tag{8.30}$$

Recall that

$$H_XH_Z=\mathop{\mathbb{E}}_{(a,b)\in\mathbb{Z}_{2^n}}\rho^X(a)\rho^Z(b)\otimes\rho^X(a)\rho^Z(b)=(|EPR\rangle\langle EPR|)^{\otimes n}.$$

Hence to show (8.29), it is sufficient to show that

$$\langle\tau|\sigma V_A^*(V_B^*\otimes\mathbb{I}_{2^{2n}}^A)(\mathbb{I}_{\mathscr{H}}\otimes(H_XH_Z)^{A_1B_1}\otimes\mathbb{I}_{2^{2n}}^{A_2B_2})(V_B\otimes\mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle\geq 1-O(\mathrm{poly}(\kappa(\mu),\varepsilon)).$$

or

$$\begin{aligned}&\left\|(\mathbb{I}_{\mathscr{H}}\otimes H_Z^{A_1B_1}\otimes\mathbb{I}_{2^{2n}}^{A_2B_2})(V_B\otimes\mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle-(\mathbb{I}_{\mathscr{H}}\otimes H_X^{A_1B_1}\otimes\mathbb{I}_{2^{2n}}^{A_2B_2})(V_B\otimes\mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle\right\|\\&\quad\leq O(\mathrm{poly}(\kappa(\mu),\varepsilon)).\end{aligned} \tag{8.31}$$

To show (8.31), by the triangle inequality, we have

$$\begin{aligned}&\left\|\left(\mathbb{I}_{\mathscr{H}}\otimes H_Z^{A_1B_1}\otimes\mathbb{I}_{2^{2n}}^{A_2B_2}\right)(V_B\otimes\mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle-\left(\mathbb{I}_{\mathscr{H}}\otimes H_X^{A_1B_1}\otimes\mathbb{I}_{2^{2n}}^{A_2B_2}\right)(V_B\otimes\mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle\right\|\\&\leq\left\|(\mathbb{I}_{\mathscr{H}}\otimes H_Z^{A_1B_1}\otimes\mathbb{I}_{2^{2n}}^{A_2B_2})(V_B\otimes\mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle-(V_B\otimes\mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle\right\|\\&\quad+\left\|(V_B\otimes\mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle-(H_X^{A_1B_1}\otimes\mathbb{I}_{2^{2n}}^{A_2B_2}\otimes\mathbb{I}_{\mathscr{H}})(V_B\otimes\mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle\right\|\\&\leq O(\mathrm{poly}(\kappa(\mu),\varepsilon)),\end{aligned}$$

where the last line follows from (8.30). This shows (8.29). Let

$$|\widehat{aux}\rangle=\left(\mathbb{I}_{\mathscr{H}}\otimes(\langle EPR|^{\otimes n})^{A_1B_1}\otimes\mathbb{I}_{2^{2n}}^{A_2B_2}\right)(V_B\otimes\mathbb{I}_{2^{2n}}^A)V_A\sigma|\tau\rangle,$$

and by setting $|aux\rangle=\frac{1}{\||\widehat{aux}\rangle\|}|\widehat{aux}\rangle$, we conclude the argument for (8.8).

Now we wish to show (8.9), since the proof for $P=A$ is the same as $P=B$, we show only the case where $P=A$ below. By labelling the space and rearranging the equation

$$\begin{aligned}&\left\|\left((V_AW^A(u)V_A^*\otimes(\mathbb{I}_{2^{2n}})^{B_1B_2})^{\mathscr{H}A_1A_2}-\mathbb{I}_{\mathscr{H}}\otimes(\rho^W(u))^{A_1}\otimes(\mathbb{I}_{2^{3n}})^{A_2B_1B_2}\right)|\mathrm{Aux}\rangle\,(|\mathrm{EPR}\rangle^{\otimes n})^{A_1B_1}\right\|^2\\&\leq 2-2\,\langle\mathrm{Aux}|\langle\mathrm{EPR}|^{\otimes n}\left((V_AW^A(u)V_A^*)(\mathbb{I}_{\mathscr{H}}\otimes\rho^W(u)\otimes\mathbb{I}_{2^n})^{\mathscr{H}A_1A_2}\right)\otimes(\mathbb{I}_{2^{2n}})^{B_1B_2}|\mathrm{Aux}\rangle\,|\mathrm{EPR}\rangle^{\otimes n}\\&=2-2\,\langle\mathrm{Aux}|\langle\mathrm{EPR}|^{\otimes n}\left((V_AW^A(u)V_A^*)\right)\otimes(\rho^W(u)\otimes\mathbb{I}_{2^n})^{B_1B_2}|\mathrm{Aux}\rangle\,|\mathrm{EPR}\rangle^{\otimes n}\\&\leq 2-2\,\langle\tau|\sigma V_A^*(V_B^*\otimes\mathbb{I}_{2^{2n}})\left((V_AW^A(u)V_A^*)\right)\otimes(\rho^W(u)\otimes\mathbb{I}_{2^n})^{B_1B_2}(V_B\otimes\mathbb{I}_{2^{2n}})V_A\sigma|\tau\rangle\\&\quad+O(\mathrm{poly}(\kappa(\mu),\varepsilon))\\&\leq 2-2\,\langle\tau|\sigma W^A(u)V_B^*(\mathbb{I}_{\mathscr{H}}\otimes\rho^X(a)\rho^Z(b)\otimes\mathbb{I}_{2^{2n}})V_B\sigma|\tau\rangle+O(\mathrm{poly}(\kappa(\mu),\varepsilon))\\&\leq 2-2\,\langle\tau|\sigma W^A(u)W^B(u)\sigma|\tau\rangle+O(\mathrm{poly}(\kappa(\mu),\varepsilon))\\&\leq O(\mathrm{poly}(\kappa(\mu),\varepsilon)),\end{aligned} \tag{8.32}$$

where the second line follows from $\rho^W(u)\otimes\mathbb{I}_{2^n}|EPR\rangle=\mathbb{I}_{2^n}\otimes\rho^W(u)\mathbb{I}_{2^n}|EPR\rangle$. The fourth line of the above equation follows from Cauchy-Schwarz, (8.8) and for two arbitrary unit vectors $|\psi\rangle,|\phi\rangle\in\mathscr{H}$ such that $\||\psi\rangle-|\phi\rangle\|^2\leq\varepsilon$ for some $\varepsilon>0$ and $A\leq\mathbb{I}$

$$\begin{aligned}\langle\psi|A|\psi\rangle-\langle\phi|A|\phi\rangle&=(\langle\psi|-\langle\phi|)A|\psi\rangle+\langle\phi|A(|\psi\rangle-|\phi\rangle)\\&\leq\||\psi\rangle-|\phi\rangle\|\cdot(\|A|\psi\rangle\|+\|A|\phi\rangle\|)\leq\sqrt{\varepsilon}\cdot 1.\end{aligned}$$

The fifth line on (8.32) follows from the commutation of the isometry, the sixth line follows from (8.27) and the Cauchy-Schwarz's inequality, and the last line follows from (8.28). This concludes the proof for Theorem 8.3.2. □

CHAPTER 9

Question Reduction part 2: The introspection protocol

We introduce the introspection protocol in this chapter, which is the protocol required for the question reduction step for the compression theorem Proposition 7.4.2. Intuitively, the introspection protocol asks both provers to sample their own question pairs for the nth game instance, and play the game based on the question pair they sampled. This procedure might first seem counter-intuitive, since the provers can always pre-select a question pair before the game rather than sampling it honestly during the interaction. As seen later in this chapter, the trick for the verifier is to use the n^α-Pauli basis test introduced in the last chapter to "force" the provers to generate fresh shared randomness between them, using Z measurements on their shared EPR pairs. In fact, it was shown that if the provers are only allowed to measure using noisy EPR in the MIP* model, the complexity ends up being correlated inversely with the noise tolerance, and if given a constant amount of noise (in fidelity), the complexity automatically drops down to NEXP, the same as classical MIP [DFN+23].

9.1 The Introspection protocol

In this subsection, we present the introspection protocol for a game that is CL samplable, which provides the algorithm $\texttt{QuestionReduction}_{\alpha,k}$ required in Proposition 7.4.2. For $\alpha, n \in \mathbb{N}$, and a CL samplable game $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$, where the question distribution μ is a (k, m, p) CL distribution with $m \cdot p \leq n^\alpha$, we define the sampling procedure for the $(\mathcal{G}, \alpha, k)$-introspection protocol in Figure 9.2, the verification procedure in Figure 9.3, and a typed graph for the input distribution in Figure 9.1. We remark that the introspection protocol is almost the same as the one presented in [JNV+22a, Figure 10], with minor adjustments for clarity.

We give a simple example to illustrate the introspection protocol. Suppose $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$ is a synchronous game, where μ is a $(1, m, 3)$ CL distribution defined by the CL functions

$$\mathtt{L}^A(s_0, s_1, s_2) = (s_0 + s_1, 0, 0) \quad \text{and} \quad \mathtt{L}^B(s_0, s_1, s_2) = (s_1 + s_2, 0, 0), \tag{9.1}$$

for all $(s_0, s_1, s_2) \in \mathbb{F}_2^3$. The introspection protocol first forces the provers to prepare three copies of the $|\mathrm{ME}_2\rangle$ by using the 3-qubit Pauli basis test as a subroutine. In the ideal scenario, the verifier wants the prover (Alice) who receives the question arising from $\mathtt{L}^A$ to perform a Pauli Z measurement on the first 2 copies of $|\mathrm{ME}_2\rangle^{\otimes 3}$ in order to sample two random bits $(s_0^A, s_1^A) \sim \{0,1\}^2$, compute $\mathtt{L}^A(s_0^A, s_1^A, 0)$ to obtain her question for $\mathcal{G}$ and play the game accordingly. Intuitively, this "samples" the first two bits, s_0 and s_1, which is the minimum amount of information Alice needs to compute $\mathtt{L}^A$. The other prover, Bob, should perform a Pauli Z measurement on the last two copies $|\mathrm{ME}_2\rangle^{\otimes 3}$ in order to sample $(s_1^B, s_2^B) \sim \{0,1\}^2$, and calculate $\mathtt{L}^B(0, s_1^B, s_2^B)$ to

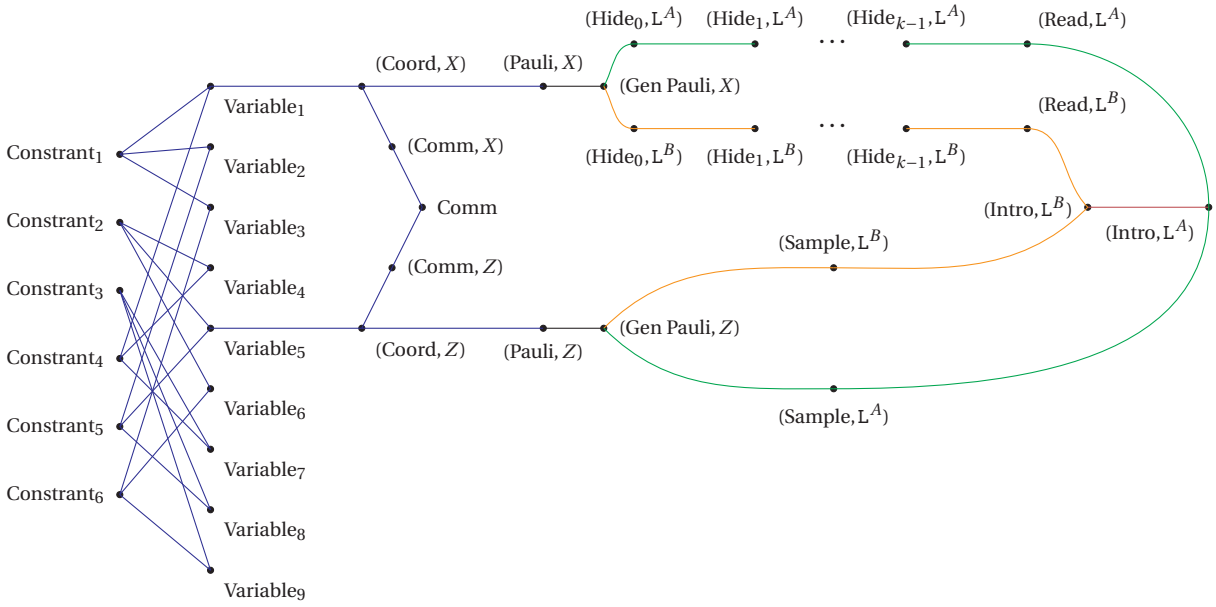

Figure 9.1: The typed graph $(\mathrm{T}_k^{\mathrm{Intro}}, \mathrm{E}_k^{\mathrm{Intro}})$ for a $(n^\alpha, k, \mathscr{G})$-introspection protocol, where each of the vertices above also contains a self-loop (which is a black edge). The purple edges are intuitively the question pair in which the provers are asked to sample honestly from the question distribution, and play the original game. The orange and green edges are questions which are designed to make sure that the provers perform the correct measurement such that $\mathrm{L}^A(s)$ and $\mathrm{L}^B(s)$ can be sampled correctly. The blue edges correspond to question pairs from the n^α-Pauli basis test. We remark that the typed graph above only depends on the parameter k, and functions $\mathrm{L}^A, \mathrm{L}^B$ within the question label are presented for clarity.

Question label	Question content	Answer format
n^α-PB question labels	See Figure 8.3	
(Gen Pauli, W)		$s_W \in \mathbb{F}_{2^p}^m$
(Hide$_0$, L^P)		$(t^\perp_{\leq 0}, r_{>0}) \in V_0 \times V_{>0}$
(Hide$_i$, L^P), $i \in [k] \setminus 0$		$(t^{\mathrm{Line}}_{<i}, t^\perp_{\leq i}, r_{>i}) \in V_{<i} \times V_{\leq i} \times V_{>i}$
(Read, L^P)		$(t^\perp_{\mathrm{Read},P}, t^{\mathrm{Line}}_{\mathrm{Read},P}, a_{\mathrm{Read},P}) \in \mathbb{F}_{2^p}^m \times \mathbb{F}_{2^p}^m \times \mathscr{A}$
(Sample, L^P)		$(s_{\mathrm{Sample}}, a_{\mathrm{Sample},P}) \in \mathbb{F}_{2^p}^m \times \mathscr{A}$
(Intro, L^P)		$(x_P, a_P) \in \mathbb{F}_{2^p}^m \times \mathscr{A}$

Figure: Q and A format for the $(\mathscr{G}, n^\alpha, k)$-introspection protocol, with $W \in \{X, Z\}$ and $P \in \{A, B\}$.

Sampling procedure

1. Sample $(u_X, u_Z) \in S_{n^\alpha}^{\mathrm{PB}} \times S_{n^\alpha}^{\mathrm{PB}}$, $(n_0, n_1) \in \mathrm{T}_k^{\mathrm{Intro}} \times \mathrm{T}_k^{\mathrm{Intro}}$, where $(\mathrm{T}_k^{\mathrm{Intro}}, \mathrm{E}_k^{\mathrm{Intro}})$ is defined in Figure 9.1, and perform rejection sampling until $(n_0, n_1) \in \mathrm{E}_k^{\mathrm{Intro}}$.
2. Send the question label and question content corresponding to n_0 to one of the provers, and send the question content corresponding to the n_1 to the other prover.

Figure 9.2: The description for the sampling procedure for the $(\mathscr{G}, n^\alpha, k)$-introspection protocol. Where $\mathscr{G} = (\mathscr{X}, \mathscr{A}, \mu, D)$, such that the distribution μ is a (k, m, p) CL distribution with $m \cdot p \leq n^\alpha$. The CL distribution is defined by two CL functions L^A and L^B with registers $\{V_i\}_{i \in [k]}$ with $\mathbb{F}_{2^p}^m = \bigcup_{i \in [k]} V_i$.

obtain his half of the question pair and output his answer accordingly. By the properties of entanglement, if Alice and Bob have performed the procedure properly, $s_1^A = s_1^V$, and hence the question distribution sampled by the two provers is precisely the same as μ. In the introspection game, the verifier wants the provers to perform this "ideal scenario" when given the (Intro, L^P) question label in Figure 9.1.

To enforce honesty from the provers, the verifier cross-references the measurements made by the provers with those made in the (Pauli, W) question pair for $W \in \{X, Z\}$ (which, recall, forces the provers to perform an all X or all Z measurement on all $|\mathrm{ME}_2\rangle$ states by the properties of the n^α Pauli basis test). In particular, the verifier wants to make sure the provers perform the following task correctly:

1. The provers should only measure the register they need in order to compute the function $\mathtt{L}^P$. On the example given in Equation (9.1), the prover receiving the question which arises from $\mathtt{L}^A$ should **only** perform the Pauli Z on the first 2 copies of $|\mathrm{ME}_2\rangle^{\otimes 3}$, and not measure the last copy (i.e. the kernel of $\mathtt{L}^A$).

2. After sampling the bits required to compute the function $\mathtt{L}^P$, the provers have to correctly apply the function (instead of using some pre-prepared question pair).

To ensure the first task is performed correctly, the verifier cross-references the (Intro, $\mathtt{L}^P$) question with the question label (Read, $\mathtt{L}^P$), in which the provers, in addition to performing the Pauli Z measurement, must also make a Pauli X measurement on the kernel space of $\mathtt{L}^A$, and are expected to output the same answer as the (Intro, $\mathtt{L}^P$) question. On the example above, since Alice, given the (Intro, $\mathtt{L}^A$) question label can only perform an X or Z measurement on the third qubit, her answer must not depend on the measurement outcome for the third qubit. The

Verification procedure

Synchronicity/EPR test.

- (Self-loop): The provers win iff they output the same answer.
- (n^α-PB question labels): If (n_0, n_1) is a blue edge, refer to Figure 8.3.
- (Pauli, W) – (Gen Pauli, W): Let $\pi_{p\cdot m}(t_W)$ be the first $p\cdot m$ bits of t_W, the provers win iff the canonical representation of s_W is equal to $\pi_{p\cdot m}(t_W)$.

Hiding test. We identify $t^{\mathrm{Line}}_{<0} = 0 \in \mathbb{F}_{2^p}$.

- (Gen Pauli, X) – (Hide$_0$, $\mathtt{L}^P$): Write $s_X = (s_X)_0 + (s_X)_0^C + (s_X)_{>0} \in \ker \mathtt{L}^P_{0,0} \oplus \ker \mathtt{L}^P_{0,0}{}^C \oplus V_{>0}$ (where the canonical complement is defined over V_0), the prover wins iff $(s_X)_0^C = t^\perp_{\le 0}$ and $(s_X)_{>0} = r_{>0}$.
- (Hide$_i$, $\mathtt{L}^P$) – (Hide$_{i+1}$, $\mathtt{L}^P$) for $i \in [k]$: Write

$$t^\perp_{\le i+1} = \tilde{t}^\perp_{\le i} + \tilde{t}^\perp_{i+1} \in V_{\le i} \oplus V_{i+1}, \quad t^{\mathrm{Line}}_{<i+1} = \tilde{t}^{\mathrm{Line}}_{<i} + \tilde{t}^{\mathrm{Line}}_{i} \in V_{<i} \oplus V_i,$$
$$r_{>i} = \bar{r}_i + \bar{r}_i^C + \bar{r}_{>i+1} \in \ker\left(\mathtt{L}^P_{i+1, t^{\mathrm{Line}}_{<i+1}}\right) \oplus \ker\left(\mathtt{L}^P_{i+1, t^{\mathrm{Line}}_{<i+1}}\right)^C \oplus V_{>i+1},$$

  In the notation above, the bar above the variable indicates that the element is decomposed from (Hide$_i$) and the tilde above the variable refers to elements from (Hide$_{i+1}$). The complement for $\ker\left(\mathtt{L}^P_{i+1, t^{\mathrm{Line}}_{<i+1}}\right)^C$ is over the subspace V_i.
  The provers win iff

$$\tilde{t}^\perp_{\le i} = t^\perp_{\le i}, \quad \bar{r}_i^C = \tilde{t}^\perp_{i+1}, \quad \tilde{t}^{\mathrm{Line}}_{<i} = t^{\mathrm{Line}}_{<i}, \quad \bar{r}_{>i+1} = r_{>i+1}.$$

- (Hide$_{k-1}$, $\mathtt{L}^P$) – (Read, $\mathtt{L}^P$): The provers win iff $t_{\mathrm{Read},P} = t_{k-1}$, and $t^\perp_{\mathrm{Read},P} = t^\perp_{k-1}$.
- (Read, $\mathtt{L}^P$) – (Intro, $\mathtt{L}^P$): The provers win iff $t^{\mathrm{Line}}_{\mathrm{Read},P} = x_P$, and $a_{\mathrm{Read},P} = a_P$.

Sampling test.

- (Gen Pauli, Z) – (Sample, $\mathtt{L}^P$): The provers win iff $s_Z = s_{\mathrm{Sample}}$.
- (Sample, $\mathtt{L}^P$) – (Intro, $\mathtt{L}^P$): The provers win iff $\mathtt{L}^P(s_{\mathrm{Sample}}) = x_P$ and $a_{\mathrm{Sample},P} = a_P$.

Introspection of $\mathcal{G}$

- (Intro, $\mathtt{L}^A$) – (Intro, $\mathtt{L}^B$): The provers win iff $D(x_A, x_B, a_A, a_B) = 1$.

Figure 9.3: The description for the verification procedure for the $(\mathcal{G}, n^\alpha, k)$-introspection protocol.

(Read, $\mathtt{L}^P$) question label is then cross-referenced with the (Pauli, X) question from the 3-qubit Pauli basis test to ensure consistency for the X measurement on the third qubit. To ensure the second task, the verifier cross-references the (Intro, $\mathtt{L}^P$) question with the question label (Sample, $\mathtt{L}^P$), in which the provers are expected to sample the entirety of the seed s by performing Pauli Z measurements on **all** of their $|\mathrm{ME}_2\rangle$ bits, compute the corresponding question $\mathtt{L}^P(s)$ and generate the corresponding answer. The (Sample, $\mathtt{L}^P$) question label is cross-referenced with the (Pauli, Z) question to ensure consistency.

In general, there are two additional problems. If the CL function $\mathtt{L}^P$ is a level k CL function, then the "Read" question cannot be cross-checked with the (Pauli, X) question, since the kernel space of the linear function for each level depends on the computation step from the previous level. Intuitively, the behaviour of the prover for the "Read" question is enforced by a series of "Hide" questions, each designed to enforce the "honest measurement" for the Read question for one level. Since a CL distribution is defined using two CL functions which map subspaces of $\mathbb{F}_{2^p}^m$ rather than of $\mathbb{F}_2^m$, generalized Pauli measurements are needed for the introspection protocol. In combination with Lemma 8.2.1, we see that for any $p \in \mathbb{N}$, the $p \cdot n$ qubit Pauli basis test can also serve as a rigidity test for generalized Pauli measurement over $|\mathrm{ME}_p\rangle$, and we use the (Gen Pauli, W) to convert between these two types of self-test.

We have the following theorem regarding the $(\mathcal{G}, n^\alpha, k)$-introspection protocol. We remark that in comparison to [JNV+22a, Theorem 8.3], there is no longer dependency on α (the variable R or $n\alpha$ in [JNV+22a]). This is due to our EPR tester (the n-qubit Pauli basis test) not using the low-degree test as a part of the subroutine.

Theorem 9.1.1 (Properties of the $(\mathcal{G}, n^\alpha, k)$-introspection protocol)**.** *Let $n, \alpha \in \mathbb{N}$, and let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$ be a k-th CL samplable game where the question distribution μ is a (k, m, p) CL distribution with $m \cdot p \leq n^\alpha$. Let $\mathcal{G}^{intro} = (\mathcal{X}^{intro}, \mathcal{A}^{intro}, \mu^{intro}, D^{intro})$ be the (typed) $(\mathcal{G}, \alpha, k)$-introspection protocol specified in Figure 9.2 and Figure 9.3. For $t \in \{*, co\}$, the following holds:*

- *(Sample complexity): $\mathcal{G}^{intro}$ is samplable via a $(\mathtt{T}_k^{Intro}, \mathtt{E}_k^{Intro}, \{\mathtt{L}^v\}_{v \in \mathtt{T}^{Intro}})$ typed CL distribution, where each $\mathtt{T}^{Intro} : \mathbb{F}_2^{2 \cdot \alpha \cdot \lfloor \log(n) \rfloor}$ is a first level CL function. Furthermore, the question distribution only depends on the parameter n^α and k.*
- *(Completeness): If there exists a perfect oracularizable strategy for $\mathcal{G}$ in model t, then there exists a perfect oracularizable strategy for $\mathcal{G}^{intro}$ in model t.*
- *(Soundness): There exists a polynomial $\mathbf{P}^{Intro}(\varepsilon, k, \alpha)$ such that*

$$\omega^t(\mathcal{G}) \leq 1 - \varepsilon \Longrightarrow \omega^t(\mathcal{G}^{Intro}) \leq 1 - \mathbf{P}^{Intro}(\varepsilon, \exp(k)).$$

We prove the "sample complexity" and the "completeness" clauses below and the "soundness" clause in Section 9.2.

Proof. Fix $n, \alpha, k \in \mathbb{N}$, model $t \in \{*, co\}$. Let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$ be a game that satisfies the description for Theorem 9.1.1, and let $\mathcal{G}^{\mathrm{Intro}} = (\mathcal{X}^{\mathrm{Intro}}, \mathcal{A}^{\mathrm{Intro}}, \mu^{\mathrm{Intro}}, D^{\mathrm{Intro}})$ be the (typed)-introspection protocol. Let $(\mathtt{T}_k^{\mathrm{Intro}}, \mathtt{E}_k^{\mathrm{Intro}})$ be the typed graph as given in Figure 9.1. Since $\mathcal{G}$ is a k-th CL samplable game with the input distribution μ being a (k, m, p) CL distribution, by Definition 7.2.5, there exist two CL functions $\mathtt{L}^A, \mathtt{L}^B : \mathbb{F}_{2^p}^m \to \mathbb{F}_{2^p}^m$ over registers $\{V_j\}_{j \in [k]}$ which can be used to sample from μ. Furthermore, recall from the preliminaries that $U_{2 \to p}$ is the unitary map given in Lemma 8.2.1, and we write $U_{2 \to p}^m$ as a unitary acting on $\mathbb{C}^{2^{n^\alpha}}$ defined by $U_{2 \to p}^{\otimes m} \otimes \mathbb{I}_{2^{n^\alpha - m \cdot p}}$.

For the "sample complexity" clause in the theorem, since the "question content" specified in Figure 9.2 are empty except for the question labels from the n^α-Pauli basis game, in which the corresponding CL functions are already specified in the proof of Theorem 8.3.3; the distribution μ^{intro} is samplable via a $(\mathtt{T}_k^{\mathrm{Intro}}, \mathtt{E}_k^{\mathrm{Intro}}, \{\mathtt{L}^v\}_{v \in \mathtt{T}^{\mathrm{Intro}}})$ typed CL distribution as specified by the theorem statement. The "furthermore" part follows from Figure 9.2 depends only on n^α for

the Pauli basis test and k for the number of "Hide" question labels. This concludes the proof for the "Sample complexity" clause of the theorem.

For the "completeness" clause in the theorem, let $\mathscr{S}$ be a perfect oracularizable strategy in model t for $\mathscr{G}$. Since $\mathscr{G}$ is synchronous, $\mathscr{S}$ is synchronous and hence by Lemma 4.3.2, we can write $\mathscr{S} = (\mathscr{L}^2(\mathscr{A}, \tau), |\tau\rangle, \{A_a^x\})$ as a projective synchronous strategy.

Before defining the perfect strategy for $\mathscr{G}^{\text{Intro}}$, we first introduce some notations for some data processed Pauli measurements which is used as a part of the perfect strategy. For $P \in \{A, B\}$, we define the data processing measurement for the CL function $\mathtt{L}^P$ as

$$\rho^{p,Z}_{[\mathtt{L}^P|x]} = \sum_{a \in V | \mathtt{L}^P(a) = x} \rho^{p,Z}_a,$$

for all $x \in V$. By the definition of a CL function given in Definition 7.3.1, the measurement above is equivalent to the following: The prover first performs the data processing measurement $\{\rho^{p,Z}_{[\mathtt{L}^P_{0,0}|x_0]}\}_{x_0 \in V_0}$ to sample some $x_0 \in V_0$. Then, for $1 \le i < k$, the prover performs the measurement

$$\left\{\rho^{p,Z}_{[\mathtt{L}^P_{i,x_{<i}}|x_i]}\right\}_{x_i \in V_i} \tag{9.2}$$

to obtain measurement outcome x_i and compute $x_{<i+1} = x_i + x_{<i}$. The final measurement outcome is $x = x_{<k}$. For $j \in [k]$, we define the measurement operator

$$\left\{\rho^{p,Z}_{[\mathtt{L}^P_{<j-1}|x_{<j}]}\right\}_{x_{<j} \in V_{\le j}} \tag{9.3}$$

similarly to $\rho^{p,Z}_{[\mathtt{L}^P|x]}$, except that only the first j measurements from (9.2) are performed.

Recall from Section 7.1 that, for a linear map $\mathtt{L}$, we use $\mathtt{L}^\perp$ to denote the linear map that projects onto $\left(\ker(\mathtt{L})^\perp\right)^C$. By Lemma 8.2.2, for a fixed $x_{<j} \in V_{<j}$, the measurement operators $\left\{\rho^{p,Z}_{[\mathtt{L}^P_{i,x_{<i}}|x_i]}\right\}_{x_i \in V_i}$ pairwise commute with the measurement operator $\left\{\rho^{p,X}_{[(\mathtt{L}^P)^\perp_{i,x_{<i}}|x_i^\perp]}\right\}_{x_i^\perp \in V_i}$.

We define a perfect symmetric oracularizable strategy $\mathscr{S}^{\text{Intro}} = (\mathbb{C}^{2^{n^\alpha+1}} \otimes \mathbb{C}^{2^{n^\alpha+1}} \otimes \mathscr{H}, |\text{ME}_2\rangle^{\otimes(n^\alpha+1)} \otimes |\tau\rangle, \{M_a^x\})$ for $\mathscr{G}^{\text{Intro}}$ as follows: For the question labels $v \in \mathtt{T}_k^{\text{Intro}}$ which intersects a blue edge as specified in Figure 9.1. We define the measurement operator M_a^x

$$M_a^v = P_a^b \otimes \mathbb{I}_A$$

where $\mathscr{S}^{\text{PB}} = (\mathbb{C}^{2^{n^\alpha+1}} \otimes \mathbb{C}^{2^{n^\alpha+1}}, |\text{ME}_2\rangle^{\otimes(n^\alpha+1)}, \{P_a^x\})$ is the perfect oracularizable strategy for the n^α-Pauli basis test guaranteed by Theorem 8.3.3. Notably, for $t_W \in \{0,1\}^{n^\alpha}$

$$M^{(\text{Pauli},W)}_{t_W} = \rho^W_{t_W} \otimes \mathbb{I}_2 \otimes \mathbb{I}_{\mathscr{A}}.$$

Let $\mathbb{I}_R = \mathbb{I}_{2^{n^\alpha - p\cdot m+1}}$. We define the measurement for the rest of the question label as follows:

$$\begin{aligned}
M^{(\text{Gen Pauli},W)}_{t_W} &= U^m_{2\to p}(\rho^{p,W}_{s_W})(U^m_{2\to p})^* \otimes \mathbb{I}_R \otimes \mathbb{I}_{\mathscr{A}} && \text{for all } s_W \in \mathbb{F}^m_{2^p},\\
M^{(\text{Sample},\mathtt{L}^P)}_{(s,a)} &= U^m_{2\to p}(\rho^{p,Z}_s)(U^m_{2\to p})^* \otimes \mathbb{I}_R \otimes A^{\mathtt{L}^P(s)}_a && \text{for all } s \in \mathbb{F}^m_{2^p}, a \in \mathscr{A},\\
M^{(\text{Intro},\mathtt{L}^P)}_{(x,a)} &= U^m_{2\to p}(\rho^{p,Z}_{[\mathtt{L}^P|x]})(U^m_{2\to p})^* \otimes \mathbb{I}_R \otimes A^x_a && \text{for all } s \in \mathbb{F}^m_{2^p}, a \in \mathscr{A},
\end{aligned}$$

We define the measurement operator for $M^{(\text{Read},\mathtt{L}^P)}_{t,t^\perp,a}$ as follows: The prover first performs the measurement $U^m_{2\to p}\left(\rho^{p,W}_{[\mathtt{L}^P|t]}\right)(U^m_{2\to p})^* \otimes \mathbb{I}_R \otimes A^x_a$ (where the procedure for performing $\rho^{p,W}_{[\mathtt{L}^P|t]}$ is defined in (9.2)) and samples $(t,a) \in \mathscr{X} \times \mathscr{A}$. Then the prover performs the measurement

$$\left\{U^m_{2\to p}\left(\rho^{p,X}_{[(\mathtt{L}^P_x)|x^\perp]}\right)(U^m_{2\to p})^* \otimes \mathbb{I}_R \otimes \mathbb{I}_{\mathscr{A}}\right\}_{t^\perp \in V}.$$

Since these measurements commute, the measurement $M^{(\text{Read},\mathrm{L}^P)}_{t,t^\perp,a}$, defined as the product of the two measurements described, is a well-defined measurement.

For $i \in [k]$, the measurement operator for $M^{(\text{Hide}_i,\mathrm{L}^P)}_{t_{<i},t^\perp_{\le i},r_{>i}}$ is defined in a similar way. The prover first performs the measurement

$$\left\{ U^m_{2\to p}\left(\rho^{p,Z}_{[\mathrm{L}^P_{<j-1}|t_{<i}]}\right)(U^m_{2\to p})^* \otimes \mathbb{I}_R \otimes \mathbb{I}_{\mathscr{A}} \right\}_{t_{<i}\in V_{\le j}}$$

to sample $t_{<i} \in V_{<j}$. Then performs the measurement

$$\left\{ U^m_{2\to p}\left(\rho^{p,X}_{r_{>i}}\right)(U^m_{2\to p})^* \otimes \mathbb{I}_R \otimes \mathbb{I}_{\mathscr{A}} \right\}_{r_{>i}\in V_{>j}},$$

to sample $r_{>i} \in V_{>j}$. We remark that by the comment after Lemma 8.2.1, these two measurements commute. Lastly, the prover performs the measurement

$$\left\{ U^m_{2\to p}\left(\rho^{p,X}_{\left[\left(\mathrm{L}^P_{\le i,x_{<j}}\right)^\perp | x^\perp_{\le i}\right]}\right)(U^m_{2\to p})^* \otimes \mathbb{I}_R \otimes \mathbb{I}_{\mathscr{A}} \right\}_{x^\perp_{\le i}\in V_{\le i}} \tag{9.4}$$

to sample $t^\perp_{\le i} \in V_{\le j}$. This measurement commutes with the first measurement as proven above and commutes with the second measurement because both are generalized Pauli X measurements. Hence $M^{(\text{Hide}_i,\mathrm{L}^P)}_{t_{<i},t^\perp_{\le i},r_{>i}}$ defined as the product of the above three measurements is a well-defined measurement. For clarity, we write all the measurement operator on the table below.

First, the measurement M^ν_a are projective, as given any $\nu \in \mathrm{T}^{\text{Intro}}_k$, the measurements M^ν_a are define by products of projective measurements which all commute with each other. Since $M^\nu_a \in \bigotimes_{i\in[n^\alpha+1]} \mathscr{M}_2(\mathbb{C}) \otimes \mathscr{A}$, we have

$$\left(M^x_a \otimes \mathbb{I}_{2^{n^\alpha+1}}\right)|\mathrm{ME}_2\rangle^{n^\alpha+1}|\tau\rangle = \left(\mathbb{I}_{2^{n^\alpha+1}} \otimes (M^x_a)^{op}\right)|\mathrm{ME}_2\rangle^{n^\alpha+1}|\tau\rangle$$

where M^x_a above is acting on one registers of the entangled state $|\mathrm{ME}_2\rangle^{n^\alpha+1}$ and the state $|\tau\rangle$. Thus, the strategy $\mathrm{T}^{\text{Intro}}_k$ succeeds with probability 1 on the consistency equations. By Theorem 8.3.3, the strategy $\mathscr{S}^{\text{Intro}}_k$ is perfect and oracularizable when restricted to the question pair restricted to the blue edge (i.e. the n^α-Pauli basis test) within Figure 9.1. By Lemma 8.2.1, the strategy $\mathscr{S}^{\text{Intro}}_k$ is perfect and oracularizable when restricted to the question pair (Pauli, W) – (Gen Pauli, W). When restricted to the question pair (Intro, L^A) – (Intro, L^B), by construction, the question pair (t_A, t_B) sampled by the measurement operator of $\mathscr{S}^{\text{Intro}}$ precisely corresponds to the question distribution μ, as $(t_A, t_B) = (\mathrm{L}^A(s), \mathrm{L}^B(s))$ for some $s \in V$. Since $\mathscr{S}$ is a perfect oracularizable strategy for the game $\mathscr{G}$, $\mathscr{S}^{\text{Intro}}$ is also a perfect oracularizable strategy when restricted to the "Intro" question pair. It is straightforward to verify that $\mathscr{S}^{\text{Intro}}$ remains a perfect and oracularizable strategy for the remainder question pairs by the table above, concluding the proof for "completeness" part of the theorem. □

9.2 Soundness proof for the introspection protocol

The proof of the "soundness" clause for Proposition 7.4.2 follows a similar structure to [JNV+22a, Section 8.4]. We first establish some notations which we use in the proof. Let $(\mathscr{A}, \tau)$ be a tracial von Neumann algebra represented under the standard form $(\chi_\tau, \mathscr{L}^2(\mathscr{A}, \tau), |\tau\rangle)$. $\mathscr{M}_n(\mathbb{C}) \otimes \mathscr{A}$ is also a tracial von Neumann algebra with the trace Tr. In this case, the standard form for $\mathscr{M}_n(\mathbb{C}) \otimes \mathscr{A}$ are represented as $\mathscr{M}_n(\mathbb{C}) \otimes \mathbb{I}_n \otimes \mathscr{A}$ with the tracial state $|\mathrm{ME}_n\rangle \otimes |\tau\rangle$. In this case the opposite map op maps elements from $(\mathscr{M}_n(\mathbb{C}))_A \otimes (\mathbb{I}_n)_B \otimes \mathscr{A}$ to $(\mathbb{I}_n)_A \otimes (\mathscr{M}_n(\mathbb{C}))_B \otimes \mathscr{A}$. Hence, when discussing measurement operators $P_{A_1\mathscr{A}}$ from $\mathscr{S}''$, $\left(P_{A_1\mathscr{A}}\right)^{op}$ are defined in $B_1\mathscr{A}$ (where

Label	V_0	V_1	V_2	$\cdots$	V_{k-1}	$\mathscr{A}$
(Gen Pauli, X)	$(s_W)_0 \sim \sigma^X$	$(s_W)_1 \sim \sigma^X$	$(s_W)_2 \sim \sigma^X$		$(s_W)_{k-1} \sim \sigma^X$	$\mathbb{I}_{\mathscr{A}}$
(Hide$_0$, $\mathtt{L}^P$)	$t_0^{\perp} \sim \sigma^X_{(\mathtt{L}_0^P)^{\perp}}$	$r_1 \sim \sigma^X$	$r_2 \sim \sigma^X$		$r_{k-1} \sim \sigma^X$	$\mathbb{I}_{\mathscr{A}}$
(Hide$_1$, $\mathtt{L}^P$)	$t_0^{\text{Line}} \sim \sigma^Z_{\mathtt{L}_0^P}$ $t_0^{\perp} \sim \sigma^X_{(\mathtt{L}_0^P)^{\perp}}$	$t_1^{\perp} \sim \sigma^X_{(\mathtt{L}_1^P)^{\perp}}$	$r_2 \sim \sigma^X$	$\cdots$ $\cdots$	$r_{k-1} \sim \sigma^X$	$\mathbb{I}_{\mathscr{A}}$
(Hide$_2$, $\mathtt{L}^P$)	$t_0^{\text{Line}} \sim \sigma^Z_{\mathtt{L}_0^P}$ $t_0^{\perp} \sim \sigma^X_{(\mathtt{L}_0^P)^{\perp}}$	$t_1^{\text{Line}} \sim \sigma^Z_{\mathtt{L}_1^P}$ $t_1^{\perp} \sim \sigma^X_{(\mathtt{L}_1^P)^{\perp}}$	$t_2^{\perp} \sim \sigma^X_{(\mathtt{L}_2^P)^{\perp}}$	$\cdots$ $\cdots$	$r_{k-1} \sim \sigma^X$	$\mathbb{I}_{\mathscr{A}}$
			$\cdots$			
(Hide$_{k-1}$, $\mathtt{L}^P$)	$t_0^{\text{Line}} \sim \sigma^Z_{\mathtt{L}_0^P}$ $t_0^{\perp} \sim \sigma^X_{(\mathtt{L}_0^P)^{\perp}}$	$t_1^{\text{Line}} \sim \sigma^Z_{\mathtt{L}_1^P}$ $t_1^{\perp} \sim \sigma^X_{(\mathtt{L}_1^P)^{\perp}}$	$t_2^{\text{Line}} \sim \sigma^Z_{\mathtt{L}_2^P}$ $t_2^{\perp} \sim \sigma^X_{(\mathtt{L}_2^P)^{\perp}}$	$\cdots$ $\cdots$	$t_{k-1}^{\perp} \sim \sigma^X_{(\mathtt{L}_{k-1}^P)^{\perp}}$	$\mathbb{I}_{\mathscr{A}}$
(Read, $\mathtt{L}^P$)	$(t^{\text{Line}})_0 \sim \sigma^Z_{\mathtt{L}_0^P}$ $(t^{\perp})_0 \sim \sigma^X_{(\mathtt{L}_0^P)^{\perp}}$	$(t^{\text{Line}})_1 \sim \sigma^Z_{\mathtt{L}_1^P}$ $(t^{\perp})_1 \sim \sigma^X_{(\mathtt{L}_1^P)^{\perp}}$	$(t^{\text{Line}})_2 \sim \sigma^Z_{\mathtt{L}_2^P}$ $(t^{\perp})_2 \sim \sigma^X_{(\mathtt{L}_2^P)^{\perp}}$	$\cdots$ $\cdots$	$(t^{\text{Line}})_{k-1} \sim \sigma^Z_{\mathtt{L}_{k-1}^P}$ $(t^{\perp})_{k-1} \sim \sigma^X_{(\mathtt{L}_{k-1}^P)^{\perp}}$	$a \sim A_a^{t^{\text{Line}}}$
(Intro, $\mathtt{L}^P$)	$t_0 \sim \sigma^Z_{\mathtt{L}_0^P}$	$t_1 \sim \sigma^Z_{\mathtt{L}_1^P}$	$t_2 \sim \sigma^Z_{\mathtt{L}_2^P}$	$\cdots$	$t_{k-1} \sim \sigma^Z_{\mathtt{L}_{k-1}^P}$	$a \sim A_a^t$
(Sample, $\mathtt{L}^P$)	$s_0 \sim \sigma^Z$	$s_1 \sim \sigma^Z$	$s_2 \sim \sigma^Z$	$\cdots$	$s_{k-1} \sim \sigma^Z$	$a \sim A_a^{\mathtt{L}^P(s)}$
(Gen Pauli, Z)	$(s_Z)_0 \sim \sigma^Z$	$(s_Z)_1 \sim \sigma^Z$	$(s_Z)_2 \sim \sigma^Z$	$\cdots$	$(s_Z)_{k-1} \sim \sigma^Z$	$\mathbb{I}_{\mathscr{A}}$

Table 9.1: Summary of the measurement operator M_a^v, and as well as the output being sampled from each measurement operator. The notation $x \sim M$ are the variable x sampled from the measurement operator. For $i \in [k]$, the measurement $\sigma^Z_{\mathtt{L}_i^P}$ (resp.$\sigma^X_{(\mathtt{L}_i^P)^{\perp}}$) above are shorthand for $\sigma^Z_{[\mathtt{L}_{i,t_{<i}}^P|x]}$ (resp. $\sigma^X_{[(\mathtt{L}_{i,t_{<i}}^P)^{\perp}|x]}$) (where the v_i depends on the previous measurement outcome). We also omit the conjugation by $U^m_{2\to p}$ for clarity.

the $\mathscr{A}$ register is defined within the Hilbert space $\mathscr{L}^2(\mathscr{A},\tau)$ and contains measurements from both $\mathscr{A}$ and $\mathscr{A}'$). We sometimes write $P^{op}_{B_1\mathscr{A}}$ for a measurement operator $P_{A_1\mathscr{A}}$ to specified the registers.

For a canonical register subspace $V \subseteq \mathbb{F}_{2^p}^m$, we define $\mathbb{C}_V \subseteq (\mathbb{C}^{2^p})^{\otimes m}$ as the subspace span by the basis $\{|v\rangle\}_{v\in V}$, and $\mathbb{I}_V = \sum_{v\in V} |v_0,\cdots v_{m-1}\rangle\langle v_0,\cdots v_{m-1}| \in \mathscr{M}_{2^{p\cdot m}}(\mathbb{C})$. Furthermore, for $A,B \in \mathscr{B}$ for some von Neumann algebra $\mathscr{B}$, we write $[A,B] = A\cdot B - B\cdot A$. For a multi-outcome measurement $\{P_{a,b}\}_{a\in\mathscr{A},b\in\mathscr{B}}$, we denote $P_a = \sum_{b\in\mathscr{B}} P_{a,b}$. For a function $\mathbf{f}\colon \mathscr{A}\to\mathscr{C}$, we also use $P_{[\mathbf{f}(a)|b]}$ to denote the data process measurement being applied on P_a. For two sets of POVM $\{P_{a_1,b}\}_{a_1\in\mathscr{A},b\in\mathscr{B}}$, $\{Q_{c,a_2}\}_{c\in\mathscr{B},a_2\in\mathscr{A}}$, we write $P_{a_1} \approx Q_{a_2=a_1}$ to emphasize the outcome variables for the measurement outcome which follows the $\approx$ relationship.

9.2.1 Preliminary lemmas

Before continuing with the proof, we begin by recalling the following important lemma from [JNV+22a].

Lemma 9.2.1 (Pauli twirl decomposition, Lemma 8.15 of [JNV+22a])**.** *Let $\mathscr{X},\mathscr{A}$ be two finite sets, μ be a distribution over $\mathscr{X}$, V be a canonical subspace of $\mathbb{F}_{2^p}^m$, and $(\mathscr{A},\tau)$ be a tracial von Neumann algebra represented in the standard form $(\chi_\tau,\mathscr{L}^2(\mathscr{A},\tau),|\tau\rangle)$. For each $x\in\mathscr{X}$, let W_x and V_x be two canonical subspace of V such that $W_x\subseteq V_x\subseteq V$, and let $\mathtt{L}_x : W_x \to W_x$ be a linear map.*

Consider the state $|\psi\rangle = |ME_{2^p}\rangle^{\otimes m}_{A_1A_2} \otimes |Aux\rangle_{\mathscr{A}} \in \left(\mathbb{C}^{2^p} \otimes \mathbb{C}^{2^p}\right)^{\otimes m}_{A_1B_1} \otimes \mathscr{L}^2(\mathscr{A},\tau)$ *where* $|Aux\rangle_{\mathscr{A}}$ *is an arbitrary vector. For each* $x \in \mathscr{X}$*, let* $\{M^x_{t,a}\}_{t\in W_x, a\in\mathscr{A}}$ *be a set of PVM on* $\mathscr{B}(\mathbb{C}_{V_x}) \otimes \mathscr{A}$*. Suppose that the following condition holds for some* $\varepsilon > 0$

$$\begin{aligned} M^x_t \otimes \mathbb{I}_{V^C_x} &\approx_\varepsilon \rho^{p,Z}_{[\mathrm{L}_x|t]} \otimes \mathbb{I}_{W^C_x} \otimes \mathbb{I}_{\mathscr{A}} \\ \left[(M^x_{t,a} \otimes \mathbb{I}_{V^C_x}), (\rho^{p,Z}_z \otimes \mathbb{I}_{W^C_x} \otimes \mathbb{I}_{\mathscr{A}})\right] &\approx_\varepsilon 0 \\ \left[(M^x_{t,a} \otimes \mathbb{I}_{V^C_x}), (\rho^{p,X}_{[\mathrm{L}^\perp_x|t^\perp]} \otimes \mathbb{I}_{W^C_x} \otimes \mathbb{I}_{\mathscr{A}})\right] &\approx_\varepsilon 0, \end{aligned}$$

where the $\approx$ *is defined over the distribution* μ *and summing over* $t, t^\perp \in W_x$ *and* $a \in \mathscr{A}$*. In the equation above, both* $\rho^{p,X}_{[\mathrm{L}^\perp_x|t^\perp]}$ *and* $\rho^{p,Z}_{[\mathrm{L}_x|t]}$ *are in* $\mathscr{B}(\mathscr{H}_{V_x\setminus W_x}) \otimes \mathscr{A}$*.*

Then for each $x \in \mathscr{X}$ *and* $t \in W_x$*, there exists a set of POVM* $\{M^{x,t}_a\}_{a\in\mathscr{A}}$ *acting on* $\mathscr{B}(\mathbb{C}_{V_x\setminus U_x}) \otimes \mathscr{A}$ *such that on average over* $x \sim \mu$

$$(M^x_{t,a} \otimes \mathbb{I}_{V^C_x}) \approx_{O(\mathrm{poly}(\varepsilon))} \rho^{p,Z}_{[\mathrm{L}_x|r]} \otimes M^{x,t}_a \otimes \mathbb{I}_{V^C_x}.$$

We also recall several lemmas from [JNV+22a, Section 8.4.2], which are useful for decomposing PVM measurements between different Hilbert spaces, as well as showing commutation relationships between measurements.

Lemma 9.2.2 (Decomposition of measurements over the $\approx$ distance, Lemma 8.16 of [JNV+22a])**.** *Let* $\mathscr{A}$ *and* $\mathscr{B}$ *be two finite sets,* $\varepsilon > 0$*,* $|\psi_Q\rangle \in \mathscr{H}_Q$ *and* $|\psi_A\rangle \in \mathscr{H}_A$*. Furthermore, let* $\{Q_a\} \subseteq \mathscr{B}(\mathscr{H}_Q)$ *be a set of PVM and for all* $a \in \mathscr{A}$*, and let* $\{A^a_b\}_{b\in\mathscr{B}}, \{B^a_b\}_{b\in\mathscr{B}} \subseteq \mathscr{B}(\mathscr{H}_A)$ *be two sets of POVMs. Then the following are equivalent:*

- $(Q_a \otimes A^a_b) \approx_\varepsilon (Q_a \otimes B^a_b)$ *on the state* $|\psi_Q\rangle \otimes |\psi_A\rangle$*.*
- *Over the distribution* $P(a) = \langle\psi_Q|Q_a|\psi_Q\rangle$ *and the state* $|\psi_A\rangle$*, we have* $A^a_b \approx_\varepsilon B^a_b$*.*

Lemma 9.2.3 (Lemma 8.18 of [JNV+22a])**.** *Let* $\mathscr{X}$*,* $\mathscr{Y}$ *and* $\mathscr{Z}$ *be three finite sets, and* $(\mathscr{A},\tau)$ *be a tracial von Neumann algebra represented in the standard form* $(\chi_\tau, \mathscr{L}^2(\mathscr{A},\tau), |\tau\rangle)$*. Furthermore, for all* $x \in \mathscr{X}$*, let* $y \in \mathscr{Y}$*, let* $\{A^y_{x,z}\} \subseteq \mathscr{A}$ *be a set of POVM,* $\{B_{x,y,z}\} \subseteq \mathscr{A}'$ *be a set of PVM Suppose that*

$$\sum_{x,y,z} \langle\psi|A^y_{x,z}B_{x,y,z}|\psi\rangle \geq 1 - \varepsilon,$$

for some state $|\psi\rangle \in \mathscr{L}^2(\mathscr{A},\tau)$ *and* $\varepsilon > 0$*. Then with respect to the state* $|\psi\rangle$*,* $B_{x,y,z} \approx_\varepsilon A^y_{x,z}B_{x,y}$*.*

Lemma 9.2.4 (Approximation relationship implies commutation, Lemma 5.25 of [JNV+22a])**.** *Let* $\mathscr{A}$ *be a von Neumann algebra, let* $\mathscr{X}$*,* $\mathscr{A}$*,* $\mathscr{B}$*, and* $\mathscr{C}$ *be four finite sets. For every* $x \in \mathscr{X}$*, let* $\{A^x_{a,b}\}_{a\in\mathscr{A},b\in\mathscr{B}}, \{C^x_{a,c}\}_{a\in\mathscr{A},c\in\mathscr{C}} \subseteq \mathscr{A}$ *be two sets of POVM and let* $\{B^x_{a,b,c}\}_{a\in\mathscr{A},b\in\mathscr{B},c\in\mathscr{C}} \subseteq \mathscr{A}'$ *be a set of PVM. Suppose that for some* $\delta > 0$

$$A^x_{a,b} \approx_\delta B^x_{a,b} \qquad C^x_{a,c} \approx_\delta B^x_{a,c}.$$

Then $[A^x_{a,b}, C^x_{a,c}] \approx_\delta 0$*.*

Lemma 9.2.5 (Decomposition measurements preserves approximate commutation over the $\approx$ distance, Lemma 8.16 of [JNV+22a])**.** *Let* $\mathscr{A}$*,* $\mathscr{B}$*, and* $\mathscr{C}$ *be three finite sets,* $\varepsilon > 0$*,* $|\psi_Q\rangle \in \mathscr{H}_Q$ *and* $|\psi_A\rangle \in \mathscr{H}_A$*. Furthermore, let* $\{Q_a\} \subseteq \mathscr{B}(\mathscr{H}_Q)$ *be a set of PVM, let* $\{A^a_b\}_{b\in\mathscr{B}}, \{B^a_b\}_{b\in\mathscr{B}} \subseteq \mathscr{B}(\mathscr{H}_A)$ *be two sets of POVMs, and let* $A_{a,b} = Q_a \otimes A^a_b$*,* $B_{a,c} = Q_a \otimes A^a_c$*.*

Suppose that $[A_{a,b}, B_{a,c}] \approx_\varepsilon 0$ *with respect to the state* $|\psi_Q\rangle \otimes |\psi_A\rangle$*. Then*

$$[A^a_b, B^a_c] \simeq_\varepsilon 0$$

where the $\approx$ *and* $\simeq$ *is defined over the distribution* $P(a) = \langle\psi_Q|Q_a|\psi_Q\rangle$*, the state* $|\psi_A\rangle$*, and summing over* $(b,c) \in \mathscr{B} \times \mathscr{C}$*.*

We remark that although the lemmas in this subsection are originally defined for finite-dimensional matrices, the proof can be trivially modified for the infinite-dimensional setting.

9.2.2 The proof

Recall, given the original game $\mathcal{G}$, the input distribution μ is described by a (k,m,p) CL distribution which is defined over two k-th level CL function $\mathtt{L}^A, \mathtt{L}^B : \mathbb{F}_{2^p}^m \to \mathbb{F}_{2^p}^m$ over registers $\{V_j\}_{j\in[k]}$, $\mathbb{F}_{2^p}^m = V = \bigoplus_{j\in[k]} V_j$. For the introspection transformation $\mathcal{G}^{\text{Intro}} = (\mathcal{X}^{\text{Intro}}, \mathcal{A}^{\text{Intro}}, \mu^{\text{Intro}}, D^{\text{Intro}})$ of $\mathcal{G}$, we call the question $x \in \mathcal{X}^{\text{Intro}}$ as an introspection question if the question label for x corresponds to a vertex which intersects with either a orange or green edge (i.e. all the vertices on the right side of "(Pauli, X)" and "(Pauli, Z)").

For "soundness", suppose that $\omega^t(\mathcal{G}^{\text{Intro}}) > 1-\varepsilon$, we wish to show that $\omega^t(\mathcal{G}) > 1-O(\text{poly}(\exp(k),\varepsilon))$. Let $\mathcal{S} = (\mathcal{L}^2(\mathcal{A},\tau), \sigma|\tau\rangle, \{A_a^x\}, \{(B_a^x)^{op}\})$ be a tracially embeddable strategy in model t with $\omega(\mathcal{G}^{\text{Intro}}, \mathcal{S}) > 1-\varepsilon$. We start the proof by first showing the following claim, this is an analogue of [JNV+22a, Lemma 8.19]

Lemma 9.2.6. *There is a symmetric strategy*

$$\mathcal{S}' = (\mathbb{C}^{2^{n^\alpha}}_{A_1} \otimes \mathbb{C}^{2^{n^\alpha}}_{B_1} \otimes \mathbb{C}^{2^{n^\alpha}}_{A_2} \otimes \mathbb{C}^{2^{n^\alpha}}_{B_2} \otimes \mathcal{L}^2(\mathcal{A},\tau), |ME_2\rangle^{\otimes n^\alpha}_{A_1B_1} \otimes |Aux\rangle_{A_2B_2\mathcal{A}}, \{\hat{P}_a^x\})$$

for some vector state $|aux\rangle_{A_2B_2\mathcal{A}} \in \mathbb{C}^{2^{n^\alpha}}_{A_2} \otimes \mathbb{C}^{2^{n^\alpha}}_{B_2} \otimes \mathcal{L}^2(\mathcal{A},\tau)$. *Furthermore,* $\omega(\mathcal{G}^{Intro}, \mathcal{S}') > 1 - \text{poly}(n,\varepsilon)$ *, and for* $W \in \{X,Z\}$

$$P_u^{(Pauli,\, W)} = \rho_u^W. \tag{9.5}$$

Proof. Consider $\mathcal{S}$ when restricted to the n^α-Pauli basis test (the blue vertices in Figure 9.1). Since by sampling a random question pair, there is a $O(k)$ probability that a question from the Pauli basis test is selected. This implies that $\mathcal{S}$ succeeds on the n^α-Pauli basis test with probability at least $1-O(k\cdot\varepsilon)$. By Theorem 8.3.3, there exist two isometries V_A, V_B with $(V_B \otimes \mathbb{I}_{2^{2\cdot n^\alpha}})V_A = V_A(V_B \otimes \mathbb{I}_{2^{2\cdot n^\alpha}})$; a state $|\text{Aux}\rangle_{A_2B_2\mathcal{A}} \in \mathbb{C}^{2^{2n^\alpha}} \otimes \mathcal{L}^2(\mathcal{A},\tau)$ such that

$$\left\| \left(V_B \otimes \mathbb{I}_{2^{2n^\alpha}}\right) V_A(\sigma|\tau\rangle) - |\text{ME}_2\rangle^{\otimes n^\alpha} |\text{Aux}\rangle \right\|^2 \leq O\left(\text{poly}(k,\varepsilon)\right), \tag{9.6}$$

and for all $W \in \{X,Z\}$ and $u \in \mathbb{F}_{2^{n^\alpha}}$

$$\begin{aligned}
&\| ((V_A A_u^{\text{Pauli, W}} V_A^*)_{A_1A_2\mathcal{A}} \otimes (\mathbb{I}_{2^{2n^\alpha}})_{B_1B_2} - \\
&\quad (\rho_u^W)_{A_1} \otimes (\mathbb{I}_{2^{n^\alpha}})_{A_2} \otimes (\mathbb{I}_{2^{2n^\alpha}})_{B_1B_2} \otimes (\mathbb{I}_{\mathcal{H}})_{\mathcal{A}}) |\text{Aux}\rangle_{\mathcal{A}A_2B_2} |\text{ME}_2\rangle^{\otimes n}_{A_1B_1} \|^2 \leq O\left(\text{poly}(\varepsilon)\right), \qquad (9.7) \\
&\| \left(\left(V_B (B_u^{(\text{Pauli, W})})^{\text{op}} V_B^* \right)_{\mathcal{A}B_1B_2} \otimes (\mathbb{I}_{2^{2n^\alpha}})_{A_1A_2} - \right. \\
&\quad \left. (\rho_u^W)_{B_1} \otimes (\mathbb{I}_{2^{n^\alpha}})_{B_2} \otimes (\mathbb{I}_{2^{2n^\alpha}})_{A_1A_2} \otimes (\mathbb{I}_{\mathcal{H}})_{\mathcal{A}} \right) |\text{Aux}\rangle_{\mathcal{A}A_2B_2} |\text{ME}_2\rangle^{\otimes n^\alpha}_{A_1B_1} \|^2 \leq O\left(\text{poly}(\varepsilon)\right). \qquad (9.8)
\end{aligned}$$

For each $(x,a) \in \mathcal{X}^{\text{Intro}} \times \mathcal{A}^{\text{Intro}}$, we define $\hat{A}_a^x = V_A A_a^x V_A^*$, and likewise $\hat{B}_a^x = V_B B_a^x V_B^*$. Define $\mathcal{S}_1$ as a strategy which uses the state $|\text{EPR}\rangle^{\otimes n^\alpha}_{A_1B_1} |\text{Aux}\rangle_{A_2B_2\mathcal{A}}$ and the measurement operator $\hat{A}_a^x$ and $\hat{B}_a^x$ for all questions instead. By (9.6), $\mathcal{S}_1$ succeeds in $\mathcal{G}^{\text{Intro}}$ with probability $1-O(\text{poly}(k,\varepsilon))$. By the description given in Table 9.1. Since $\mathcal{G}^{\text{Intro}}$ is $O(k)$-balance, this implies that $\mathcal{S}_1$ is $O(\text{poly}(k,\varepsilon))$-synchronous; hence by Corollary 4.3.6, there exists a projective, symmetric strategy

$$\mathcal{S}_2 = (\mathbb{C}^{2^{n^\alpha}}_{A_1} \otimes \mathbb{C}^{2^{n^\alpha}}_{B_1} \otimes \mathbb{C}^{2^{n^\alpha}}_{A_2} \otimes \mathbb{C}^{2^{n^\alpha}}_{B_2} \otimes \mathcal{L}^2(\mathcal{A},\tau), |\text{EPR}\rangle^{\otimes n^\alpha}_{A_1B_1} |\text{Aux}\rangle_{A_2B_2\mathcal{A}}, \{\hat{P}_a^x\})$$

such that $\hat{A}_a^x \approx_{O(\delta)} \hat{P}_a^x$ with $\omega(\mathcal{S}_2, \mathcal{G}^{\text{Intro}}) > 1 - O(\text{poly}(n))$.

Define $\mathcal{S}'$ as the same measurement operator as $\mathcal{S}_2$, except for the question label (Pauli, W) where instead the Pauli measurements $(\rho_u^W)_{A_1}$ (resp. $(\rho_u^W)_{B_1}$) are used instead. The lemma then follows by combining Equation (9.7) and $\hat{A}_a^x \approx_{O(\delta)} \hat{P}_x^a$. □

We wish to transform the underlying state of $\mathcal{S}'$ in a way that the underlying entangled state is $|\text{ME}_{2^p}\rangle^{\otimes m}$ instead, consider the strategy $\mathcal{S}''$, which is defined on the same Hilbert space as $\mathcal{S}'$ except that the under lying state is

$$\left((U^m_{2\to p})_{A_1} \otimes (U^m_{2\to p})_{B_1} \otimes \mathbb{I}_{A_2B_2\mathcal{A}} \right) |\text{ME}_2\rangle^{\otimes n^\alpha}_{A_1B_1} \otimes |\text{Aux}\rangle_{A_2B_2\mathcal{A}}$$

and the measurement operator is defined as $P_a^x = (U_{2\to p}^m \otimes \mathbb{I}_{2^{n^\alpha}} \otimes \mathbb{I}_{\mathscr{A}})^* \hat{P}_a^x (U_{2\to p}^m \otimes \mathbb{I}_{2^{n^\alpha}} \otimes \mathbb{I}_{\mathscr{A}})$. Since $U_{2\to p}^m$ is a unitary, $\omega(\mathscr{S}', \mathscr{G}^{\text{Intro}}) = \omega(\mathscr{S}'', \mathscr{G}^{\text{Intro}})$. By Lemma 8.2.1, we can rewrite the state in $\mathscr{S}''$ as

$$\left(|\text{ME}_{2^p}\rangle^{\otimes m} \otimes |\text{ME}_2\rangle^{\otimes n^\alpha - p\cdot m}\right)_{A_1B_1} \otimes |\text{Aux}\rangle_{A_2B_2\mathscr{A}}$$

and $P_a^{\text{Pauli, W}} = (\rho^{p,W}_{\kappa^{-1}(\pi_{\le p\cdot k}(a))})_{A_1} \otimes (\rho^W_{\pi_{>p\cdot k}(a)} \otimes \mathbb{I})_{\mathscr{A}}$ with $(P_a^{\text{Pauli, W}})^{op} = (\rho^{p,W}_{\kappa^{-1}(\pi_{\le p\cdot k}(a))})_{B_1} \otimes (\rho^W_{\pi_{>p\cdot k}(a)})_{\mathscr{A}}$. Since the question pair (Pauli, W) – (Gen Pauli, W) occurs with probability $O(\frac{1}{k})$,

$$(\rho_s^{p,W} \otimes \rho^W_{\pi_{>p\cdot k}(a)}) \simeq_{O(\text{poly}(k,\varepsilon))} P_s^{(\text{Gen Pauli, W})},$$

and since $\rho^W_{\pi_{>p\cdot k}(a)})$ is a set of PVM, by summing over ρ^W and apply Lemma 3.3.3,

$$\rho_s^{p,W} \approx_{O(\text{poly}(k,\varepsilon))} P_s^{(\text{Gen Pauli, W})}.$$

For simplicity of notation, we rewrite $\mathscr{S}'$ as follows. We shrink the registers A_1 and B_1 to include only the first $|\text{ME}_{2^p}\rangle^{\otimes m}$ pair, and combine the remaining parts of A_1 and B_1, as well as the registers A_2 and B_2, into the "$\mathscr{A}$" infinite-dimensional register. Hence, we can write

$$\mathscr{S}'' = \left(\mathbb{C}^{2^{p\cdot m}}_{A_1} \otimes \mathbb{C}^{2^{p\cdot m}}_{B_1} \otimes \mathscr{L}^2(\mathscr{A}, \tau), |\psi\rangle = |\text{ME}_{2^p}\rangle^{\otimes m}_{A_1B_1} \otimes |\text{Aux}\rangle_{\mathscr{A}}, \{P_a^x\}\right), \tag{9.9}$$

such that $\omega(\mathscr{G}^{\text{Intro}}, \mathscr{S}'') > 1 - O(\text{poly}(n)) = \delta_1$, with

$$\rho_s^{p,W} \approx_{O(\text{poly}(k,\varepsilon))} P_s^{(\text{Gen Pauli, W})} \tag{9.10}$$

for $W \in \{X, Z\}$. In this case, $\mathscr{A}$ also includes the extra finite-dimensional registers A_2B_2, and the extra EPR pair guaranteed by the soundness clause of Theorem 8.3.3. Our goal is construct a strategy for $\mathscr{G}$ which succeed with probability $1 - O(\text{poly}(\varepsilon))$ using the measurement $P^{\text{Intro},\text{L}^A}$ and $P^{\text{Intro},\text{L}^B}$. Unless otherwise stated, the $\approx$ and $\simeq$ relationship in this subsection are define over the state $|\psi\rangle$ used to define the strategy $\mathscr{S}''$

For every $s \in \mathbb{F}_{2^p}^m$, we partition $s = \sum_{j\in[k]} s_j$ where each $s_j \in V_j$, and we write

$$s_{<j} = \sum_{i\in[j]} s_i, \qquad s_{\ge j} = \sum_{j\le i<k} s_i. \tag{9.11}$$

Furthermore, for $W \in \{X, Z\}$ and $s_j \in V_j$, we use the notation $\rho_{s_j}^{p,W} = \sum_{t\in\mathbb{F}_{2^p}^m | t_j = s_j} \rho_{s_j}^{p,W}$, and we define $\rho_{s_{<j}}^{p,W}$ (resp. $\rho_{s_{\ge j}}^{p,W}$) in a similar manner for $s_{<j} \in V_{<j}$ (resp. $s_{\ge j} \in V_{\ge j}$). Since $\rho^{p,W}$ is projective, we have $\rho_s^{p,W} = \Pi_{i\in[k]}\rho_{s_i}^{p,W}$ by definition. Since V_j is a canonical basis subspace,

$$\rho_s^{p,W} = \rho^{p,W}_{s_0,\cdots,s_{k-1}} = \bigotimes_{i\in[k]} \rho_{s_i}^{p,W} \tag{9.12}$$

where each $s_i \in V_i$. When decomposing the generalized Pauli measurement in this manner, we often times write it as $\bigotimes_{i\in[k]} \left(\rho_{s_i}^{p,W}\right)_{V_i} \in \bigotimes_{i\in[k]} \mathbb{C}_{V_i} = \mathbb{C}^{2^{p\cdot m}}$ to emphasize the underlying Hilbert space. For a canonical basis subspace $V \subseteq \mathbb{F}_{2^p}^m$, we write as the (normalized) state $\left(|\text{ME}_{2^p}\rangle^{\otimes m}_{A_1B_1}\right)_V := \sum_{x\in V} |x\rangle \otimes |x\rangle$.

For $j \in [k]$, $s_{\le j}, t_{\le j} \in V_{\le j}$ and $r_{<j} \in V_{<j}$, and $P \in \{A, B\}$, define

$$\left(\rho^{p,Z}_{[\text{L}^P_{\le j}(s_{\le j})|t_{\le j}]}\right)_{V_{\le j}} := \left(\rho^{p,Z}_{[\text{L}^P_{0,0}(s_{\le j})_0|(t_{\le j})_0]}\right)_{V_0} \otimes \left(\bigotimes_{1\le i\le j} \left(\rho^{p,Z}_{\left[\text{L}^P_{i,(t_{\le j})_{<i-1}}((s_{\le j})_i)|(t_{\le j})_i\right]}\right)_{V_j}\right),$$

$$\left(\rho^{p,X}_{[(\text{L}^P)^\perp_{\le j,r_{<j}}(s_{\le j})|t_{\le j}]}\right)_{V_{\le j}} := \left(\rho^{p,X}_{\left[\left(\text{L}^P_{0,0}\right)^\perp(s_{\le j})_0|(t_{\le j})_0\right]}\right)_{V_0} \otimes \left(\bigotimes_{1\le i<j} \left(\rho^{p,X}_{\left[\left(\text{L}^P_{i,(r_{<j})_{<i-1}}\right)^\perp((s_{\le j})_i)|(t_{\le j})_i\right]}\right)_{V_j}\right),$$

where

$$s_{\leq j} = \sum_{i\in[j]} \left(s_{\leq j}\right)_i \in \bigoplus_{i\in[j]} V_i,\ t_{\leq j} = \sum_{i\in[j]} \left(t_{\leq j}\right)_i \in \bigoplus_{i\in[j]} V_i, \text{ and } r_{<j} = \sum_{i\in[j-1]} \left(r_{<j}\right)_i \in \bigoplus_{i\in[j-1]} V_i.$$

By definition, for $s,t \in \mathbb{F}_{2^p}^m$, $\left(\rho^{p,Z}_{[\mathrm{L}^P_{<k}(s)|t]}\right) = \left(\rho^{p,Z}_{[\mathrm{L}^P(s)|t]}\right)$. Base on the synchronicity condition of Figure 9.3, we have the following claim.

Claim 9.2.7 (Consistency of measurement output)**.** *For all $x \in \mathcal{X}^{Intro}$ where x is an introspection question, $P^x_a \approx_{O(\mathrm{poly}(k,\delta_1))} (P^x_a)^{op}$ over the state $|ME_{2^p}\rangle^{\otimes m}_{A_1A_2} \otimes |Aux\rangle_{\mathscr{A}}$.*

Proof. Fix an introspection question x, given a question pair sampled from μ^{Intro}, there is a $O(\frac{1}{k^2})$ probability that the sampled question pair is (x,x). Since $\mathscr{S}''$ is a strategy for $\mathscr{G}^{\mathrm{Intro}}$ which succeed with probability $1-\delta_1$, this implies that $P^x_a \simeq_{O(\mathrm{poly}(k,\delta_1))} (P^x_a)^{op}$ by the "synchronicity" clause given by Figure 9.3. The claim then follows from Lemma 3.3.3 to convert between $\simeq$ distance to $\approx$ distance. □

Based on the verification procedure given in Figure 9.3, we have the following claim about the approximation related to the Pauli X measurement

Claim 9.2.8 (Approximation of strategies related to the Pauli X measurement)**.** *Let $P \in \{A,B\}$ and $0 < j < k$, then*

$$P^{Hide_0,\mathrm{L}^P}_{t^{\perp}_{\leq 0},r_{>0}} \approx_{O(k,\delta_1)} \left(\rho^{p,X}_{[\left(\mathrm{L}^P_{0,0}\right)^{\perp}(s_0)|t^{\perp}_{\leq 0}]}\right)_{V_0} \otimes \left(\rho^{p,X}_{s_{>0}=r_{>0}}\right)_{V_{>0}} \otimes \mathbb{I}_{\mathscr{A}}, \tag{9.13}$$

where we partition the measurement outcome for $\rho^{p,X}$ as $\sum_{i\in[k]} s_i \in \bigoplus_{i\in[k]} V_i$ and $s_{>0} \in V_{>0}$ in the above equation.

Proof. Fix $P \in \{A,B\}$, since the question pair (Gen Pauli, X) – (Hide_0, L^P) is sampled with probability $O(k)$ from the distribution μ^{Intro}, and $\mathscr{S}''$ is a strategy for $\mathscr{G}^{\mathrm{Intro}}$ which succeed with probability $1-\delta_1$, combining with (9.10)

$$P^{\mathrm{Hide}_0,\mathrm{L}}_{(t^{\perp}_{\leq 0},r_{>0})} \simeq_{O(k,\delta_1)} \left(\rho^{p,X}_{s_0=t^{\perp}_{\leq 0},s_{>0}=r_{>0}}\right)^{op},$$

where $s = s_0 + s_0^C + s_{>0} = \ker \mathrm{L}^P_{0,0} \oplus \ker \mathrm{L}^{P\ C}_{0,0} \oplus V_{>0}$ according to the verification procedure from Figure 9.3. By applying Lemma 3.3.3, Claim 9.2.7 and the triangle inequality for $\approx$ distance,

$$P^{\mathrm{Hide}_0,\mathrm{L}}_{(t^{\perp}_{\leq 0},r_{>0})} \approx_{O(k,\delta_1)} \rho^{p,X}_{s_0=t^{\perp}_{\leq 0},s_{>0}=r_{>0}} \otimes \mathbb{I}_{\mathscr{A}}.$$

Finally, by the definition of $\left(\mathrm{L}^P_{0,0}\right)^{\perp}$, and $\rho^{p,X}$ is projective, we obtain (9.13). □

Similarly, we have the following claim about the approximation about the Pauli Z measurement.

Claim 9.2.9 (Approximation of strategies related to the Pauli Z measurement)**.** *For every $P \in \{A,B\}$ and $1 \leq j < k$ the following hold*

$$P^{Sample,\mathrm{L}^P}_{s_{sample}} \approx_{O(\mathrm{poly}(k,\delta_1))} \rho^{p,Z}_{s_{sample}} \otimes \mathbb{I}_{\mathscr{A}}, \tag{9.14}$$

$$P^{Intro,\mathrm{L}^P}_{x_P,a_P} \approx_{O(\mathrm{poly}(k,\delta_1))} \rho^{p,Z}_{[\mathrm{L}^P(s)|x_P]} \otimes \mathbb{I}_{\mathscr{A}}, \tag{9.15}$$

$$P^{Read,\mathrm{L}^P}_{t^{Line}_{Read}} \approx_{O(\mathrm{poly}(k,\delta_1))} \rho^{p,Z}_{[\mathrm{L}^P(s)|t^{Line}_{Read}]} \otimes \mathbb{I}_{\mathscr{A}}, \tag{9.16}$$

$$P^{Hide_j,\mathrm{L}^P}_{t^{Line}_{<j}} \approx_{O(\mathrm{poly}(k,\delta_1))} \left(\rho^{p,Z}_{[\mathrm{L}^P_{<j-1}(s_{<j})|t^{Line}_{<j}]}\right)_{V_{<j}} \otimes \mathbb{I}_{V_{\geq j}} \otimes \mathbb{I}_{\mathscr{A}}. \tag{9.17}$$

Proof. Since, the proof for each approximation follows a similar structure as (9.13) from Claim 9.2.8, we only give a rough sketch for the proof for each of the equation below.

- Equation (9.14) follows from the verification procedure for (Gen Pauli, Z) – (Sample, $\mathtt{L}^P$) question pair from Figure 9.3.
- Equation (9.15) follows from the verification procedure for (Sample, $\mathtt{L}^P$) – (Intro, $\mathtt{L}^P$) question pair, and applying the triangle inequality of $\approx$ distance to Equation (9.14).
- Equation (9.16) follows from the verification procedure for (Read, $\mathtt{L}^P$) – (Intro, $\mathtt{L}^P$) question pair, and applying the triangle inequality of $\approx$ distance to Equation (9.15).
- For Equation (9.17) in the case where $j = k-1$. The equation follows from the verification procedure for (Hide$_{k-1}$, $\mathtt{L}^P$) – (Read, $\mathtt{L}^P$) question pair.
- For Equation (9.17) in the case where $0 < j < k-1$. The equation follows from an inductive proof using j = k-1 as the base case, and the inductive step follows from the verification procedure for (Hide$_i$, $\mathtt{L}^P$) – (Hide$_{i+1}$, $\mathtt{L}^P$) and Lemma 9.2.3. □

Based on the commutation with the Pauli-X and Pauli-Z measurement, we conclude the following approximation relationship related to the $P^{\text{Hide}_j,\mathtt{L}^P}$ and $P^{\text{Read},\mathtt{L}^P}$. The proof of the below claim is almost identical to [JNV+22a, Lemma 8.22, Lemma 8.23], except we use Claim 9.2.7 to swap the measurement operator instead.

Claim 9.2.10. *For $P \in \{A, B\}$ and all $j \in [k]$,*

$$P^{Hide_j,\mathtt{L}^P}_{t^{Line}_{<j},t^{\perp}_{\leq j},r_{>j}} \approx_{O(\text{poly}(k,\delta_1))} \left(\rho^{p,X}_{[(\mathtt{L}^P)^{\perp}_{<j,t^{Line}_{<j}}(s_{\leq j})|t^{\perp}_{\leq j}]} \cdot \rho^{p,Z}_{[\mathtt{L}^P_{<j-1}(s_{<j})|t^{Line}_{<j}]}\right)_{V_{\leq j}} \otimes \left(\rho^{p,X}_{s_{>j}=r_{>j}}\right)_{V_{>j}} \otimes \mathbb{I}_{\mathscr{A}}, \tag{9.18}$$

$$P^{Read,\mathtt{L}^P}_{t^{\perp}_{Read,P},t^{Line}_{Read,P}} \approx_{O(\text{poly}(k,\delta_1))} \left(\rho^{p,X}_{\left[(\mathtt{L}^P)^{\perp}_{\leq k,t^{Line}_{Read,P}}(s)|t^{\perp}_{Read,P}\right]} \cdot \rho^{p,Z}_{[\mathtt{L}^P(s)|t^{Line}_{Read}]}\right)_{V} \otimes \mathbb{I}_{\mathscr{A}}. \tag{9.19}$$

Proof. We start by showing Equation (9.18) by an inductive argument. For the case where $j = 0$, this precisely follows from Equation (9.13). For the inductive step, fix $1 \leq i < k$ and assume Equation (9.18) holds for all $0 \leq j \leq i$, we wish to show Equation (9.18) for $i+1$. Since the question pair (Hide$_i$, $\mathtt{L}^P$) –(Hide$_{i+1}$, $\mathtt{L}^P$) in $\mathscr{G}^{\text{Intro}}$ are selected with probability $O(\frac{1}{k})$, by the "Hiding test" verification procedure given in Figure 9.3

$$\sum_{\substack{(v_{\leq i},v^{\perp}_{\leq i+1},u_{>i+1}) \\ \in V_{<i+1}\times V_{\leq i+1}\times V_{>i+1}}} \langle\psi|P^{\text{Hide}_i,\mathtt{L}^P}_{\begin{pmatrix} t^{\text{line}}_{<i}=v_{<i}, \\ t^{\perp}_{\leq i}=v^{\perp}_{\leq i}, \\ \left[\mathtt{L}^{\perp}_{i+1,v_{\leq i}}(r_i)|v^{\perp}_{i+1},\right]. \\ r_{>i+1}=u_{>i+1} \end{pmatrix}} \cdot \left(P^{\text{Hide}_{i+1}\mathtt{L}^P}_{v_{\leq i},v^{\perp}_{\leq i+1},u_{>i+1}}\right)^{op} |\psi\rangle \geq 1 - O(\text{poly}(k,\delta_1))$$

where we partition the variables according to the convention that $v_{\leq i} = v_{<i} + v_i \in V_{<i} \oplus V_i$ and likewise with the other variables in the sum. We wish to apply Lemma 9.2.3 where the "A" measurement is $P^{\text{Hide}_i,\mathtt{L}^P}$, the "B" measurement is $P^{\text{Hide}_{i+1},\mathtt{L}^P}$, the "x" variable is $v_{<i}$, the "y" variable is v_i and the "z" variable are $(v^{\perp}_{\leq i+1}, u_{>i+1})$. By this formulation, for $u_i \in V_i$, we can rewrite the "A" measurement as

$$P^{\text{Hide}_i,\mathtt{L}^P,u_i}_{v_{<i},v^{\perp}_{\leq i},\left[\mathtt{L}^{\perp}_{i+1,v_{\leq i}}(u_i)|v^{\perp}_{i+1},\right],u_{>i+1}} \tag{9.20}$$

using the summation for the measurement outcome above, where in this case, we divide up the output $r_{>i} = r_i + r_{>i+1}$ and write it as the third and fourth output of $P^{\text{Hide}_i,\mathtt{L}^P,v_{i+1}}$. By the

inductive hypothesis,

$$
\begin{aligned}
&P^{\mathrm{Hide}_i,\mathrm{L}^P,v_{i+1}}_{v_{<i},v^{\perp}_{\leq i},\left[\mathrm{L}^{\perp}_{i+1,v_{\leq i}}(u_i)|v^{\perp}_{i+1},\right],u_{>i+1}}\\
&\approx_{O(\mathrm{poly}(k,\delta_1))}\left(\rho^{p,X}_{[(\mathrm{L}^P)^{\perp}_{<i,u_{<i}}(s_{\leq i})|u^{\perp}_{\leq i}]}\cdot\rho^{p,Z}_{[\mathrm{L}^P_{<i-1}(s_{<i})|t_{<i}]}\right)_{V_{\leq i}}\\
&\qquad\otimes\left(\rho^{p,X}_{\left[\mathrm{L}^{\perp}_{i+1,v_{\leq i}}(u_i)|v^{\perp}_{i+1},\right]}\right)_{V_{i+1}}\otimes\left(\rho^{p,X}_{s_{>i+1}=r_{>i+1}}\right)_{V_{>i+1}}\otimes\mathbb{I}_{\mathscr{A}}. \qquad (9.21)
\end{aligned}
$$

Similarly, since $P^{\mathrm{Hide}_{i+1},\mathrm{L}^P}$ is projective, we can write the corresponding $B_{x,y}$ as $P^{\mathrm{Hide}_{i+1},\mathrm{L}^P}_{v_{<i},v_i} = P^{\mathrm{Hide}_{i+1},\mathrm{L}^P}_{v_{<i+1}}$. Hence

$$
\begin{aligned}
P^{\mathrm{Hide}_{i+1}\mathrm{L}^P}_{t^{\mathrm{Line}}_{<i+1},t^{\perp}_{\leq i+1},r_{>i+1}} &\approx_{O(\mathrm{poly}(k,\delta_1))}\left(P^{\mathrm{Hide}_{i+1}\mathrm{L}^P}_{t^{\mathrm{Line}}_{<i+1},t^{\perp}_{\leq i+1},r_{>i+1}}\right)^{op}\\
&\approx_{O(\mathrm{poly}(k,\delta_1))} P^{\mathrm{Hide}_i,\mathrm{L}^P,u_i}_{t^{\mathrm{Line}}_{<i},t^{\perp}_{\leq i},\left[\mathrm{L}^{\perp}_{i+1,t^{\mathrm{Line}}_{<i+1}}(r_i)|t^{\perp}_{i+1},\right],r_{>i+1}}\cdot\left(P^{\mathrm{Hide}_{i+1},\mathrm{L}^P}_{t^{\mathrm{Line}}_{<i+1}}\right)^{op}\\
&\approx_{O(\mathrm{poly}(k,\delta_1))}\left(\rho^{p,X}_{[(\mathrm{L}^P)^{\perp}_{<i,u_{<i}}(s_{\leq i})|u^{\perp}_{\leq i}]}\cdot\rho^{p,Z}_{[\mathrm{L}^P_{<i-1}(s_{<i})|t_{<i}]}\right)_{V_{\leq i}}\\
&\qquad\otimes\left(\rho^{p,X}_{\left[\mathrm{L}^{\perp}_{i+1,v_{\leq i}}(u_i)|v^{\perp}_{i+1},\right]}\right)_{V_{i+1}}\otimes\left(\rho^{p,X}_{s_{>i+1}=r_{>i+1}}\right)_{V_{>i+1}}\otimes\mathbb{I}_{\mathscr{A}}\cdot\left(P^{\mathrm{Hide}_{i+1},\mathrm{L}^P}_{t^{\mathrm{Line}}_{<i+1}}\right)^{op}\\
&\approx_{O(\mathrm{poly}(k,\delta_1))}\left(\rho^{p,X}_{[(\mathrm{L}^P)^{\perp}_{<i,u_{<i}}(s_{\leq i})|u^{\perp}_{\leq i}]}\cdot\rho^{p,Z}_{[\mathrm{L}^P_{<i-1}(s_{<i})|t_{<i}]}\right)_{V_{\leq i}}\cdot\left(\rho^{p,Z}_{[\mathrm{L}^P_{<i}(s_{<i+1})|t^{\mathrm{Line}}_{<i+1}]}\right)_{V_{\leq i}}\\
&\qquad\otimes\left(\rho^{p,X}_{\left[\mathrm{L}^{\perp}_{i+1,v_{\leq i}}(u_i)|v^{\perp}_{i+1},\right]}\right)_{V_{i+1}}\otimes\left(\rho^{p,X}_{s_{>i+1}=r_{>i+1}}\right)_{V_{>i+1}}\otimes\mathbb{I}_{\mathscr{A}}\\
&=\left(\rho^{p,X}_{[(\mathrm{L}^P)^{\perp}_{<i+1,t^{\mathrm{Line}}_{<i+1}}(s_{\leq i+1})|t^{\perp}_{\leq i+1}]}\cdot\rho^{p,Z}_{[\mathrm{L}^P_{<i}(s_{<i+1})|t^{\mathrm{Line}}_{<i+1}]}\right)_{V_{\leq i}}\otimes\left(\rho^{p,X}_{s_{>i}=r_{>i+1}}\right)_{V_{>i}}\otimes\mathbb{I}_{\mathscr{A}}
\end{aligned}
$$

where the first inequality follows from Claim 9.2.7 and Lemma 3.3.3, the second line follows from Lemma 9.2.3 labelled above. Line 3 follows from applying Lemma 3.3.4 along with (9.21). Line 4 follows from applying Lemma 3.3.4 along with (9.17) and Claim 9.2.7. The last line follows from the definition of L^P and $\left(\mathrm{L}^P\right)^{\perp}$. This shows the case for $i+1$, which show (9.18). (9.19) follows from a similar argument using the question pair $(\mathrm{Hide}_{k-1},\mathrm{L}^P)$ – $(\mathrm{Read},\mathrm{L}^P)$. □

Finally, we show that the measurement $P^{\mathrm{Intro},\mathrm{L}^P}$ can be decomposed as a tensor product of a "question sampling" PVM using the Generalized Pauli Z measurement and a "game strategy" PVM. We remark that this is an analogue for [JNV+22a, Lemma 8.24] for the commuting operator model and the proof follows a similar structure.

Claim 9.2.11. *Fix $P\in\{A,B\}$. For every $j\in[k+1]$ and $\left(x_p\right)_{<j}\in V_{<j}$, there exist a set of POVM $\left\{\left(P^{Intro,\mathrm{L}^P,(x_p)_{<j}}_{x_{\geq j},a_P}\right)_{V_{\geq j}\mathscr{A}}\right\}_{x_{\geq j}\in V_{\geq j},a_P\in\mathscr{A}}\in\mathscr{B}(\mathbb{C}_{V_{\geq j}})\otimes\mathscr{A}$ such that*

$$
\left(P^{Intro\mathrm{L}^P}_{(x_P)_{<j},(x_P)_{\geq j},a_P}\right)_{V\mathscr{A}}\approx_{O(\mathrm{poly}(k,\delta_1)^{1/2^j})}\left(\rho^{p,Z}_{[\mathrm{L}^P_{<j}(s_{<j})|(x_P)_{<j}]}\right)_{V_{<j}}\otimes\left(P^{Intro,\mathrm{L}^P,(x_p)_{<j}}_{(x_P)_{\geq j},a_P}\right)_{V_{\geq j}\mathscr{A}},
$$

where the summation are over $((x_P)_{<i},(x_P)_{\geq i},a_P)\in V_{<i}\times V_{\geq i}\times\mathscr{A}$.

Proof. We show this claim via induction. For $k=0$, the claim trivially follows by setting

$$
P^{\mathrm{Intro}\mathrm{L}^P,0}_{(x_P)_{<i},(x_P)_{\geq i},a_P}=P^{\mathrm{Intro},\mathrm{L}^P}_{(x_P)_{<i}+(x_P)_{\geq i},a_P}.
$$

For the inductive step, fix $1 \le i < k+1$ and assume Equation (9.18) holds for all $0 \le j \le i$, we wish to show the claim for $i+1$. For $(x_P)_{<i}$, let $\left(P^{\text{Intro}\mathtt{L}^P}_{(x_P)_{<j},(x_P)_{\ge j},a_P}\right)_{V\mathscr{A}}$ be the PVM guaranteed by the inductive hypothesis. Let $x \sim \mathtt{L}^P_{<i}$ denote the distribution where s is first sampled uniformly randomly from $V_{<i}$, and the first i-th levels of $\mathtt{L}^P$ are then applied to s. The goal is to use Lemma 9.2.1 in order to construct the measurement require for the lemma. To do so, we show that on average over $(x_P)_{<i} \sim \mathtt{L}^P_{<i}$, the following set of equations hold:

$$\left(P^{\text{Intro},\mathtt{L}^P,(x_p)_{<i}}_{(x_P)_i}\right)_{V_{\ge i}\mathscr{A}} \approx_{O(\text{poly}(k,\delta_1)^{1/2^j})} \left(\rho^{p,Z}_{\left[\mathtt{L}^P_{i,(x_p)<i}(s_i)|(x_P)_i\right]}\right)_{V_i} \otimes \mathbb{I}_{V_{>i}} \otimes \mathbb{I}_{\mathscr{A}} \tag{9.22}$$

$$= \left[\left(P^{\text{Intro},\mathtt{L}^P,(x_p)_{<i}}_{(x_P)_{\ge i},a_P}\right)_{V_{\ge i}\mathscr{A}}, \left(\rho^{p,Z}_z\right)_{V_i} \otimes \mathbb{I}_{V_i^C} \otimes \mathbb{I}_{\mathscr{A}})\right] \approx_{O(\text{poly}(k,\delta_1)^{1/2^j})} 0 \tag{9.23}$$

$$\left[\left(P^{\text{Intro},\mathtt{L}^P,(x_p)_{<i}}_{(x_P)_{\ge i},a_P}\right)_{V_{\ge i}\mathscr{A}}, \left(\rho^{p,X}_{\left[\left(\mathtt{L}^P_{i,(x_p)<i}\right)^{\perp}(s_i)|(x_P)^{\perp}\right]}\right)_{V_i} \otimes \mathbb{I}_{V_{>i}} \otimes \mathbb{I}_{\mathscr{A}})\right] \approx_{O(\text{poly}(k,\delta_1)^{1/2^j})} 0, \tag{9.24}$$

where we partition the output $(x_P)_{\le i}$ from $P^{\text{Intro},\mathtt{L}^P,(x_p)_{<i}}$ as $(x_P)_{<i} + (x_P)_i \in V_{<i} \oplus V_i$. In this case, the "x" variable from Lemma 9.2.1 correspond to $(x_p)_{<i}$, the "t" variable corresponds to $(x_P)_i$ and the "a" variable corresponds to $((x_P)_{>i}, a_P)$.

For (9.22), by Equation (9.16) when restricted to the output to $(x_P)_{<i+1}$,

$$\begin{aligned}\left(P^{\text{Intro}\mathtt{L}^P}_{(x_P)_{<i+1}}\right)_{V\mathscr{A}} &\approx_{O(\text{poly}(k,\delta_1)^{1/2^j})} \left(\rho^{p,Z}_{[\mathtt{L}^P_{<i+1}|(x_P)_{<i+1}]}\right)_{V_{<i+1}} \otimes \mathbb{I}_{V_{>i}} \otimes \mathbb{I}_{\mathscr{A}},\\ &= \left(\rho^{p,Z}_{[\mathtt{L}^P_{<i}|(x_P)_{<i}]}\right)_{V_{<i}} \otimes \left(\rho^{p,Z}_{\left[\mathtt{L}^P_{i,(x_p)<i}|(x_P)_i\right]}\right)_{V_i} \otimes \mathbb{I}_{V_{>i}} \otimes \mathbb{I}_{\mathscr{A}}\end{aligned}$$

where the second equality follows from the definition of $\mathtt{L}^P$. (9.22) then follows from Lemma 9.2.4.

For (9.23). By using the fact that $\mathscr{S}''$ succeed with probability $1-\delta_1$, the fact that the question pair (Gen Pauli, W) – (Sample, $\mathtt{L}^P$) is sampled with probability $O(k)$ from the distribution μ^{Intro}, and Lemma 3.3.3, (9.10) and the triangle inequality of $\approx$ distance, we have

$$\left(P^{\text{Sample}\mathtt{L}^P}_{s_{<i+1}}\right)_{V\mathscr{A}} \approx_{O(\text{poly}(k,\delta_1)^{1/2^j})} \left(\rho^{p,Z}_{s_{<i+1}}\right)_{V_{<i+1}} \otimes \mathbb{I}_{V_{\ge i+1}} \otimes \mathbb{I}_{\mathscr{A}}. \tag{9.25}$$

By applying the same argument with the (Sample, $\mathtt{L}^P$) – (Intro, $\mathtt{L}^P$) along with the inductive hypothesis

$$\left(P^{\text{Sample}\mathtt{L}^P}_{a_P}\right) \approx_{O(\text{poly}(k,\delta_1)^{1/2^j})} \left(\rho^{p,Z}_{[\mathtt{L}^P_{<i}(s_{<i})|(x_P)_{<i}]}\right)_{V_{<i}} \otimes \left(P^{\text{Intro},\mathtt{L}^P,(x_p)_{<i}}_{a_P}\right)_{V_{\ge i}\mathscr{A}}. \tag{9.26}$$

Hence, by Lemma 9.2.4

$$[\left(\rho^{p,Z}_{s_{<i+1}}\right)_{V_{<i+1}} \otimes \mathbb{I}_{V_{\ge i+1}} \otimes \mathbb{I}_{\mathscr{A}}, \left(\rho^{p,Z}_{[\mathtt{L}^P_{<i}(s_{<i})|(x_P)_{<i}]}\right)_{V_{<i}} \otimes \left(P^{\text{Intro},\mathtt{L}^P,(x_p)_{<i}}_{a_P}\right)_{V_{\ge i}\mathscr{A}}] \approx_{O(\text{poly}(k,\delta_1)^{1/2^j})} 0.$$

Finally, since the underlying state are $|\text{ME}_{2^p}\rangle^{\otimes m}$ on the registers V, (9.23) follows from Lemma 9.2.5. For (9.24), by restricting the output from Equation (9.19)

$$P^{\text{Read},\mathtt{L}^P}_{\left(t^{\perp}_{\text{Read},P}\right)_i,\left(t^{\text{Line}}_{\text{Read},P}\right)_{<i}} \approx_{O(\text{poly}(k,\delta_1)^{1/2^j})} \left(\rho^{p,Z}_{\left[\mathtt{L}^P_{<i}(s_{<i})|\left(t^{\text{Line}}_{\text{Read},P}\right)_{<i}\right]}\right)_{V_{<j}} \otimes \left(\rho^{p,X}_{\left[\left(\mathtt{L}^P_{i,\left(t^{\text{Line}}_{\text{Read},P}\right)_{<i}}\right)^{\perp}(s_i)|\left(t^{\perp}_{\text{Read},P}\right)_i\right]}\right)_{V_i}.$$

Similarly, by considering the (Read, $\mathtt{L}^P$) – (Intro, $\mathtt{L}^P$) along with the inductive hypothesis

$$\left(P^{\text{Read}\mathtt{L}^P}_{a_P}\right) \approx_{O(\text{poly}(k,\delta_1)^{1/2^j})} \left(\rho^{p,Z}_{[\mathtt{L}^P_{<i}(s_{<i})|(x_P)_{<i}]}\right)_{V_{<i}} \otimes \left(P^{\text{Intro},\mathtt{L}^P,(x_p)_{<i}}_{a_P}\right)_{V_{\ge i}\mathscr{A}}. \tag{9.27}$$

(9.24) then follows from Lemma 9.2.4 to obtain the commutation relationship, followed by Lemma 9.2.5 on the register $V_{<i}$.

By Lemma 9.2.1, by taking the "t" variable as $(x_P)_i \in V_i$ and the "a" variable as $((x_P)_{>i+1}, a_P) \in V_{>i+1} \times \mathscr{A}$, for every $(x_p)_{<i} + (x_P)_i = (x_p)_{<i+1} \in V_{i+1}$, there exists a set of POVM measurement $\left(P^{\text{Intro},\widehat{\mathrm{L}^P,(x_p)_{<i+1}}}_{(x_P)_{>i+1},a_P}\right)_{V_{>i}\mathscr{A}}$ such that, on expectation over $(x_p)_{<i} \sim \mathrm{L}^P_{<i}$

$$\left(P^{\text{Intro},\mathrm{L}^P,(x_p)_{<i}}_{(x_P)_{\geq i},a_P}\right)_{V_{\geq i}\mathscr{A}} \approx_{O(\text{poly}(\varepsilon))} \left(\rho^{p,Z}_{\left[\mathrm{L}^P_{i,(x_p)_{<i}}|(x_P)_i\right]}\right)_{V_i} \otimes \left(P^{\text{Intro},\widehat{\mathrm{L}^P,(x_p)_{<i+1}}}_{(x_P)_{\geq i+1},a_P}\right)_{V_{>i}\mathscr{A}}.$$

Since $\langle ME_{2^P}|^{\otimes m}_{V_i}\left(\left(\rho^{p,Z}_{[\mathrm{L}^P_{<i}(s_{<i})|(x_P)_{<i}]}\right)_{V_{<i}} \otimes \mathbb{I}_V\right)_{A_1B_1} |\mathrm{ME}_{2^P}\rangle^{\otimes m}_{V_i}$ precisely describes the distribution $(x_p)_{<i} \sim \mathrm{L}^P_{<i}$, by Lemma 9.2.2

$$\begin{aligned}
&\left(\rho^{p,Z}_{[\mathrm{L}^P_{<i}(s_{<i})|(x_P)_{<i}]}\right)_{V_{<i}} \otimes \left(P^{\text{Intro},\mathrm{L}^P,(x_p)_{<i}}_{(x_P)_{\geq i},a_P}\right)_{V_{\geq i}\mathscr{A}} \\
&\approx_{O(\text{poly}(\varepsilon))} \left(\rho^{p,Z}_{[\mathrm{L}^P_{<i}(s_{<i})|(x_P)_{<i}]}\right)_{V_{<i}} \otimes \left(\rho^{p,Z}_{\left[\mathrm{L}^P_{i,(x_p)_{<i}}(s_i)|(x_P)_i\right]}\right)_{V_i} \otimes \left(P^{\text{Intro},\widehat{\mathrm{L}^P,(x_p)_{<i+1}}}_{(x_P)_{\geq i+1},a_P}\right)_{V_{>i}\mathscr{A}} \\
&= \left(\rho^{p,Z}_{[\mathrm{L}^P_{<i+1}(s_{<i+1})|(x_P)_{<i+1}]}\right)_{V_{<i+1}} \otimes \left(P^{\text{Intro},\widehat{\mathrm{L}^P,(x_p)_{<i+1}}}_{(x_P)_{\geq i+1},a_P}\right)_{V_{>i}\mathscr{A}},
\end{aligned}$$

where the last line follows from the definition of L^P. Thus, combining with the inductive hypothesis,

$$\left(P^{\text{Intro}\,\mathrm{L}^P}_{(x_P)_{<i+1},(x_P)_{\geq i+1},a_P}\right)_{V\mathscr{A}} \approx_{O(\text{poly}(k,\delta_1)^{1/2^j})} \left(\rho^{p,Z}_{[\mathrm{L}^P_{<i+1}(s_{<i+1})|(x_P)_{<i+1}]}\right)_{V_{<i+1}} \otimes \left(P^{\text{Intro},\widehat{\mathrm{L}^P,(x_p)_{<i+1}}}_{(x_P)_{\geq i+1},a_P}\right)_{V_{\geq i+1}\mathscr{A}}, \tag{9.28}$$

Finally, we wish to replace each of the $\left(P^{\text{Intro},\widehat{\mathrm{L}^P,(x_p)_{<i+1}}}_{(x_P)_{\geq i+1},a_P}\right)_{V_{\geq i+1}\mathscr{A}}$ with a set of PVM via Lemma 3.3.1. To do so, we show the following lemma

Lemma 9.2.12 (Approximation of almost projective measurements)**.** *Let $\mathscr{A}$ be finite sets, and $\mathscr{A} \subseteq \mathscr{B}(\mathscr{H})$ be a von Neumann algebra. Let $\{A_a\}_a \subseteq \mathscr{A}$ be a set of PVM and $\{B_a\}_a \subseteq \mathscr{A}$ be a set of POVM such that*

$$A_a \approx_\varepsilon B_a$$

for some state $|\psi\rangle \in \mathscr{H}$ and some $\varepsilon > 0$. Then $\langle\psi|B_a^2|\psi\rangle \geq 1 - O(\sqrt{\varepsilon})$.

Proof. Since $\{A_a^x\}$ is a set of PVM, by Lemma 3.3.3, $A_a \approx_{\sqrt{\varepsilon}} B_a$ over $|\psi\rangle$ and $\varepsilon > 0$. By definition, we have $\sum_a \langle\psi|A_aB_a|\psi\rangle \geq 1 - \sqrt{\varepsilon}$, or

$$\sqrt{\varepsilon} \geq 1 - \sum_a \langle\psi|A_aB_a|\psi\rangle \geq 1 - \sqrt{\sum_a \langle\psi|A_a^2|\psi\rangle} \cdot \sqrt{\sum_a \langle\psi|B_a^2|\psi\rangle} = 1 - \sqrt{\sum_a \langle\psi|B_a^2|\psi\rangle}$$

where the last equality follows from $\{A_a\}_{a\in\mathscr{A}}$ being a set of PVM. This implies that $\sum_a \langle\psi|B_a^2|\psi\rangle \geq (1-\sqrt{\varepsilon})^2 = 1 - O(\sqrt{\varepsilon})$, as desired. □

Applying the above lemma to (9.28) along with Lemma 9.2.2, this implies that on expectation over $(x_P)_{<i+1} \sim \mathrm{L}^P_{<i+1}$, we have $\left(P^{\text{Intro},\widehat{\mathrm{L}^P,(x_p)_{<i+1}}}_{(x_P)_{\geq i+1},a_P}\right)^2_{V_{\geq i+1}\mathscr{A}} \approx_{O(\text{poly}(k,\delta_1))^{1/2^{j+1}}} 0$. Hence, by Lemma 3.3.1, there exist sets of PVM $\left\{\left(P^{\text{Intro},\mathrm{L}^P,(x_p)_{<i+1}}_{(x_P)_{\geq i+1},a_P}\right)_{V_{\geq i+1}\mathscr{A}}\right\}_{(x_P)_{\geq i+1}\in V_{\geq i+1},\, a_P\in\mathscr{A}}$ indexed by $(x_p)_{<i+1} \in V_{<i+1}$ such that, on expectation over $(x_P)_{<i+1} \sim \mathrm{L}^P_{<i+1}$,

$$\left(P^{\text{Intro}\,\mathrm{L}^P}_{(x_P)_{<i+1},(x_P)_{\geq i+1},a_P}\right)_{V\mathscr{A}} \approx_{O(\text{poly}(k,\delta_1))^{1/2^{j+1}}} \left(P^{\text{Intro},\mathrm{L}^P,(x_p)_{<i+1}}_{(x_P)_{\geq i+1},a_P}\right)_{V_{\geq i+1}\mathscr{A}}.$$

By applying Lemma 9.2.2 again, and the triangle inequality of $\approx$ distance applied to (9.28),

$$\left(P^{\text{Intro}\,\mathtt{L}^P}_{(x_P)_{<i+1},(x_P)_{\geq i+1},a_P}\right)_{V\mathscr{A}} \approx_{O(\text{poly}(k,\delta_1))^{1/2^{j+1}}} \left(\rho^{p,Z}_{[\mathtt{L}^P_{<i+1}(s_{<i+1})|(x_P)_{<i+1}]}\right)_{V_{<i+1}} \otimes \left(P^{\text{Intro},\mathtt{L}^P,(x_p)_{<i+1}}_{(x_P)_{\geq i+1},a_P}\right)_{V_{\geq i+1}\mathscr{A}}.$$

This shows the claim for $i+1$, thus concluding the proof. □

We remark that the dependency of $\frac{1}{1/2^j}$ in the above lemma arises from using Lemma 3.3.1 in order to force the POVM guaranteed by Lemma 9.2.1 to be PVMs. Since k is assumed to be a constant in this thesis and $j \in [k+1]$, this power dependency does not change the result of this thesis.

By Claim 9.2.11 for the case where $j = k$ and $P = A$, for every $x_P \in \mathbb{F}_{2^p}^m$, there exist a set of PVM $\left\{\left(P^{\text{Intro},\mathtt{L}^P,x_A}_{a_A}\right)_{\mathscr{A}}\right\}_{a_A\in\mathscr{A}} \subseteq \mathscr{A}$ such that

$$\left(P^{\text{Intro}\,\mathtt{L}^A}_{x_A,a_A}\right)_{A_1\mathscr{A}} \approx_{O(\text{poly}(k,\delta_1)^{1/2^k})} \left(\rho^{p,Z}_{[\mathtt{L}^A|(x_A)]}\right)_{A_1} \otimes \left(P^{\text{Intro},\mathtt{L}^P,x_A}_{a_A}\right)_{\mathscr{A}} \tag{9.29}$$

and by the same argument as Claim 9.2.11 applied to $\left(P^{\text{Intro}\,\mathtt{L}^P}_{x_P,a_P}\right)^{op}_{B_1\mathscr{A}} \subseteq \mathscr{B}(\mathbb{C}_V)\otimes\mathscr{A}'$ and take the case where $j = k$ and $P = B$, for every $x_P \in \mathbb{F}_{2^p}^m$, there exist a set of PVM $\left\{\left(Q^{\text{Intro},\mathtt{L}^P,x_B}_{a_B}\right)_{\mathscr{A}}\right\}_{a_B\in\mathscr{A}} \subseteq \mathscr{A}$ such that

$$\left(P^{\text{Intro}\,\mathtt{L}^B}_{x_B,a_B}\right)^{op}_{B_1\mathscr{A}} \approx_{O(\text{poly}(k,\delta_1)^{1/2^k})} \left(\rho^{p,Z}_{[\mathtt{L}^B|(x_B)]}\right)_{B_1} \otimes \left(Q^{\text{Intro},\mathtt{L}^P,x_B}_{a_B}\right)^{op}_{\mathscr{A}}. \tag{9.30}$$

Since the question pair (Intro, $\mathtt{L}^A$) – (Intro, $\mathtt{L}^B$) is sampled with probability $O(k)$ from the distribution μ^{Intro}. This means that whenever the question/answer pair (x_A, x_B, a_A, a_B) are sampled by the measurement $\langle\psi|\left(P^{\text{Intro}\,\mathtt{L}^A}_{x_A,a_A}\right)_{A_1\mathscr{A}}\left(P^{\text{Intro}\,\mathtt{L}^B}_{x_B,a_B}\right)^{op}_{B_1\mathscr{A}}|\psi\rangle$, by the "Introspection of $\mathscr{G}$" question clause, from Figure 9.3, the question/answer pair succeed in $\mathscr{G}$ (i.e. $D(x_A, x_B, a_A, b_B) = 1$) with probability at least $1 - O(\text{poly}(k,\delta_1))$. Combining with (9.29), (9.30) and the definition of $|\psi\rangle$ from Equation (9.9), this shows that the question/answer pair sampled by

$$\left(\langle \text{ME}_{2^p}|^{\otimes m}\left(\rho^{p,Z}_{[\mathtt{L}^A|(x_A)]}\right)_{A_1} \otimes \left(\rho^{p,Z}_{[\mathtt{L}^B|(x_B)]}\right)_{B_1}|\text{ME}_{2^p}\rangle^{\otimes m}\right)_{A_1B_1}$$
$$\otimes\left(\langle \text{Aux}|\left(P^{\text{Intro},\mathtt{L}^P,x_A}_{a_A}\right)\cdot\left(Q^{\text{Intro},\mathtt{L}^P,x_B}_{a_B}\right)^{op}\text{Aux}\rangle\right)_{\mathscr{A}}$$

succeed on $\mathscr{G}$ with probability at least $1-O(\text{poly}(k,\delta_1))-O(\text{poly}(k,\delta_1)^{1/2^k}) = 1-O(\text{poly}(k,\delta_1)^{1/2^k})$. To conclude the proof, we see that (x_A, x_B) sampled from the measurement

$$\left(\langle \text{ME}_{2^p}|^{\otimes m}\left(\rho^{p,Z}_{[\mathtt{L}^A|(x_A)]}\right)_{A_1} \otimes \left(\rho^{p,Z}_{[\mathtt{L}^B|(x_B)]}\right)_{B_1}|\text{ME}_{2^p}\rangle^{\otimes m}\right)_{A_1B_1}$$

are equal to $(x_A, x_B) \sim \mu$. This shows that the strategy

$$\mathscr{S}^{\mathscr{G}} = \left(\mathscr{L}^2(\mathscr{A},\tau), |\text{Aux}\rangle_{\mathscr{A}}, \{P^{\text{Intro},\mathtt{L}^P,x_A}_{a_A}\}, \{Q^{\text{Intro},\mathtt{L}^P,x_B}_{a_B}\}\right), \tag{9.31}$$

satisfies $\omega(\mathscr{G},\mathscr{S}^{\mathscr{G}}) > 1 - O(\text{poly}(k,\delta_1)^{1/2^k}) = 1 - O(\text{poly}(\exp k,\varepsilon))$ by the initial definition of δ_1. This shows the "soundness" clause for Theorem 9.1.1.

9.3 Proof of Proposition 7.4.2

In this section, we give a proof for Proposition 7.4.2. We restate the theorem below for convenience.

Proposition 9.3.1 (Question Reduction)**.** *For all constants $\alpha, k \in \mathbb{N}$, there exists a polynomial time algorithm $\texttt{QuestionReduction}_{\alpha,k}$ that takes, as input, a pair of Turing machines $(\mathtt{Q},\mathtt{D})$ and outputs a tuple of Turing machine machines $(\mathtt{Q}^{QR},\mathtt{D}^{QR})$ such that the following holds:*

For models $t \in \{*, co\}$, $\texttt{QuestionReduction}_{\alpha,k}$ outputs a pair of Turing machine $(\mathtt{Q}^{QR},\mathtt{D}^{QR})$ which is a fourth-level CL-verifier $\mathscr{V}^{QR}$ for a infinite sequence of game $\mathscr{G}^{QR} = \{\mathscr{G}_n^{QR}\}_{n\in\mathbb{N}}$ with the following properties: There exist an integer $\gamma^{QR} = O(poly(\alpha))$ such that

1. *(Computation time):* $\mathsf{TIME}_{\texttt{QuestionReduction}_{\alpha,k}}(\mathrm{poly}(\alpha), |\mathtt{Q}|, |\mathtt{D}|)$.
2. *(Synchronicity) The game sequence $\mathscr{V}^{QR}$ is a synchronous game sequences.*
3. *(Complexity bounds for the output):*
   - $\mathtt{Q}^{QR}(n, Parameter) \le \max\left\{\log^{\gamma^{QR}}(n), C^{trivial}\right\}$,
   - $\mathsf{TIME}_{\mathtt{Q}^{QR}}(n) \le \max\left\{\log^{\gamma^{QR}}(n), C^{trivial}\right\}$,
   - $\mathsf{TIME}_{\mathtt{D}^{QR}} \le \max\left\{n^{\gamma^{QR}}, C^{trivial}\right\}$,

   for some universal constant $C^{trivial}$.
4. *(Independency) $\mathtt{Q}^{QR}$ is a 3-th level CL sampler which only depends on the parameter α and k and does not depend on $|\mathtt{D}|$, $|\mathtt{D}^{QR}| = O(\mathrm{poly}(\alpha, k))$.*

Furthermore if the input $\mathscr{V} = (\mathtt{Q},\mathtt{D})$ is a k'th CL verifier for the infinite sequence of synchronous game $\mathscr{G} = \{\mathscr{G}_n\}_{n\in\mathbb{N}}$ for some constant $k' \in \mathbb{N}$ such that $k' < k$, and there exist some constant $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$

$$\max\left\{\mathsf{TIME}_{\mathtt{Q}}, \mathsf{TIME}_{\mathtt{D}}\right\} \le n^{\alpha}.$$

Then there exist some $n_0^{QR} = \mathrm{poly}(\gamma^{QR}, n_0)$ with $n_0 \le n_0^{QR}$ such that for all $n \ge n_0^{QR}$

1. *(Completeness) If there exists a perfect oracularizable synchronous strategy for $\mathscr{G}_n$ in model t, then there exists a perfect oracularizable synchronous strategy for $\mathscr{G}_n^{QR}$ in model t.*
2. *(Soundness) There exist some universal function $\mathbf{s}_\alpha^{QR}$ which depends on k and ε with $\mathbf{s}_\alpha^{QR} = O(\exp(k), \mathrm{poly}(\varepsilon))$ such that, for any polynomial $\varepsilon : \mathbb{N} \to [0,1]$*

$$\omega^t(\mathscr{G}_n) \le 1 - \varepsilon(n) \Longrightarrow \omega^t(\mathscr{G}_n^{QR}) \le 1 - \mathbf{s}_\alpha^{QR}(k, \varepsilon(n)).$$

Proof. Fix the constant $\alpha, k \in \mathbb{N}$. We define the algorithm $\mathtt{QuestionReduction}_{\alpha,k}$ as follows. Given a pair of Turing machine $(\mathtt{Q},\mathtt{D})$, we first describe a sequence of typed samplable games $\mathscr{G}_n^{\mathrm{Intro}}$, then we use Lemma 7.2.11 to convert $\mathscr{G}_n^{\mathrm{Intro}}$ into a CL samplable game as desired. Hence, fix some input $(\mathtt{Q},\mathtt{D})$ and integer n, we define $\mathscr{G}_n^{\mathrm{Intro}}$ as the following:

The game $\mathscr{G}_n^{\mathrm{Intro}}$ has the sampling procedure for the $(\mathscr{G}, n^\alpha, k)$-introspection game as pre given Figure 9.1 for any arbitrary game $\mathscr{G}$. By Theorem 9.1.1, the input distribution is independent of the game $\mathscr{G}$. For the decision process, given $(v_0, v_1) \in \mathtt{E}_k^{\mathrm{Intro}}$, $u_x, u_z \in \{0,1\}^{\alpha\lceil\log(n)\rceil}$, the question label that the verifier sends to the two provers. Let $a, b \in \{0,1\}^*$ be the answer that the verifier receives the answers based on the question label. The verifier computes the following: If at any point in the computation process, $|a|, |b| \ge 3 \cdot n^\alpha$ (since each input have at most 3 item of length at most n^α), or the computation step for running $\mathtt{Q}$, $\mathtt{D}$ either returns an invalid output or runs for time more than n^α steps, the verifier terminates and returns 0 (i.e. the verifier

rejects). The verifier first computes $(k_n, m_n, p_n) = \mathtt{Q}(n, \text{parameter})$, and rejects if $k_n > k$ and $m_n \cdot p_n > n^\alpha$.

Based on the vertices $(v_0, v_1) \in \mathbb{E}_k^{\text{Intro}}$, the verifier first divides the answer associate with each question label into the format given in Figure 9.2. Then the verifier does the following based on (v_0, v_1):

- (n^α-Pauli Basis): The verifier accepts according to the rules described in Figure 8.3, this can be done uniformly in time $O(\text{poly}(n))$ by Theorem 8.3.3.
- (Self-loop): The verifier accepts iff $a = b$; otherwise reject. This can be done in $O(n^\alpha)$ time by the terminating assumption above.
- (Pauli, W) – (Gen Pauli, W): The verifier accepts iff $|b| = p_n \cdot m_n$ and the first $p_n \cdot m_n$ bits of a are equal to b (where recall, the elements $b \in \mathbb{F}_{2^p}^n$ is represented using the canonical representation in this thesis).
- (Gen Pauli, X) – (Hide$_0$, $\mathtt{L}^P$): The verifier first uses $\mathtt{Q}(n, \text{Function}, \vec{1})$ to compute the canonical basis which spans V_0^n (where $\vec{1} \in \{0,1\}^{m(n)\cdot p_n}$ is the all 1 vector). For each $\hat{e}_j$, the canonical basis which spans V_0, compute $\mathtt{Q}(n, \text{Function}, P, 0, 0, \hat{e}_i)$, and run the standard Gaussian elimination to find the description of the subspace $\ker \mathtt{L}_{0,0}^{P,n}$ and $\ker \mathtt{L}_{0,0}^{P,n\perp}$. Finally, parse $s_X = (s_X)_0 + (s_X)_0^\perp + (s_X)_{>0}^n \in \ker \mathtt{L}_{0,0}^{P,n} \oplus \ker \mathtt{L}_{0,0}^{P,n\perp} \oplus V_{>0}$. The verifier accepts iff the answers from the provers are in the correct subspace according to Figure 9.3 (i.e. $t_{\le 0}^\perp \in V_0^n$) and $(s_X)_0^\perp = t_{\le 0}^\perp$ and $(s_X)_{>0} = r_{>0}$.
- (Hide$_i$, $\mathtt{L}^P$) – (Hide$_{i+1}$, $\mathtt{L}^P$) for $i \in [k]$: If $i \ge k_n$, treat this as a consistency check. Otherwise using the same technique as above, compute the description for $V_{<i}^n$, V_i^n and $V_{>i+1}$. Parse

  $$t_{\le i+1}^\perp = \tilde{t}_{\le i}^\perp + \tilde{t}_{i+1}^\perp \in V_{\le i}^n \oplus V_{i+1}^n, \quad t_{<i+1}^{\text{Line}} = \tilde{t}_{<i}^{\text{Line}} + \tilde{t}_i^{\text{Line}} \in V_{<i}^n \oplus V_i^n,$$

  and use a similar computation step to compute the description for $\ker\left(\mathtt{L}_{i+1, t_{<i+1}^{\text{Line}}}^{P,n}\right)$ and $\ker\left(\mathtt{L}_{i+1, t_{<i+1}^{\text{Line}}}^{P,n}\right)^\perp$. Then the verifier parse

  $$r_{>i} = \bar{r}_i + \bar{r}_i^\perp + \bar{r}_{>i+1} \in \ker\left(\mathtt{L}_{i+1, t_{<i+1}^{\text{Line}}}^{P,n}\right) \oplus \ker\left(\mathtt{L}_{i+1, t_{<i+1}^{\text{Line}}}^{P,n}\right)^\perp \oplus V_{>i+1}^n,$$

  The verifier accepts iff the answers from the provers are in the correct subspace

  $$\tilde{t}_{\le i}^\perp = t_{\le i}^\perp, \quad \bar{r}_i^\perp = \tilde{t}_{i+1}^\perp, \quad \tilde{t}_{<i}^{\text{Line}} = \tilde{t}_{<i}^{\text{Line}}, \quad \bar{r}_{>i+1} = r_{>i+1}.$$

- (Hide$_{k-1}$, $\mathtt{L}^P$) – (Read, $\mathtt{L}^P$): The verifier accepts iff the answers from the provers are in the correct subspace, $t_{\text{Read},P} = t_{k-1}$, and $t_{\text{Read},P}^\perp = t_{k-1}^\perp$.
- (Read, $\mathtt{L}^P$) – (Intro, $\mathtt{L}^P$): The verifier accepts iff $t_{\text{Read},P}^{\text{Line}} = x_P$, and $a_{\text{Read},P} = a_P$.
- (Gen Pauli, Z) – (Sample, $\mathtt{L}^P$): The verifier accepts iff $s_Z = s_{\text{Sample}}$.
- (Sample, $\mathtt{L}^P$) – (Intro, $\mathtt{L}^P$): The verifier computes $\mathtt{L}^{P,n}(s_{\text{Sample}})$ by using the algorithm provided in Section 7.3 with $\mathtt{Q}$. The verifier accepts iff the output provided by $\mathtt{L}^P(s_{\text{Sample}})$ is equal to x_P and $a_{\text{Sample},P} = a_P$.
- (Intro, $\mathtt{L}^A$) – (Intro, $\mathtt{L}^B$): The verifier accepts iff $\mathtt{D}(n, x_A, x_B, a_A, a_B) = 1$.

The above procedure uniformly defines a $(\mathscr{G}_n, k, p)$-introspection game for $n \geq n_0$ assuming $(\mathtt{Q}, \mathtt{D})$ are valid Turing machines as given in the second part of Proposition 7.4.2. By Lemma 7.1.1 and the hardcoded computation bound on $\mathtt{Q}$ and $\mathtt{D}$, the above procedure can be computed in $O(\text{poly}(n,k))$ time.

Since $\mathscr{G}_n^{\text{Intro}}$ is typed samplable and $|\mathrm{E}_k^{\text{Intro}}| = 30 + 2k$, by applying Lemma 10.2.1 to each $\mathscr{G}_n^{\text{Intro}}$, we obtain a sequence of $(3, \alpha\lceil\log(n)\rceil + C^{\text{detype}}, 2)$ CL samplable games $\mathscr{G}_n^{\text{QR}}$, where C is some constant that depends linearly on k. Since the detyping procedure given in Definition 7.2.8 only adds extra string parsing and synchronization checks to the sampling and decision procedure. Each $\mathscr{G}_n^{\text{QR}}$ can still be sampled in $O(\text{poly}(\log(n), k))$ time and verified in $O(\text{poly}(n))$ time. Pick $\gamma^{\text{QR}} \in \mathbb{N}$ to be sufficiently large so that $\mathscr{G}_n^{\text{QR}}$ can be sampled in $O(\log^{\gamma^{\text{QR}}}(n))$ time, $k \cdot \alpha \cdot \lceil\log(n)\rceil + C^{\text{detype}} = O(\log^{\gamma^{\text{QR}}}(n))$ and $\mathscr{G}_n^{\text{QR}}$ can be decided in $O(\text{poly}(n))$ time.

Define $(\mathtt{Q}^{\text{QR}}, \mathtt{D}^{\text{QR}})$ in the following way: let n_0^{Run} be the constant such that for all $n > n_0^{\text{Run}}$, $\mathscr{G}_n^{\text{QR}}$ can be sampled in fewer than $\log^{\gamma^{\text{QR}}}(n)$ steps (no big-O notation here!) and decided in time fewer than $n^{\gamma^{\text{QR}}}(n)$ steps, and furthermore $k \cdot \alpha \cdot \lceil\log(n)\rceil + C^{\text{detype}} \leq \log^{\gamma^{\text{QR}}}(n)$. For all $n < n_0^{\text{Run}}$, $(\mathtt{Q}^{\text{QR}}(n), \mathtt{D}^{\text{QR}}(n))$ returns an encoding of the rejecting game $\mathscr{G}^{\text{reject}}$ defined in Definition 6.2.3, otherwise return an encoding for $\mathscr{G}_n^{\text{QR}}$.

We now verify all the properties listed in Proposition 7.4.2.

1. (Computation time): This follows since the description of $(\mathtt{Q}^{\text{QR}}, \mathtt{D}^{\text{QR}})$ only depends on the description of $(\mathtt{Q}, \mathtt{D})$ and some fixed constant.

2. (Synchronicity): This follows from the fact that $\mathscr{G}_n^{\text{Intro}}$ is always synchronous.

3. (Complexity bounds for the output): Let C^{trivial} be the constant such that $\mathscr{G}^{\text{reject}}$ can be both sampled and decided in time C^{trivial}. This follows from the definition of $(\mathtt{Q}^{\text{QR}}, \mathtt{D}^{\text{QR}})$.

4. (Independency) This follows since both $\mathscr{G}_n^{\text{Intro}}$ and the detyping procedure can be defined uniformly for all verifier sequences $(\mathtt{Q}, \mathtt{D})$, and the sampling procedure for $\mathscr{G}_n^{\text{Intro}}$ does not depend on $\mathtt{D}$

Let $\mathscr{V} = (\mathtt{Q}, \mathtt{D})$ and $n_0 \in \mathbb{N}$ be as pre described in the theorem. Let $n_0^{\text{QR}} = \max\{n_0^{\text{Run}}, n_0\}$, which both depend on n^λ, and n_0^{Run} depends on the constant k. For all $n \geq n_0^{\text{QR}}$:

1. (Completeness): This follows from Theorem 9.1.1 and Lemma 7.2.11.

2. (Soundness): By Theorem 9.1.1: There exists a polynomial $\mathbf{s}_\alpha^{\text{Intro}}$ such that

$$\omega^*(\mathscr{G}_n) \leq 1 - \varepsilon(n) \Longrightarrow \omega_s^t(\mathscr{G}_n^{\text{Intro}}) \leq 1 - \mathbf{s}_\alpha^{\text{Intro}}(k, \varepsilon(n))$$

Let $\mathbf{s}_\alpha^{\text{QR}}(k, \varepsilon(n)) = \frac{\mathbf{s}_\alpha^{\text{Intro}}(k, \varepsilon(n))}{4(30+2k)^2 \cdot 16^{(30+2k)}}$. By Lemma 7.2.11

$$\omega^*(\mathscr{G}_n) \leq 1 - \mathbf{s}_\alpha^{\text{Intro}}(k) \Longrightarrow \omega_s^t(\mathscr{G}_n^{\text{Intro}}) \leq 1 - \mathbf{s}_\alpha^{\text{QR}}(\exp(k), \varepsilon(n)).$$

This concludes the proof for Proposition 7.4.2.

□

CHAPTER 10

Answer Reduction part 1: The low-individual degree test

The next two chapters will be focused on the answer reduction procedure. We start by introducing an important subroutine for the answer reduction protocol, the quantum simultaneous low-individual degree test. The simultaneous low-individual degree test is an important part in showing NEXP ⊆ MIP in [BFL91], and the quantum variant, intuitively, will be used in a similar manner in the answer reduction protocol when proving Proposition 7.4.3.

Before introducing the quantum simultaneous low-degree test, we start by introducing some additional preliminaries in coding theory in Section 10.1. We then introduce the quantum low-individual degree test in Section 10.2. The simultaneous variant can be thought of as the "parallel repetition" variant of the quantum low-degree test, and it was already known to be "quantum sound" for synchronous strategies (for both the tensor product and the commuting operator model) in [JNV+22b]. In Section 10.3, we show how one can use the rounding theorem (Theorem 5.1.1) to extend the quantum soundness result from synchronous strategies to general tracially embeddable strategies. Finally, we introduce the simultaneous low-individual degree test in Section 10.4 and prove the quantum soundness of this protocol.

10.1 Preliminaries in Coding theory

We use the notation from [JNV+22b]. Recall from Section 8.3.2 that a linear $[n,c,d]_{\mathbb{F}_q}$ code is a set $\mathfrak{C}$ of functions $g:[p]\to\mathbb{F}_q$ with size $|\mathfrak{C}|=q^c$ that is closed under linear combination, such that for any two distinct $g\neq g'$, the number of coordinates $i\in[n]$ such that $g(i)\neq g'(i)$ is at least d. Given $\mathfrak{C}$, a linear $[n,c,d]_{\mathbb{F}_q}$ code, the tensor code $\mathfrak{C}^{\otimes m}$ is the set of all functions $f:[n]^m\to\mathbb{F}_q$ such that the restriction $f|_{\mathbf{l}_j}$ to any axis-parallel line $\mathbf{l}_j$ is a codeword in $\mathfrak{C}$, where for $j\in m$, an axis parallel line $\mathbf{l}_j$ is defined as

$$\mathbf{l}_j=\{(s_0,\cdots,s_{j-1},x,s_{j+1},\cdots,s_{m-1}):x\in\mathbb{F}_q\}$$

Given constants $p,c\in\mathbb{N}$, the set $\mathfrak{C}$ consists of all degree c polynomials $\mathbf{f}:\mathbb{F}_{2^p}\to\mathbb{F}_{2^p}$ is a $[2^p,c,c]_{\mathbb{F}_{2^p}}$ code by the Schwartz-Zippel lemma. We further see that the set of low-individual degree polynomials $\mathbf{f}:\mathbb{F}_{2^p}^m\to\mathbb{F}_{2^p}$ with individual degree at most c is a tensor code $\mathfrak{C}^{\otimes}$ (since $\mathbf{f}|_{\mathbf{l}_j}$ always gives a 1-variant polynomial of degree at most c).

Low-individual degree polynomials can also be used to define an error correction code with good distance properties in the context of the *generalized Reed-Muller code*. Given a string $s\in\{0,1\}^m$, we define the *indicator polynomial* for m over the field $\mathbb{F}_{2^p}$ as the m-variate polynomial with $ind_{m,a}:\mathbb{F}_{2^p}^m\to\mathbb{F}_{2^p}$ as

$$ind_{m,a}(x):=\prod_{i:a_i=1}x_i\cdot\prod_{i:a_i=0}(1-x_i).$$

where we identify $\{0,1\}^m$ as elements of $\mathbb{F}_{2^p}^m$. The indicator polynomial has individual degree of 1 and has the properties that for all $s \in \{0,1\}^m \subseteq \mathbb{F}_{2^p}^m$, $ind_a(s) = 0$ except when $a = s$.

Let $M = 2^m$, for any elements $b \in \{0,1\}^M$, we define the generalized Reed-Muller encoding of a to be the polynomial

$$\mathrm{RM}_b(x) := \sum_{y \in \{0,1\}^m} b_{\mathbf{bininv}(y)} ind_{m,y}(x), \tag{10.1}$$

where recall, $\mathbf{bininv}(\cdot)$ is the map which maps the corresponding binary representation back to an integer. Since each $ind_{m,y}(x)$ has an individual degree of 1, RM_b also has an individual degree of 1. For any $y \in \{0,1\}^m \subseteq \mathbb{F}_q^m$, evaluating RM_b with y returns the $\mathbf{bininv}(y)$th coordinate of the string b.

10.2 The quantum low-individual degree test

We start this section by giving some high-level intuition about the quantum low-individual degree test. The quantum low-individual degree test is first introduced in [JNV+22b], and it is based on the classical low-degree test given in [BFL91]. The classical low-individual degree test was used in the first proof of PCP theorem [AS98; ALM+98], and the quantum low-individual degree is used in a similar way in the answer reduction protocol in this thesis.

The quantum low-individual degree test is parameterized by $(p,m,d) \in \mathbb{N}^3$, where $m = 2^c$ for some constant c. Intuitively, the goal of the quantum low-individual degree test is to force two entangled provers to prove to the verifier that they both share the same global m-variant polynomial over 2^p with individual degree of at most d. To give a better intuition on how the quantum low-individual degree test works, we first give a brief description of its classical counterpart.

Suppose the two provers agree on a global m-variant low-individual degree polynomial $\mathbf{g}$: $\mathbb{F}_{2^p}^m \to \mathbb{F}_{2^p}$ with individual degree of at most d. To demonstrate to the verifier that they both share the same polynomial, both provers can simply send g to the verifier. However, since $\mathbf{g}$ can have at most $(d+1)^m$ monomials, this means that the prover needs to send a messages with potential length $O((d+1)^m \cdot p)$-bits which is inefficient if m is large. Instead, the verifier can send one of the prover $u \in \mathbb{F}_{2^p}^m$ and ask him to evaluate the polynomial $\mathbf{g}$ on u and return the outcome $\mathbf{g}(u) \in \mathbb{F}_{2^p}$. The verifier then gives the other prover a random "parallel axis line" $\mathbf{l} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$ that pass through u, where for $i \in [m]$, an axis parallel line passing through u is defined as

$$\mathbf{l} = \{u + x \cdot e_i | x \in \mathbb{F}_{2^p}\}. \tag{10.2}$$

The prover receiving the axis parallel line is expected to return the polynomial $\mathbf{g_l} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$, where $\mathbf{g_l}$ is the $m \cdot d$-th degree polynomial corresponding to $\mathbf{g}$ restricted to the range of $\mathbf{l}$, or

$$\mathbf{g_l}(x) = \mathbf{g}(u + x \cdot e_i).$$

Intuitively, this is asking the prover to evaluate $\mathbf{g}$ on *all* of $\mathbf{l}$. If the two provers share the same low-individual degree polynomial in the beginning of the protocol, then $\mathbf{g_l}(0)$ would be consistent with $\mathbf{g}(u)$. If, on the other hand, the two provers do not share the same low-individual degree polynomial, then by Lemma 7.1.4, for two different m-variant low-individual degree polynomials $\mathbf{g}$ and $\mathbf{g}'$ and $u \in \mathbf{l}$, $\mathbf{g}(u) = \mathbf{g}'(u)$ occurs with probability at most $\frac{d}{q}$. This means that the restricted polynomial $\mathbf{g_l}$ generated by the axis parallel line prover is unlikely to agree with the prover with who is expected to evaluate $\mathbf{g}$ somewhere on $\mathbf{l}$. This protocol allows the verifier to check the consistency of a global low-individual degree polynomial with message size $O(d \cdot k)$, significantly more efficient than the previous protocol.

In order to make sure the above protocol remains quantum sound (i.e. the provers will have a low success rate if they do not share the same low-individual degree test polynomial) [JNV+22b]

The quantum low-individual degree test is parametrized by $(p, m, d) \in \mathbb{N}^3$ where p, m are both an odd integer. Perform the following test with probability $\frac{1}{8}$ each:

- **Axis parallel line test.** The verifier uniformly samples $s = (s_0, \cdots, s_{m-1}) \in \mathbb{F}_{2^p}^m$ and $j \in [m]$. Let $\mathbf{l}_j$ be the jth axis-parallel line given in (10.2). Recall from Definition 7.1.3 that $\mathrm{Can}(\mathbf{l}_j)$ is the canonical representation of a line.
  - The verifier sends (Point, s) to one of the provers, and receive $a \in \mathbb{F}_m$ as a response.
  - The verifier sends $(\mathrm{Aline}, \mathrm{Can}(\mathbf{l}_j))$ to the other prover, and receive a degree d polynomial $\mathbf{f} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$ as a response.

  The verifier accepts if $\mathbf{f}(s) = a$.
- **Diagonal line test.** The verifier uniformly samples $s = (s_0, \cdots, s_{m-1}) \in \mathbb{F}_{2^p}^m$, $j \in [m]$ and $v \sim \mathbb{F}_{2^p}^j$. Extend v as an element of $\mathbb{F}_{2^p}^m$ by appending 0 on the last $m - j$ coordinates. Define the line $\mathbf{d}_{j,v} = \{s + x \cdot v : x \in \mathbb{F}_{2^p}\}$ to be the line passing through s in the direction of v.
  - The verifier sends (Point, s) to one of the provers, and receive $a \in \mathbb{F}_m$ as a response.
  - The verifier sends $(\mathrm{Dline}, \mathrm{Can}(\mathbf{d}_{j,v}))$ to the other prover, and receive a degree $d \cdot m$ polynomial $\mathbf{g} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$ as a response.

  The verifier accepts if $\mathbf{g}(s) = a$.

Perform the following test with probability $\frac{1}{4}$ each:

- **Point consistency test.** The verifier uniformly samples $s \in \mathbb{F}_{2^p}^m$. The verifier sends (Point, s) to both provers, and receive $(a, b) \in \mathbb{F}_{2^p}^2$. The verifier accepts if $f(s) = a$.
- **Axis parallel line consistency test.** The verifier uniformly samples $s \in \mathbb{F}_{2^p}^m$ and $j \in [m]$. Let $\mathbf{l}_j$ be the axis parallel line define in the "Axis parallel line test". The verifier sends $\mathrm{Can}(\mathbf{l}_j)$ to both provers, and receive two degree c polynomial $\mathbf{f}^A, \mathbf{f}^B : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$. The verifier accepts if $\mathbf{f}^A = \mathbf{f}^B$.
- **Diagonal line line consistency test.** The verifier uniformly samples $s \in \mathbb{F}_{2^p}^m$, $j \in [m]$ and $v \sim \mathbb{F}_{2^p}^j$. Let $\mathbf{d}_{j,v}$ be the line define in the "Diagonal line test". The verifier sends $\mathrm{Can}(\mathbf{d}_{j,v})$ to both provers, and receive two degree $d \cdot m$ polynomial $\mathbf{g}^A, \mathbf{g}^B : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$. The verifier accepts if $\mathbf{g}^A = \mathbf{g}^B$.

Figure 10.1: The sampling/decision procedure for the (p, m, d)-quantum low-individual degree test, the only change in comparison to the original test as presented in [JNV+22a] is the distribution of the synchronicity test, which only changes the constant in Theorem 10.3.1 (see [JNV+22b, Section 4.1] for more details).

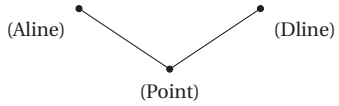

Figure 10.2: The typed graph for the quantum low-individual degree test

added two additional questions types. The first is the "diagonal line test", where, one of the provers still receives a random point $u \in \mathbb{F}_{2^p}^m$ and similarly is expected to return $\mathbf{g}(u)$ in the ideal case. The other prover is given a random "diagonal line" intersecting with u, where we define the notion of a diagonal line below, and is expected to return an $m \cdot d$-th degree polynomial similarly as to above. This addition is mostly used to enforce a commutation relationship for the proof of soundness for the quantum low-degree test. Secondly, to ensure synchronicity, a "consistency test" is added, where the two provers are given the same question (which can be a line, a point, or a diagonal line) and they are expected to output the same answer. We formally give the definition for a quantum low-individual degree test on Figure 10.1, and we state the soundness result in the next section.

We remark that in our description, the verifier sends the canonical representation of the line. This is equivalent to the original formulation where the verifier sends a set containing all points in the line to the prover. Both formulations are designed to hide the point u (for the "point" player) when sending the line. As seen in the description from the classical low-degree test, if both provers share an m-variate polynomial $\mathbf{h}$ over $\mathbb{F}_{2^p}$ with individual degree of at most d. The provers can simply plug their respective input into $\mathbf{h}$ to obtain a perfect (classical) strategy.

For the remainder of this subsection, we show that the quantum low-individual degree test can be sampled via a typed CL distribution, and hence can be converted to a game with a CL distribution as the input distribution via Lemma 7.2.11.

Lemma 10.2.1 (The quantum low-individual degree test can be sample via a CL distribution)**.** *Let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, V)$ be a (p, m, d)-quantum low-individual degree test, then there exists a game $\mathcal{G}' = (\mathcal{X}', \mathcal{A}, \mu', V')$ which is $(5, 9 + m' + 2 \cdot m, p)$ CL samplable where $m' = \left\lceil \frac{\log(m)}{p} \right\rceil$, and $\mathcal{X} \subseteq \mathcal{X}'$ such that the following holds: For any $t \in \{*, co\}$ and $\varepsilon > 0$. Any synchronous, oracularizable strategy $\mathcal{S}$ in model t such that $\omega(\mathcal{G}', \mathcal{S}) \geq 1 - \varepsilon$. $\mathcal{S}$, when restricted on the question pair from $\mathcal{G}$, $\omega(\mathcal{G}, \mathcal{S}) \geq 1 - c\varepsilon$ for some constant c.*

Proof. We first show that the quantum low-individual degree test can be sampled via a typed CL distribution. Let (p, m, d) be the parameter specified and let $q = 2^p$. For simplification, we first assume $\frac{\log(m)}{p} \in \mathbb{N}$ (and hence $m' \in \mathbb{N}$). The typed graph associated with the quantum low-individual degree test can be specified by the types stated in Figure 10.2. The CL function maps $\mathbb{F}_{2^p}^{m'+2\cdot m} \to \mathbb{F}_{2^p}^{m'+2\cdot m}$, where we write $\mathbb{F}_{2^p}^{m'+2\cdot m} = \mathbb{F}_{2^p}^{m'} \oplus \mathbb{F}_{2^p}^{m} \oplus \mathbb{F}_{2^p}^{m} = V_0 \oplus V_1 \oplus V_2$. The CL functions is a level 3-CL function with register $\{V_0, V_1, V_2\}$. For the description below, we assume the function have the input $s = (s_0, s_1, s_2) \in V_0 \oplus V_1 \oplus V_2$. We define the collection of CL functions $\{\mathtt{L}^v\}$ as the following:

- Define $\mathtt{L}^{\text{Point}}$ to be the third level CL function

$$\mathtt{L}^{\text{Point}}(s_0, s_1, s_2) = (0, 0, s_2).$$

  In other words, $\mathtt{L}^{\text{Point}}$ projects the input s onto V_2, and s_2 corresponds to the point in the "point question" for the quantum low-individual degree test. $\mathtt{L}^{\text{Point}}$ can be realized as a third level CL function with registers $\{V_i\}_{i\in[3]}$ in the following way: We define

$$\mathtt{L}_{0,0}^{\text{Point}}(s_0, 0, 0) = (0, 0, 0), \qquad \mathtt{L}_{1,x_0}^{\text{Point}}(0, s_1, 0) = (0, 0, 0), \qquad \mathtt{L}_{2,x_0+x_1}^{\text{Point}}(0, 0, s_2) = (0, 0, s_2),$$

  for all $x_0 \in V_0$ and $x_1 \in V_1$.

- Define $\mathtt{L}^{\mathrm{Dline}}$ to be the third level CL function. For any input $(s_0, s_1, s_2) \in \mathbb{F}_{2^p}^{m'+2\cdot m}$, let $\hat{j} = \kappa(s_0)$. The function $\mathtt{L}^{\mathrm{Dline}}$ is defined as

$$\mathtt{L}^{\mathrm{Dline}}(s_0, s_1, s_2) = (s_0, 0, \mathrm{Null}^{\mathrm{LN}}_{e_{\mathbf{bininv}(\hat{j})}}(s_2)),$$

  where $\mathrm{Null}^{\mathrm{LN}}$ is the function used in Definition 7.1.3 to defined the canonical representation of a line. In the example above, $(e_{\mathbf{bininv}(\hat{j})}, \mathrm{Null}^{\mathrm{LN}}_{e_{\mathbf{bininv}(\hat{j})}}(s_2)$ defines an axis parallel line thought the $\mathbf{bininv}(\hat{j}) = j$th axis through the point s_2. $\mathtt{L}^{\mathrm{Dine}}$ can be realized as a third level CL function with registers $\{V_i\}_{i\in[3]}$ in the following way: We define

$$\mathtt{L}^{\mathrm{Dline}}_{0,0}(s_0, 0, 0) = (s_0, 0, 0), \qquad \mathtt{L}^{\mathrm{Dline}}_{1,s_0}(0, s_1, 0) = (0, 0, 0)$$

  for all $x_0 \in V_0$. For the second level, for all $x_0 \in V_0$, $x_1 \in V_1$, we define the second level linear function for $\mathtt{L}^{\mathrm{Dline}}$ as

$$\mathtt{L}^{\mathrm{Dline}}_{2,x_0+x_1}(0, 0, s_2) = (0, 0, \mathrm{Null}^{\mathrm{LN}}_{e_{\mathbf{bininv}(x_0)}}(s_2)).$$

- Let $\hat{j}$ be the same as the definition above. Recall that $\pi^m_{\leq j}$ refers to the zero-out map for the basis element $e_1 \cdots e_m$ in $\mathbb{F}^m_{2^p}$ and let $v = \pi^m_{\leq \mathbf{bininv}(\hat{j})}(s_0)$. Define $\mathtt{L}^{\mathrm{Aline}}$ to be the third level CL function

$$\mathtt{L}^{\mathrm{Aline}}(s_0, s_1, s_2) = (s_0, v, \mathrm{Null}^{\mathrm{LN}}_v(s_2)).$$

  In this case, $(v, \mathrm{Null}^{\mathrm{LN}}_v(s_2))$ corresponds to the diagonal line $D_{\mathbf{bininv}(s_0),v}$ which passes through s_2, and with the last $m - \mathbf{bininv}(\hat{j})$ coordinates being zero. $\mathtt{L}^{\mathrm{Aline}}$ can be realized as a third level CL function with registers $\{V_i\}_{i\in[3]}$ in the following way: We define

$$\mathtt{L}^{\mathrm{Aline}}_{0,0}(s_0, 0, 0) = (s_0, 0, 0),$$

  to be the 0th level linear function for $\mathtt{L}^{\mathrm{Aline}}$. Since each $\pi^m_{\leq j}$ is a linear function, we define the first level linear function as

$$\mathtt{L}^{\mathrm{Aline}}_{1,x_0}(0, s_1, 0) = (0, \pi_{\mathbf{bininv}(x_0)}(s_1), 0),$$

  for all $x_0 \in V_0$. We define the second level linear function for $\mathtt{L}$ as

$$\mathtt{L}^{\mathrm{Aline}}_{2,x_0+x_1}(0, 0, s_2) = (0, 0, \mathrm{Null}^{\mathrm{LN}}_{x_1}(s_2))$$

  for all $x_0 \in V_0$ and $x_1 \in V_1$.

This shows that $\mathscr{G}$ is typed CL samplable. In the case where $\frac{\log(m)}{p} \notin \mathbb{N}$, we can simply treat the space $\mathbb{F}^{m'}_{2^p}$ as a $\{0,1\}^{p\cdot m'}$ bit string using the canonical representation, and only use the first $\log(m)$ bits to define the CL function. Finally, we use the detyped transformation and apply Lemma 7.2.11 to complete the proof of this lemma. □

When discussing the quantum low-individual degree test in this thesis, we refers to the version which is CL samplable given by the above lemma. This version of the quantum low-individual degree test still retains the soundness property from Theorem 10.3.1 (by changing the a in the aforementioned theorem to the constant $(4 * 4^2 + 16^4)^b \cdot a = (65600)^b \cdot a$).

We remark that the only time we take advantage of the structure of $\mathbb{F}^m_{2^p}$ (instead of treating it as a bit string of $\{0,1\}^{pm}$) when using the CL distribution is to sample a diagonal line intersecting the point s in the quantum low-individual degree test. As mentioned previously, the diagonal line test was not used in the classical low-degree test, and the only purpose conceptually is to enforce a single commutation relationship within the proof of quantum soundness (see [JNV+22b, Lemma 6.1] for more details). It would be interesting to see if the quantum soundness of the quantum low-individual degree test still holds without the diagonal line test, as this would allow us to show the compression theorem for a simpler class of question distribution. This also shows that only Pauli X and Z measurements on fixed EPR pairs are sufficient for the gap compression theorem.

10.3 Quantum soundness of the quantum low-individual degree test

We recall the following soundness result related to the quantum low-individual degree test:

Theorem 10.3.1 (Quantum soundness for the quantum low-individual degree test under synchronous strategies, Theorem 4.1 of [JNV+22a])**.** *There exist universal constants $c_{LD,1} \leq 1$ and $0 < c_{LD,2} \leq 1$ and a function*

$$\eta_{LD}(p,m,d,\varepsilon) = c_{LD,1}(dm)^{c_{LD,1}}(\varepsilon^{c_{LD,2}} + 2^{-c_{LD,2}p} + 2^{-c_{LD,2}md})$$

such that the following holds. Let $\mathcal{G}^{LD}$ be the (p,m,d)-low-individual degree test with $q = 2^p$, and let $\mathcal{S} = (\mathcal{L}^2(\mathcal{A},\tau), |\tau\rangle, \{A_a^x\})$ be a synchronous strategy for $\mathcal{G}^{LD}$ which succeed with probability $1-\varepsilon$. There exist a set of PVM $\{G_{\boldsymbol{g}}\} \subseteq \mathcal{A}'$ with outcome labelled by m-variant polynomials $\boldsymbol{g} \in IdPoly(p,m,d)$, such that

$$\mathop{\mathbb{E}}_{s\sim\mathbb{F}_{2^p}^m} \sum_{\boldsymbol{g}\in IdPoly(p,m,d)} \langle\tau| A^{(point,s)}_{\boldsymbol{g}(s)} \cdot G_{\boldsymbol{g}} |\tau\rangle \geq 1 - \eta_{LD}(p,m,d,\varepsilon).$$

In other words, if the provers succeed on the (p,m,c)-low-individual degree test with high probability, then the provers, in essence, are secretly sampling a low-individual degree polynomial which are then used as a part of the strategy. Although [JNV+22b, Theorem 4.1] is originally proven for tensor codes, as shown in Section 7.1, the set of m-variant polynomials over $\mathbb{F}_{2^p}^m$ with low-individual degree of at most c is a tensor code $\mathfrak{C}^{\otimes m}$ for a $[2^p, c, c]_{\mathbb{F}_{2^p}}$ linear code $\mathfrak{C}$, and hence the same statement can be directly applied to the quantum low-individual degree test. We remark that the proof of the above theorem follows a similar structure as [BFL91, Theorem 5.4], as a quantum variant of the two key lemmas, the self-improvement lemma and the pasting lemma, is also a key lemma in [JNV+22b].

By using the rounding theorem Theorem 5.1.1, one could extend Theorem 10.3.1 to the case where the provers have access to tracially embeddable strategies. We remark that the proof of the theorem below follows a similar structure as [Vid22, Corollary 4.1].

Theorem 10.3.2 (Quantum soundness for the quantum low-individual degree test under tracially embeddable strategy)**.** *There exist universal constants $1 \geq c_{LD,1}$ and $0 < c_{LD,2} \leq 1$ and a function*

$$\eta_{LD}(p,m,d,\varepsilon) = c_{LD,1}(dm)^{c_{LD,1}}(\varepsilon^{c_{LD,2}} + 2^{-c_{LD,2}p} + 2^{-c_{LD,2}md})$$

such that the following holds. Let $\mathcal{G}^{LD}$ be the (p,m,d)-low-individual degree test with $q = 2^p$, and let $\mathcal{S} = (\mathcal{L}^2(\mathcal{A},\tau), |\tau\rangle, \{A_a^x\})$ be a synchronous strategy for $\mathcal{G}^{LD}$ which succeed with probability $1-\varepsilon$. There exist a set of PVM $\{G_{\boldsymbol{g}}\} \subseteq \mathcal{A}'$ with outcome labelled by m-variant polynomials $\boldsymbol{g} \in IdPoly(p,m,d)$, such that

$$\mathop{\mathbb{E}}_{s\sim\mathbb{F}_{2^p}^m} \sum_{\boldsymbol{g}\in IdPoly(p,m,d)} \langle\tau| A^{(point,s)}_{\boldsymbol{g}(s)} \cdot G_{\boldsymbol{g}} |\tau\rangle \geq 1 - \eta_{LD}(p,m,d,\varepsilon).$$

Proof. We first assume the strategy $\mathcal{S} = (\mathcal{L}^2(\mathcal{A},\tau), \sigma|\tau\rangle, \{A_a^x\})$ is symmetric, projective with $\sigma \in \mathcal{A}^+$. Let $\{P_\lambda\}$ and $\{A_a^{\lambda,x}\}$ be the set of projectors and PVMs promised by Proposition 5.2.2. Let $\mathcal{S}^\lambda = (\mathcal{L}^2(P_\lambda \mathcal{A} P_\lambda, P_\lambda\tau), \frac{1}{\sqrt{\tau(P_\lambda)}} P_\lambda|\tau\rangle, \{A_a^{\lambda,x}\})$ be the synchronous strategy from $\{A_a^{\lambda,x}\}$. Given $\lambda \in [0,\infty)$ with $\tau(P_\lambda) \neq 0$. By the "soundness property" from the assumption, for each $\lambda \in [0,\infty)$, since each $\mathcal{S}^\lambda$ is synchronous, and $(P_\lambda \mathcal{A} P_\lambda)' = \mathcal{A}' P_\lambda$, there exists a family of POVM $\{G_b^{\lambda,y}\}_{(y,b)\in\mathcal{Y}\times\mathcal{B}} \subseteq P_\lambda \mathcal{A} P_\lambda$ such that

$$\frac{1}{\tau(P_\lambda)} \mathop{\mathbb{E}}_{(x,y)\sim\rho} \sum_{a\in\mathcal{A}} \langle\tau| P_\lambda A_a^{x,\lambda} (G^{\lambda,y}_{g_{xy}(a)})^{op} P_\lambda |\tau\rangle \geq \kappa(\omega^{co}(\mathcal{G},\mathcal{S}^\lambda)),$$

where op from above is defined on the tracial von Neumann algebra $(\mathscr{A},\tau)$. For each $(y,b)\in \mathscr{B}\times\mathscr{Y}$, since $\int_0^\infty P_\lambda G_b^{\lambda,y} P_\lambda d\lambda \le \sigma^2$, by Lemma 3.2.1, there exist $H_b^y \le \mathbb{I}$ such that $\sigma H_b^y \sigma = \int_0^\infty P_\lambda G_b^{\lambda,y} P_\lambda d\lambda$. We see that

$$\sum_b \int_0^\infty P_\lambda G_b^{\lambda,y} P_\lambda d\lambda = \int_0^\infty P_\lambda \left(\sum_b G_b^{\lambda,y}\right) P_\lambda d\lambda = \sigma^2,$$

and hence, by a similar continuity argument as in Lemma 3.2.1, $\{H_b^y\}$ forms a valid POVM in $\mathscr{A}$. We argue that $\{H_b^y\}$ is the desired family of measurements for the corollary. We see that

$$\begin{aligned}\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|\sigma A_a^x (H^y_{g_{xy}(a)})^{op}\sigma|\tau\rangle &= \mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|A_a^x\sigma H^y_{g_{xy}(a)}\sigma|\tau\rangle \\ &= \int_0^\infty P(\lambda)\left(\frac{1}{\tau(P_\lambda)}\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|P_\lambda A_a^x (G^{\lambda,y}_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right)d\lambda \end{aligned} \tag{10.3}$$

where $P=\tau(P_\lambda)$. We claim that

$$\int_0^\infty P(\lambda)\left(\frac{1}{\tau(P_\lambda)}\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|P_\lambda A_a^x (G^{\lambda,y}_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right)d\lambda \ge O(\kappa(\varepsilon-\mathrm{poly}(\delta))-\delta^{\frac{1}{8}}). \tag{10.4}$$

To show the above inequality, we prove the following two inequalities:

$$\int_0^\infty P(\lambda)\cdot\left|\frac{1}{\tau(P_\lambda)}\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|P_\lambda (A_a^{\lambda,x}-A_a^x)(G^{\lambda,y}_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right|d\lambda \le O(\delta^{\frac{1}{8}}), \tag{10.5}$$

$$\int_0^\infty P(\lambda)\cdot\left(\frac{1}{\tau(P_\lambda)}\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|P_\lambda A_a^{\lambda,x} (G^{\lambda,y}_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right)d\lambda \ge \kappa\left(\omega^{co}(\mathscr{G},\mathscr{S})-\mathrm{poly}(\delta)\right). \tag{10.6}$$

For (10.5)

$$\begin{aligned}&\int_0^\infty P(\lambda)\cdot\left|\frac{1}{\tau(P_\lambda)}\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|P_\lambda (A_a^{\lambda,x}-A_a^x)(G^{\lambda,y}_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right|d\lambda \\ &\le \int_0^\infty \mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\left|\langle\tau|P_\lambda (A_a^{\lambda,x}-A_a^x)(G^{\lambda,y}_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right|d\lambda \\ &\le \sqrt{\int_0^\infty \mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\left|\langle\tau|P_\lambda (A_a^{\lambda,x}-A_a^x)^2P_\lambda|\tau\rangle\right|d\lambda}\cdot\sqrt{\int_0^\infty \mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\left|\langle\tau|P_\lambda (G^{\lambda,y}_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right|d\lambda} \\ &\le \sqrt{\int_0^\infty \mathop{\mathbb{E}}_{x\sim\mu}\sum_{a\in\mathscr{A}}\|(A_a^{\lambda,x}-A_a^x)P_\lambda\|_2^2 d\lambda}\cdot 1 \le O(\delta^{\frac{1}{8}}).\end{aligned}$$

For (10.6), by Theorem 5.1.1, $\int_0^\infty P(\lambda)\cdot\omega^{co}(\mathscr{G},\mathscr{S}^\lambda)d(\lambda)\ge\omega^{co}(\mathscr{G},\mathscr{S})-\mathrm{poly}(\delta)$, since κ is convex

$$\begin{aligned}\int_0^\infty P(\lambda)\cdot\left(\frac{1}{\tau(P_\lambda)}\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|P_\lambda A_a^{\lambda,x}(H^y_{g_{xy}(a)})^{op}P_\lambda|\tau\rangle\right)d\lambda &\ge \kappa\left(\int_0^\infty\omega^{co}(\mathscr{G},\mathscr{S}^\lambda)\right) \\ &\ge \kappa\left(\omega^{co}(\mathscr{G},\mathscr{S})-\mathrm{poly}(\delta)\right).\end{aligned}$$

Hence, by the triangle inequality, (10.4) follows from (10.5) and (10.6). Combining (10.3) and (10.4),

$$\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|\sigma A_a^x (H^y_{g_{xy}(a)})^{op}\sigma|\tau\rangle \ge O\left(\kappa(\varepsilon-\mathrm{poly}(\delta))-\delta^{\frac{1}{8}}\right),$$

which concludes the corollary for the special case claimed above.

For the general case where $\mathscr{S} = (\mathscr{L}^2(\mathscr{A},\tau),\sigma|\tau\rangle,\{A_a^x\},\{(B_a^x)^{op}\})$ is a tracially embeddable strategy. By Lemma 4.3.3, the symmetric strategy $\mathscr{S}' = (\mathscr{L}^2(\mathscr{A},\tau),\sigma|\tau\rangle,\{A_a^x\})$ is an $O(\delta)$-synchronous strategy, with

$$\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma A_a^x(B_b^y)^{op}\sigma|\tau\rangle - \langle\tau|\sigma A_a^x(A_b^y)^{op}\sigma|\tau\rangle| \leq O(\sqrt{\delta}),$$

which implies $\omega^{co}(\mathscr{G},\mathscr{S}') \geq \omega^{co}(\mathscr{G},\mathscr{S}) - O(\text{poly}(\delta))$. By the special case proven above, there exists a family of POVMs $\{H_{g^y_{xy}(a)}\}$ such that

$$\mathop{\mathbb{E}}_{(x,y)\sim\rho}\sum_{a\in\mathscr{A}}\langle\tau|\sigma A_a^x(H^y_{g_{xy}(a)})^{op}\sigma|\tau\rangle \geq O(\kappa(\omega^{co}(\mathscr{G},\mathscr{S}) - \text{poly}(\delta)) - \delta^{\frac{1}{8}}). \tag{10.7}$$

which concludes the proof for the general case. □

10.4 The simultaneous quantum low-individual degree test

We describe the simultaneous quantum low-individual degree test in this section. Intuitively, the (p,m,d,k)-simultaneous quantum low-individual degree test is a generalization of the (p,m,d) quantum low-individual degree test defined in Section 10.2, where the goal is to test whether the two provers agree on k global m-variant low-individual degree polynomial $\mathbf{g}:\mathbb{F}_q^m \to \mathbb{F}_q$ with individual degree of at most d. We define the (p,m,d,k)-simultaneous quantum low-individual degree test in Figure 10.3. We have the following lemma regarding the (p,m,d,k)-simultaneous quantum low-individual degree test.

Lemma 10.4.1 (Properties of the (p,m,d,k)-simultaneous quantum low-individual degree test)**.** *Let $p,m,d,k\in\mathbb{N}$, and let $\mathscr{G}^{SLD} = (\mathscr{X}^{SLD},\mathscr{A}^{SLD},\mu^{SLD},D^{SLD})$ be the (p,m,d,k)-simultaneous quantum low-individual degree test specified in Figure 10.3, then the following holds:*

Question label	Question content	Answer format
(Point)	$s\in\mathbb{F}_{2^p}^m$	$(a_0,\cdots,a_{k-1})\in\mathbb{F}_{2^p}^k$
(Dline)	$(j,s_{\text{Dline}})\in[m]\times\mathbb{F}_{2^p}^m$	k degree d polynomial, $\mathbf{f}_i:\mathbb{F}_{2^p}\to\mathbb{F}_{2^p}$, $i\in[k]$, where each $\mathbf{f}_i$ is encoded as $\mathbb{F}_{2^p}^d$
(Aline)	$(j,v,s_{\text{Aline}})\in[m]\times\mathbb{F}_{2^p}^{2m}$	k degree dm polynomial, $\mathbf{g}_i:\mathbb{F}_{2^p}\to\mathbb{F}_{2^p}$, $i\in[k]$, where each $\mathbf{g}_i$ is encoded as $\mathbb{F}_{2^p}^{dm}$

Figure: Q and A format for the (p,m,d,k)-simultaneous quantum low-individual degree test.

Sampling procedure

The sampling procedure for the (p,m,d,k)-simultaneous quantum low-individual degree test is the same as the sampling procedure for the(p,m,d) quantum low-individual degree test given in the proof for Lemma 10.2.1

Verification procedure

1. (Same question label): The provers win iff they output the same answer.
2. (Point) – (DLine): The provers win iff $\mathbf{f}_i(s) = a_i$ for all $i\in[k]$.
3. (Point) – (ALine): The provers win iff $\mathbf{g}_i(s) = a_i$ for all $i\in[k]$.

On any other input, the prover wins automatically.

Figure 10.3: The description for the (p,m,d,k)-simultaneous quantum low-degree test.

- *(Sample complexity): $\mathcal{G}^{SLD}$ is samplable via a $(5, 9+m'+2\cdot m, p)$ typed CL distribution, where $m' = \left\lceil \frac{\log(m)}{p} \right\rceil$.*
- *(Verification complexity): There exists a polynomial time Turing machine $\mathtt{D}^{SLD}$ which implements D^{SLD} and runs in $O(\mathrm{poly}(p,m,d,k))$ time.*
- *(Soundness): There exists a universal constant $1 \geq c_{SLD,1}$ and $0 < c_{SLD,2} \leq 1$ and a function*

$$\eta_{SLD}(p,m,d,k,\varepsilon) = c_{SLD,1}(kdm)^{c_{SLD,1}}(\varepsilon^{c_{SLD,2}} + 2^{-c_{SLD,2}p} + 2^{-c_{SLD,2}md})$$

such that the following holds. Let $\mathcal{S} = (\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{A_a^x\})$ be a synchronous strategy for $\mathcal{G}^{SLD}$ which succeed with probability $1-\varepsilon$. There exist a set of PVM $\{G_{(\mathbf{g}_0,\cdots\mathbf{g}_{k-1})}\} \subseteq \mathscr{A}'$ with outcome labelled by k (potentially the same) $\mathbf{g}_i \in IdPoly(p,m,d)$ such that

$$\mathop{\mathbb{E}}_{s\sim\mathbb{F}_q^m} \sum_{\mathbf{g}_0,\cdots\mathbf{g}_{k-1}\in IdPoly(p,m,d)} \langle\tau| A^{(point,s)}_{(\mathbf{g}_0(s),\cdots,\mathbf{g}_{k-1}(s))} G_{(\mathbf{g}_0,\cdots\mathbf{g}_{k-1})}|\tau\rangle \geq 1-\eta_{SLD}(p,m,d,k,\varepsilon).$$

Although the simultaneous quantum low-degree test is a more sophisticated version of the quantum low-individual degree test. Its sampling procedure is exactly the same as the quantum low-individual degree test, and its verification procedure is essentially repeating the verification procedure for the quantum low-individual degree test k times. Hence, the "sample complexity" and the "decision complexity" follows trivially from Lemma 10.2.1 and Lemma 7.1.1 respectively. To show the "soundness" clause for the above lemma, we recall the following condition from [NW19].

Definition 10.4.2 (Exactly linear functions, Definition 3.17 of [NW19])**.** *Let $m, p \geq 0$. A function $\boldsymbol{f}\colon \mathbb{F}_{2^p}^m \times \mathbb{F}_{2^p}^t \to \mathbb{F}_{2^p}$ is exactly linear in y if it can be written as*

$$\boldsymbol{f}(x,y) = y_1 \cdot \boldsymbol{f}_1(x) + y_2 \cdot \boldsymbol{f}_2(x) \cdots y_{k-1} \cdot \boldsymbol{f}_{k-1}(x),$$

for a set of functions $\{\boldsymbol{f}_i\}_{i\in[k]}$, and we call a function $\boldsymbol{g}\colon \mathbb{F}_{2^p}^k \to \mathbb{F}_{2^p}$ to be exactly linear if

$$\boldsymbol{g}(x) = \sum_{i\in[k]} c_i \cdot x_i,$$

for some $c_i \in \mathbb{F}_{2^p}$, $i \in [k]$.

We recall the following proposition about almost exactly linear individual degree d polynomials.

Proposition 10.4.3 (Almost exactly linear individual degree polynomials, Proposition 3.18 of [NW19])**.** *Suppose that $\boldsymbol{f}(x,y) : \mathbb{F}_{2^p}^m \times \mathbb{F}_{2^p}^k \to \mathbb{F}_{2^p}$ is a polynomial with individual degree of at most d which is not exactly linear in y. Then the probability that, given a uniformly random $z \sim \mathbb{F}_{2^p}^m$, the probability that the polynomial $\boldsymbol{f}_z(y) : \mathbb{F}_{2^p}^k \to \mathbb{F}_{2^p}$, $\boldsymbol{f}_z(y) = \boldsymbol{f}(z,y)$, being exactly linear is at most $\frac{md}{2^p}$.*

We are now ready to give a proof for the "soundness clause" for Lemma 10.4.1, we remark that the proof below is a slightly modified version of [NW19, Theorem 4.43] to account for the difference between the quantum low-degree test and the quantum low-individual degree test.

Proof. Fix constant $p, m, d, k \in \mathbb{N}$. Let $\mathcal{G}^{\mathrm{SLD}}$ be the (p,m,d,k)-simultaneous quantum low-individual degree test, and let $\mathcal{S} = (\mathscr{L}^2(\mathscr{A},\tau), |\tau\rangle, \{A_a^x\})$ be a tracially embeddable strategy for $\mathcal{G}^{\mathrm{SLD}}$ with $\omega(\mathcal{G}^{\mathrm{SLD}}, \mathcal{S}) \geq 1-\varepsilon$. We wish to show that $\mathcal{S}$ can be used as a part of a strategy for the $(p, m+k, d)$ quantum low-individual degree test. For simplicity of notation, for a set of functions $\{\mathbf{f}_i : \mathbb{F}_{2^p}^m \to \mathbb{F}_{2^p}\}_{i\in[k]}$, we define $\mathrm{combine}_{\mathbf{f}}(x,y) : \mathbb{F}_{2^p}^m \times \mathbb{F}_{2^p}^k \to \mathbb{F}_{2^p}$ as the function

$$\mathrm{combine}_{\mathbf{f}}(x,y) = x_0\mathbf{f}_0(y) + \cdots x_{k-1}\mathbf{f}_{k-1}(y). \tag{10.8}$$

In the definition above $\mathbf{f}_i$ could be potentially set as a constant (i.e. $\mathbf{f}_i(x) = c_i$ for some $c_i \in \mathbb{F}_{2^p}$), and for $a \in \mathbb{F}_{2^p}^k$, we write $\text{combine}_a(y) : \mathbb{F}_{2^p}^k \to \mathbb{F}_{2^p}$ as the function $\text{combine}_a(y) = \sum_{i\in[k]} a_i \cdot y_i$. For $o \in [k]$ define $\vec{1}_0 \in \mathbb{F}_{2^p}^o$ to be the vector with all coordinates being $1 \in \mathbb{F}_{2^p}$. We define a strategy for an instance of $(p, m+k, d)$-quantum low-individual degree test depending on $\mathscr{S}$ as follows: We specify the provers' behaviour based on the question label below

1. (Point) Given the question content $(s^1, s^2) \in \mathbb{F}_{2^p}^m \times \mathbb{F}_{2^p}^k$, the prover first performs the strategy $\mathscr{S}$ using the question label "(Point)" and the question content s^1 and obtain points $a = (a_0, \cdots, a_{k-1})$. The prover then returns $\text{combine}_a(s^2)$ as the answer.

2. (DLine) Given the question content (j, s_{Dline}), write $j = (j^1, j^2) \in \mathbb{F}_{2^p}^m \times \mathbb{F}_{2^p}^k$, and let l be the axis on which the axis-parallel line is defined. The prover does the following depending on l

   - If $l \in [m]$, then the prover performs the strategy $\mathscr{S}$ using the question label "(DLine)" and the question content (j^1, s_{Dline}) and obtain k degree-d polynomials $(\mathbf{f}_0, \cdots, (\mathbf{f}_{k-1})$. The prover then returns the degree d polynomial $\text{combine}_{\mathbf{f}}(j^2)$ as the answer.
   - Otherwise, the prover first performs the strategy $\mathscr{S}$ using the question label "(Point)" and the question content j^1 and obtain points $a = (a_0, \cdots, a_{k-1})$. The prover then returns the degree 1 polynomial $\text{combine}_a(j_0^2, \cdots j_{l-m-1}^2, x, j_{l-m+1}^2, j_{m+k-1}^2)$ as their answer.

3. (ALine) Given the question content (j, s_{Aline}), write $j = (j^1, j^2) \in \mathbb{F}_{2^p}^m \times \mathbb{F}_{2^p}^k$, and let l be the coordinates in which the diagonal line is defined. The prover does the following depending on j

   - If $l \in [m]$, then the prover performs the strategy $\mathscr{S}$ using the question label "(ALine)" and the question content (j^1, s_{Aline}) and obtain k degree $d \cdot m$ polynomials $(\mathbf{g}_0, \cdots, (\mathbf{g}_{k-1})$. The prover then returns the degree $d \cdot m + 1$ polynomial $\text{combine}_{\mathbf{g}}(x \cdot \vec{1}_k))$ as the answer.
   - Otherwise, the prover performs the strategy $\mathscr{S}$ using the question label "(Point)" and the question content j^1 and obtain points $a = (a_0, \cdots, a_{k-1})$. The prover then returns the degree 1 polynomial $\text{combine}_a(j_0^2, \cdots, j_{l-m-1}^2, x \cdot \vec{1}_{m+k-l})$ as their answer.

Since $\mathscr{S}$ succeeds in $\mathscr{G}^{\text{SLD}}$ with probability at least $1 - \text{poly}(\varepsilon)$, given the same question label and content, the probability that the provers give the same answer is at least $1 - \text{poly}(\varepsilon)$, as well as given the label pair "(point)" and "(DLine)" (resp. "(ALine)"), the answer pair satisfies $\mathbf{f}_i = a_i$ (resp. $\mathbf{g}_i = a_i$). Using this, one can see that the above strategy for the $(p, m+k, d)$-quantum low-individual degree test succeeds with probability at least $1 - \text{poly}(\varepsilon)$. For $(s^1, s^2) \in \mathbb{F}_{2^p}^m \times \mathbb{F}_{2^p}^k$ and $v \in \mathbb{F}_{2^p}$, we define the measurement $A^{(\text{point},s^1)}_{[\text{combine}_a(s^2)|v]}$ to be the data processing measurement in which the prover applies the function $\mathbf{f}(a_0, \cdots, a_{k-1}) = \text{combine}_a(s^2)$ to the measurement outcome of $P^{(\text{point},s^1)}_{a_0,\cdots,a_{k-1}}$. By Theorem 10.3.1, there exists a measurement $\{H_{\mathbf{f}}\} \subseteq \mathscr{A}'$ with outcomes $\mathbf{f}$ being $m+k$-variate polynomial with individual degree of at most d such that

$$\mathop{\mathbb{E}}_{(s^1,s^2)\sim\mathbb{F}_{2^p}^m\times\mathbb{F}_{2^p}^k} \sum_{\mathbf{f}\in\text{IdPoly}(p,m+1,d)} \langle\tau| A^{(\text{point},s^1)}_{[\text{combine}_a(s^2)|\mathbf{f}(s^1,s^2)]} \cdot H_{\mathbf{f}} |\tau\rangle \geq 1 - \eta_{\text{LD}}(p, m+k, d, \varepsilon). \tag{10.9}$$

Now we wish to show that the measurement outcome $\mathbf{f}(x, y)$ from $\{H_{\mathbf{f}}\}$ is exactly linear in y with high probability. Fix a g such that the measurement outcome from $\mathbf{f}$ is not exactly linear. For a fixed $s^1 \in \mathbb{F}_{2^p}^m$ and a fixed measurement outcome $a = (a_0, \cdots, a_{k-1})$ from $P^{(\text{point},s^1)}_{(a_0,\cdots,a_{k-1})}$, by construction, the function $\text{combine}_a(y) : \mathbb{F}_{2^p}^k \to \mathbb{F}_{2^p}$ is exactly linear. However, by Proposition 10.4.3 the probability that $\mathbf{f}_{s^1}(y) = \mathbf{f}(s^1, y)$ is exactly linear is at most $\frac{md}{2^p}$. As a result, for a uniformly chosen $s_1 \sim \mathbb{F}_{2^p}$, the probability that $\text{combine}_a = \mathbf{f}_{s^1}$ is at most $\frac{md}{2^p}$. Hence by Lemma 7.1.4, for

$(s^1, s^2) \sim \mathbb{F}_{2^p}^m \times \mathbb{F}_{2^p}^k$, the probability that $\text{combine}_a(s^2) = \mathbf{f}(s^1, s^2)$ is at most $\frac{md}{2^p} \cdot \frac{kd}{2^p}$ regardless of the measurement outcome for a. Combining the above fact with Equation (10.9), we see that the output of $\{H_{\mathbf{f}}\}$ is exactly linear with probability at least $1 - \eta_{\text{LD}}(p, m+k, d, \varepsilon) - \left(\frac{md}{2^p}\right)^2 - \frac{m \cdot kd}{2^{2p}}$.

Define the measurement $\{G_{\mathbf{g}_0, \cdots, \mathbf{g}_{k-1}}\} \subseteq \mathscr{A}'$ with the outcome set the same as $\{H\}$ as follows: The prover first measures according to $\{H\}$ to receive a polynomial $\mathbf{f}(x, y)$. If $\mathbf{f}(x, y)$ is exactly linear in y, it can be written as $\mathbf{f}(x, y) = \sum_k y_i \mathbf{g}_i(x)$, such that each $\mathbf{g}_i$ is an m-variant polynomial with individual degree at most d. In this case, G outputs the polynomials $\{\mathbf{g}_i\}_{i \in [k]}$ as the output. If the measurement output from $\{H\}$ is not exactly linear, G simply outputs k random m-variate polynomials with individual degree d, $\{\mathbf{g}_i\}_{i \in [k]}$. Since $\text{combine}_{\mathbf{g}}$ is equal to $\mathbf{g}$ whenever $\mathbf{g}$ is exactly linear by definition, by replacing H with G on Equation (10.9), we see that

$$A^{(\text{point}, s^1)}_{[\text{combine}_a(s^2) | \mathbf{f}(s^1, s^2)]} \simeq_{\eta_{\text{LD}}(p, m+k, d, \varepsilon) + \frac{m \cdot kd}{2^{2p}}} G_{\text{combine}_{\mathbf{g}} = \mathbf{f}}, \tag{10.10}$$

where $\simeq$ is with respect to the distribution $\mathbb{E}_{(s^1, s^2) \sim \mathbb{F}_{2^p}^m \times \mathbb{F}_{2^p}^k}$ and the state $|\tau\rangle$. Now, for any fixed s^1 and $a = (a_i)_{i \in [k]}$, if $\mathbf{g}_i(s^1) \neq a_i$ for any i, then the polynomial $\text{combine}_{\mathbf{g}}(s^1, y)$ and $\text{combine}_a(y)$ are not equal and hence again by Lemma 7.1.4, the probability is at most $\frac{1}{2^p}$ (since both $\text{combine}_{\mathbf{g}}(s^1, y)$ and $\text{combine}_a(y)$ are a multi-linear function). This implies that

$$A^{(\text{point}, s^1)}_{a_0, \cdots, a_{k-1}} \simeq_{\eta_{\text{LD}}(p, m+k, d, \varepsilon) + \frac{m \cdot kd}{2^{2p}} + \frac{1}{2^p}} G_{\text{combine}_{\mathbf{g}} = \mathbf{f}}.$$

Hence, the "soundness" condition follows by setting

$$\eta_{\text{SLD}}(p, m, d, k, \varepsilon) = \eta_{\text{LD}}(p, m+k, d, \varepsilon) + \frac{m \cdot kd}{2^{2p}} + \frac{1}{2^p}. \qquad \square$$

We remark that since the above proof does not rely on the strategy $\mathscr{S}$ being synchronous, and thus the "commuting operator soundness" condition in Lemma 10.4.1 also holds for general tracially embeddable strategies strategy (by replacing Theorem 10.3.1 above with Theorem 10.3.2). However, for simplicity of analysis, working with synchronous strategies is sufficient for the remainder of this thesis.

CHAPTER 11

Answer reduction part 2: Oracularization and time-bounded probabilistically checkable proofs by proximity (PCPP)

In this section, we give a proof for Proposition 7.4.3. The goal of the answer reduction protocol is to transform a synchronous CL verifier into another synchronous CL verifier with a more efficient verification complexity. We remark that the transformation used in this section is the same as the one given in [JNV+22a, Section 10]. We first give some intuition for the answer reduction transformation in the first section of this chapter.

11.1 Overview of the Answer reduction protocol

On a high level, the goal of the answer reduction protocol is to let the verifier delegate the task of computing $\mathtt{D}(n, x, y, a, b)$ to the provers. Of course, since the provers are by definition dishonest, the verifier cannot simply give this task to the provers. One important observation about an interactive protocol which makes the answer reduction protocol possible is that the verifier actually does not care how the computation step is being performed, *he only cares whether* $\mathtt{D}(n, x, y, a, b)$ *outputs 1 at the end of the computation step!* Hence, the goal for the verifier is to design a protocol in which the provers can somehow output "sufficient evidence" to show that they have, indeed, run the computation step of $\mathtt{D}(n, x, y, a, b)$ honestly.

Fortunately, the verifier can already use a probabilistically checkable proof (PCP), a common tool in the computer science literature [ALM+98]. Roughly speaking, let $\mathtt{TM}$ be a two-input Turing machine which runs in $O(\exp(n))$ time and $x \in \{0,1\}^*$ be a string with $|x| = O(\text{polylog}(n))$. Existing PCP in the computer science literature allows a polylogarithmic-time verifier, with the help of two (computationally unbounded) provers, to verify that there exists a string $a \in \{0,1\}^*$ with $|a| = O(\text{poly}(n))$ such that $\mathtt{TM}(x, a) = 1$. This construction can be easily modified to hold for a pair of strings (x, y), which is polylogarithmic-sized as the initial input, and a pair of polynomial-size strings (a, b), which is polynomial-sized as the initial input. We remark in this case, since $|a|$ is exponential in size, a polynomial-time verifier cannot process the entire string a even if he receives it from the prover! However, there are several challenges with directly using a PCP construction within the answer reduction-procedure, which we list below:

1. For a prover to compute the given PCP instance, it needs both question labels (x, y). This is a problem in the non-local game setting, since each prover is expected to receive only its own question label.

2. The PCP construction only checks whether there exists an answer pair (a, b) which causes the verifier to accept. In this case, the verifier also needs to check that each answer within the answer pair (a, b) depends only on its corresponding question label (x or y); i.e. the prover cannot generate the answer a based on both the question labels (x, y). This is a bigger problem for the verifier than it might at first appear, since, as previously mentioned, the verifier does not have the runtime to even process the answer labels a and b.

3. Lastly, the PCP construction must also be a CL sampleable game in order to be used as a part of the proof for the gap compression theorem.

In order to address the first problem, before applying the PCP procedure, a transformation known as *oracularization* is first applied. This transformation was first introduced in [JNV+22a, Section 9], and is also part of the answer reduction procedure in the "gapless compression" introduced in [MNY22]. The goal of this transformation is to give both provers the two question labels (x, y) and force them to generate the same answer pair (a, b) in such a way that the answer label a (resp. b) only depends on the question label x (resp. y). To ensure consistency, the provers will sometimes give one prover only one of the two question labels in order to perform a consistency check with the other prover who receives both question labels.

Unfortunately, as pointed out by point two above, since the verifier cannot process the entire question label a and b, the consistency check mentioned is also not as straightforward as it might have initially seemed. To keep the verification complexity low, the provers are expected to encode the answers a and b as a low-individual degree polynomial using the generalized Reed-Muller code introduced in Section 10.1. In this case, the consistency test for the verifier becomes verifying that the two provers share the same low-individual degree polynomial, which can be done through the quantum low-individual degree test given in Section 10.2.

To tackle the third problem, [NW19] uses a special type of PCP known as a probabilistically checkable proof of proximity (PCPP), which allows one to check whether a specific string a satisfies $\mathtt{TM}(x, a) = 1$ (rather than merely asserting the existence of such a string using a standard PCP). In this thesis, we use the tailor-made PCPP protocol constructed in [JNV+22a, Section 10], which reduces the proof checking task to an instance of the *simultaneous quantum individual low-degree test*, where the simultaneous quantum individual low-degree test (SLDT), in essence, is a parallel repeated version of the quantum individual low-degree test, which is designed to test if the provers share multiple low-individual degree polynomials. As shown later in this section, since the SLDT has the same sampling procedure as a regular quantum individual low-degree test, this solves the last problem listed above by Lemma 10.2.1.

We organize the remainder of this chapter as follows. In Section 11.2, we first go over the oracularization transformation mentioned above from [JNV+22a, Section 9] and show that the completeness/soundness properties from the tensor product model also hold for the commuting operator model. In Section 11.4, we give a summary of result of the PCPP construction from [JNV+22a, Section 10] which will be the bases of the answer reduction protocol that shows Proposition 7.4.3.

11.2 Oracularization

In this subsection, we recall the oracularization transformation used in [JNV+22a, Section 9], and show that the appropriate completeness/soundness conditions also hold for the commuting operator model. Given a non-local game $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$, we define the corresponding oracularization transformation in Figure 11.1.

We have the following lemma regarding the oracularization transformation for the game $\mathcal{G}$. We remark that the "soundness" condition for the below lemma also preserves the normal (non-synchronous) value of the game.

Lemma 11.2.1 (Properties related to the oracularization transformation)**.** *Let $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$ be a non-local game, and let $\mathcal{G}^{Ora}$ be the oracularization transformation for the game $\mathcal{G}$, then for $t \in \{*, co\}$, the following holds*

- *(Completeness): If there exists a perfect oracularizable strategy for $\mathcal{G}$ in model t, then there exists a perfect oracularizable strategy for $\mathcal{G}^{Ora}$ in model t.*
- *(Soundness): There exists a polynomial $\mathbf{P}^{Ora}(\varepsilon)$ such that*

$$\omega^t(\mathcal{G}^{Ora}) \geq 1-\varepsilon \Longrightarrow \omega^t(\mathcal{G}) \geq 1-\mathbf{P}^{Ora}(\varepsilon).$$

- *(Sample complexity) If $\mathcal{G}$ is samplable via a (k, m, p) CL distribution, then $\mathcal{G}^{Ora}$ is samplable via a $(k+1, m+4, p)$ CL distribution.*

Proof. Let $\mathcal{G}$ and $\mathcal{G}^{\text{Ora}}$ be the non-local game as specified in the lemma and fix $t \in \{*, co\}$. We also shorten the question label "(Oracularization)" to "(Ora)" in this proof for convenience.

For the "completeness" property in the lemma statement, let $\mathcal{S}$ be a perfect oracularizable strategy for the game $\mathcal{G}$. By Theorem 3.5.1, $\mathcal{S}$ is synchronous and hence by Lemma 4.3.2 can also be assumed to be a projective strategy defined by $\mathcal{S} = (\mathcal{L}^2(\mathcal{A}, \tau), |\tau\rangle, \{P_a^x\})$.

We construct a perfect (synchronous) oracularizable strategy on the Hilbert space $\mathcal{L}^2(\mathcal{A}, \tau)$ as follows: for all $(x, y, a, b) \in \mathcal{X}^2 \times \mathcal{A}^2$, define the synchronous strategy $\mathcal{S}^{\text{Ora}} = (\mathcal{L}^2(\mathcal{A}, \tau), |\tau\rangle, \{M_a^x\})$ for the game $\mathcal{G}^{\text{Ora}}$ as follows:

$$M_{a_A}^{(\text{Prover, A}),x} = P_{a_A}^x, \qquad M_{b_B}^{(\text{Prover, B}),y} = P_{b_B}^y, \qquad M_{(a,b)}^{(\text{Ora}),(x,y)} = P_a^x P_b^y.$$

We first show that $\mathcal{S}^{\text{Ora}}$ is projective. Since $\mathcal{S}$ is a projective strategy, both $M_{a_A}^{(\text{Prover, A}),x}$ and $M_{b_B}^{(\text{Prover, B}),y}$ are projective. By the definition of an oracularizable strategy Definition 4.4.1,

Question label	Question content	Answer format
(Prover, A)	$x \in \mathcal{X}$	$a_A \in \mathcal{A}$
(Prover, B)	$y \in \mathcal{X}$	$b_B \in \mathcal{A}$
(Oracularization)	$(x, y) \in \mathcal{X}^2$	$(a, b) \in \mathcal{A}$

Figure: Q and A format for the oracularization transformation for $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$

Sampling procedure

1. Sample $(x, y) \sim \mu$, and $(n_0, n_1) \in \{(\text{Prover, A}), (\text{Prover, B}), (\text{Oracularization})\}^2$.
2. Send the question label and question content corresponding to n_0 to one of the provers, and send the question content corresponding to the n_1 to the other prover.

Verification procedure

1. (Oracularization) – (Oracularization): The provers win iff they output the same answer and $D(x, y, a, b) = 1$.
2. (Prover, P) – (Prover, P) for P $\in$ {A, B}: The provers win iff they output the same answer.
3. (Prover, A) – (Oracularization): The provers win iff $D(x, y, a, b) = 1$ and $a = a_A$.
4. (Prover, B) – (Oracularization): The provers win iff $D(x, y, a, b) = 1$ and $a = b_A$.
5. (Prover, A) – (Prover, B): The provers win automatically.

On any other input, the prover win automatically.

Figure 11.1: The description for the oracularization transformation $\mathcal{G}_\perp$ for the game $\mathcal{G} = (\mathcal{X}, \mathcal{A}, \mu, D)$.

for (x,y) with $\mu(x,y) > 0$, $[P_b^y, P_a^x] = 0$. Since the question label (Oracularization), (x,y) only occurs whenever $\mu(x,y) > 0$, this shows that the measurement operator $M_{(a,b)}^{(\text{Oracularization}),(x,y)}$ is projective; hence $\mathscr{S}^{\text{Ora}}$ is a projective strategy.

The fact that $[P_b^y, P_a^x] = 0$ whenever $\mu(x,y) > 0$ also implies that the set of measurement operators $\{M_a^{(\text{Prover, A}),x}\}_{a\in\mathscr{A}} \cup \{M_b^{(\text{Prover, B}),y}\}_{b\in\mathscr{A}} \cup \{M_{(a,b)}^{(\text{Oracularization}),(x,y)}\}_{(a,b)\in\mathscr{A}^2} \subseteq \mathscr{A}$ pairwise commute with each other whenever $\mu(x,y) > 0$. This shows that $\mathscr{S}^{\text{Ora}}$ is oracularizable.

To show that $\mathscr{S}^{\text{Ora}}$ is perfect, we verify each of the possible question pairs below:

- (Same label): This follows because the strategy $\mathscr{S}^{\text{Ora}}$ is projective and uses the tracial state $|\tau\rangle$ as a part of the strategy.
- (Prover, A) – (Ora): Given $(x,y) \in \mathscr{X}$ with $\mu(x,y) > 0$, and $a, a_A, b \in \mathscr{A}$, the probability of outputting $((a,b),a_A)$ given the question pair $(((\text{Prover, A}),x))$ is

  $$\langle\tau|M_{(a,b)}^{(\text{Ora}),(x,y)}(M_{a_A}^{(\text{Prover, A}),x})^{op}|\tau\rangle = \langle\tau|(P_a^x P_b^y)(P_{a_A}^x)^{op}|\tau\rangle = \langle\tau|(P_b^y P_a^x)P_{a_A}^x|\tau\rangle$$

  where the third equality follows from $A|\tau\rangle = A^{op}|\tau\rangle$. Since $\mathscr{S}$ is a projective, perfect strategy, the resulting answer $((a,b),a_A)$ must satisfy $D(x,y,a,b) = 1$ and $a = a_A$.
- (Prover, B) – (Ora): This follows from the same argument as above.

This shows the completeness clause in the lemma.

For the "soundness" property in the lemma statement, suppose that $\omega^t(\mathscr{G}^{\text{Ora}}) \geq 1-\varepsilon$. Let $\mathscr{S}' = (\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{A_a^{v,x}\}, \{\left(B_a^{V,y}\right)\})$ be a tracially embeddable strategy for $\mathscr{G}^{\text{Ora}}$ in model t such that $\omega(\mathscr{G}^{\text{Ora}}, \mathscr{S}) \geq 1-\varepsilon$. By the sampling procedure as specified in Figure 11.1, the game $\mathscr{G}^{\text{Ora}}$ is $\frac{1}{3}$-balanced. Hence, by Lemma 5.3.1, there exist a symmetric strategy $\mathscr{S}^{\text{sym}} = (\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{P_a^{V,y}\})$ such that

$$\omega(\mathscr{G}^{\text{Ora}}, \mathscr{S}^{\text{sym}}) \geq 1-\varepsilon-(3\varepsilon)^{\frac{1}{4}} = 1-O(\text{poly}(\varepsilon)). \tag{11.1}$$

We wish to argue that the strategy $\mathscr{S} = (\mathscr{L}^2(\mathscr{A},\tau), \sigma|\tau\rangle, \{P_a^{(\text{Prover, A}),x}\}, \{P_b^{(\text{Prover, B}),y}\})$ for the game $\mathscr{G}$ satisfies $\omega(\mathscr{G},\mathscr{S}) \geq 1-O(\text{poly}(\varepsilon))$. For simplicity of notation, we write $P_a^{(\text{Prover, Q}),x}$ as $P_a^{\text{Q},x}$ for $\text{Q} \in \{\text{A},\text{B}\}$. Since the question label pair (Ora) – (Prover A) and (Ora) – (Prover B) are selected with probability $1/9$, we have

$$P_{a,b}^{(\text{Ora}),(x,y)} \approx_{O(\text{poly}(\varepsilon))} (P_a^{\text{A},x})^{op}, \qquad P_{a,b}^{(\text{Ora}),(x,y)} \approx_{O(\text{poly}(\varepsilon))} (P_b^{\text{B},y})^{op},$$

over the distribution $(x,y) \sim \mu$. Since P is projective, by Lemma 3.3.3

$$P_{a,b}^{(\text{Ora}),(x,y)} \approx_{O(\text{poly}(\varepsilon))} (P_a^{\text{A},x})^{op}, \qquad P_{a,b}^{(\text{Ora}),(x,y)} \approx_{O(\text{poly}(\varepsilon))} (P_b^{\text{B},y})^{op}. \tag{11.2}$$

Since the question label pair (Prover A) – (Prover A) is also selected with probability $1/9$,

$$P_a^{\text{A},x} \approx_{O(\text{poly}(\varepsilon))} (P_a^{\text{A},x})^{op}. \tag{11.3}$$

Since $\{P\}$ is a projective strategy, $P \leq \mathbb{I}$. Hence

$$\begin{aligned} P_{a,b}^{(\text{Ora}),(x,y)} = (P_{a,b}^{(\text{Ora}),(x,y)})^2 &\approx_{O(\text{poly}(\varepsilon))} P_{a,b}^{(\text{Ora}),(x,y)}(P_b^{\text{B},y})^{op} \\ &\approx_{O(\text{poly}(\varepsilon))} (P_a^{\text{A},x})^{op}(P_b^{\text{B},y})^{op}, \\ &\approx_{O(\text{poly}(\varepsilon))} P_a^{\text{A},x}(P_b^{\text{B},y})^{op}, \end{aligned} \tag{11.4}$$

where the first approximation follows from Lemma 3.3.4 with C being $\{P_{a,b}^{(\text{Ora}),(x,y)}\}$, and treating the set $\mathscr{C}$ as the singleton set, and the second and third approximation follows from Lemma 3.3.4

with C being $\{(P_b^{\mathrm{B},y})^{op}\}$ in conjunction with (11.2) and (11.3) respectively. Since P is projective and $P_a^{\mathrm{A},x}$ commutes with $(P_b^{\mathrm{B},y})^{op}$, we have

$$P_a^{\mathrm{A},x}(P_b^{\mathrm{B},y})^{op} = \left(P_a^{\mathrm{A},x}(P_b^{\mathrm{B},y}\right)^{op})^2.$$

Hence

$$\begin{aligned}
&\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b} |\langle\tau|\sigma P_{a,b}^{(\mathrm{Ora}),(x,y)}\sigma|\tau\rangle - \langle\tau|\sigma\left(P_a^{\mathrm{A},x}(P_b^{\mathrm{B},y})^{op}\right)\sigma|\tau\rangle| \\
&\leq \mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b} |\langle\tau|\sigma P_{a,b}^{(\mathrm{Ora}),(x,y)}\left(P_{a,b}^{(\mathrm{Ora}),(x,y)} - P_a^{\mathrm{A},x}(P_b^{\mathrm{B},y})^{op}\right)\sigma|\tau\rangle| + \\
&\qquad \mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b} |\langle\tau|\sigma\left(P_{a,b}^{(\mathrm{Ora}),(x,y)} - P_a^{\mathrm{A},x}(P_b^{\mathrm{B},y})^{op}\right)P_a^{\mathrm{A},x}(P_b^{\mathrm{B},y})^{op}\sigma|\tau\rangle| \\
&\leq \sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}\langle\tau|\sigma P_{a,b}^{(\mathrm{Ora}),(x,y)}\sigma|\tau\rangle}\sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma\left(P_{a,b}^{(\mathrm{Ora}),(x,y)} - P_a^{\mathrm{A},x}(P_b^{\mathrm{B},y})^{op}\right)^2\sigma|\tau\rangle|} + \\
&\qquad \sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\langle\tau|\sigma\left(P_{a,b}^{(\mathrm{Ora}),(x,y)} - P_a^{\mathrm{A},x}(P_b^{\mathrm{B},y})^{op}\right)^2\sigma|\tau\rangle|}\cdot\sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}\langle\tau|\sigma P_a^{\mathrm{A},x}(P_b^{\mathrm{B},y})^{op}\sigma|\tau\rangle} \\
&= 2\sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu}\sum_{a,b}|\left(P_{a,b}^{(\mathrm{Ora}),(x,y)} - P_a^{\mathrm{A},x}(P_b^{\mathrm{B},y})^{op}\right)\sigma|\tau\rangle|^2} = O(\mathrm{poly}(\varepsilon)),
\end{aligned} \tag{11.5}$$

where the second line follows from the triangle inequality and the third line follows from Cauchy-Schwarz. The last line follows from $\{P_{a,b}^{(\mathrm{Ora}),(x,y)}\}$, $\{P_a^{\mathrm{A},x}\}$ and $\{(P_b^{\mathrm{B},y})^{op}\}$ being three sets of POVMs and (11.4).

Finally, since the question label (Ora) is selected with probability 1/3 by the first prover, by (11.1), with expectation over the distribution μ, the measurement $\langle\tau|\sigma P_{a,b}^{(\mathrm{Ora}),(x,y)}\sigma|\tau\rangle$ will produce an answer (a,b) such that $D(x,y,a,b)$ with probability at least $1-O(\mathrm{poly}(\varepsilon))$ (or else the players will lose $\mathcal{G}^{\mathrm{Ora}}$). This, in conjunction with (11.5), shows the "soundness" clause in the lemma.

For the "sample complexity" property in the lemma statement, assume μ, the input distribution for $\mathcal{G}$, is samplable via a (k,m,p) CL distribution. Let $\mathtt{L}^A$ and $\mathtt{L}^B$ be the two (k,m,p) conditional linear functions used to define μ. We show that $\mathcal{G}^{\mathrm{Ora}}$ is $(k+1,m+4,p)$ CL samplable by using the series composition of two sets of CL functions (given in Definition 7.2.3). Let $V_0 = \mathbb{F}_{2^p}^4$ and write V, and $V_{>0} = \mathbb{F}_{2^p}^m$, we define the two $(k+1,m+4,p)$ conditional linear functions $\mathtt{L}^{A,\mathrm{Ora}}$ and $\mathtt{L}^{A,\mathrm{Ora}}$ as follows:

For simplicity of notation, we assume elements in V_0 are represented under the canonical representation (i.e. as elements of $\{0,1\}^{4\cdot p}$). Write every elements $s \in V_0$ as $s = (s_0,s_1,s_2,s_3,s_4)$, where $s_i \in \{0,1\}$ for $i \in [4]$ and $s_4 \in \{0,1\}^{4\cdot(p-1)}$. Define

$$\mathtt{L}_{0,0}^{A,\mathrm{Ora}}(s_0,s_1,s_2,s_3,s_4) = (s_0,s_1,\vec{0}), \quad \mathtt{L}_{0,0}^{B,\mathrm{Ora}}(s_0,s_1,s_2,s_3,s_4) = (s_2,s_3,\vec{0}),$$

where $\vec{0}$ is the all 0 string in $\{0,1\}^{4\cdot p-2}$, we define

$$\mathtt{L}_{>0,(0,0,\vec{0})}^{A,\mathrm{Ora}} = \mathtt{L}_{>0,(0,0,\vec{0})}^{B,\mathrm{Ora}} = \mathtt{L}^A \quad \mathtt{L}_{>0,(0,1,\vec{0})}^{A,\mathrm{Ora}} = \mathtt{L}_{>0,(0,1,\vec{0})}^{B,\mathrm{Ora}} = \mathtt{L}^B \quad \mathtt{L}_{>0,(1,0,\vec{0})}^{A,\mathrm{Ora}} = \mathtt{L}_{>0,(1,0,\vec{0})}^{B,\mathrm{Ora}} = \mathbb{I},$$

where $\mathbb{I}$ is the identity function on $V_{>0}$. Intuitively, $(0,0)$ label corresponds to the question label "Prover A"; the $(0,1)$ label corresponds to "Prover B"; and the $(1,0)$ label corresponds to "Oracularization". If the prover receives the label "Oracularization", we give the entire seed to that prover in order for them to compute the question label (x,y) themselves. We remark that we can treat the label $(1,1)$ as a free win for the provers, which occurs only with a constant probability; this raises the soundness condition by only a constant factor. This completes the argument for the "sample complexity" claim. □

Throughout the thesis, the "completeness" condition for almost all transformations of games always has a requirement that preserves perfect *oracularizable* strategies. As remarked in Section 4.4, the only use for this requirement in this thesis is to show the "completeness" condition for the oracularization transformation to hold.

11.3 Time-bounded classical PCPP

We recall the following PCPP protocol given in [JNV+22a, Section 10] (which is a modification of the work by [Har04]). Since we do not modify the construction from [JNV+22a], we do not go through the details of this construction in this thesis, and instead refer to the original reference for more details. Recall from the preliminaries that given a string $a \in \{0,1\}^n$, one can encode a into a low-individual-degree $\log(n)$-variate polynomial with a low-individual degree of 2 using the generalize Reed-Muller encoding given in (10.1). The following theorem is the main result of [JNV+22a, section 10.1-10.5].

Theorem 11.3.1 (Time bounded decider PCPP)**.** *Let* $\mathtt{D}$ *be a decider for a CL verifier* $\mathscr{V}$*, and* $\alpha \in \mathbb{N}$*. There exist two Turing machines* $(\mathtt{PCPParameter}_\alpha, \mathtt{ComputePCP}_\alpha)$*.* $\mathtt{PCPParameter}_\alpha$ *takes, as input* $n \in \mathbb{N}$ *and outputs a tuple of parameters* $(m^{ans}, m^{PCPP}, g, p)$ *such that the following holds:*

- $\mathsf{TIME}_{\mathtt{PCPParameter}_\alpha} = O(\mathrm{poly}(\alpha, \log(n)))$
- $m^{ans}, g = O(\mathrm{poly}(\alpha, \log(n)))$
- $m^{PCPP} = 5 \cdot m^{ans} + 5 + g$*, where* g *is padded such that* m^{PCPP} *is of the form* 2^i *for some integer* i*.*
- *Let* $1 \geq c_{SLD,1}$ *and* $0 < c_{SLD,2} \leq 1$ *be the universal constant defined within Lemma 10.4.1. The field size* p *is chosen to be the smallest integer which satisfies the following:*
  1. $p \geq \frac{\alpha c_{SLD,2} + 3c_{SLD,1}) \cdot \log(g)}{c_{SLD,2}}$,
  2. $\frac{(2+5p) \cdot m^{ans}}{2^p} < \frac{1}{2}$,
  3. $\frac{p m^{PCPP}}{2^p} \leq s^{-c_{SLD,2}\alpha}$,
  4. 2^p *is divisible by* m^{PCPP}*.*

  By the choice of parameters, one can check that $p = O(\mathrm{polylog}(\alpha, \log(n)))$

The Turing machine $\mathtt{ComputePCPP}_\alpha$ *takes, as input, another Turing machine* $\langle \mathtt{D} \rangle$*, a natural number* $n \in \mathbb{N}$ *and a pair of strings* (x, y) *such that* $|x|, |y| \leq \log^\alpha(n)$*, and outputs a description of a* m^{PCPP}*-variate polynomial* $\mathbf{g}_\mathtt{D} \in IdPoly(p, m^{PCPP}, p)$*.* $\mathtt{ComputePCPP}_\alpha$ *has the following properties:*

- *(Time complexity:* $\mathtt{ComputePCPP}_\alpha$ *takes time* $O(\mathrm{poly}(\alpha, \log^\alpha(n), |\langle \mathtt{D} \rangle|))$*.*
- *(The complexity of the PCPP formula): The description of* $\mathbf{g}_\mathtt{D}$ *can be represented using* $O(\mathrm{poly}(\alpha, \log(n)))$ *bits, and evaluating* $\mathbf{g}_\mathtt{D}$ *at a single point takes time* $O(\mathrm{poly}(\alpha, \log(n)))$*.*

Furthermore, if $|\mathtt{D}| = \mathrm{poly}(\alpha)$ *and there exist two strings* $(a, b) \in \{0,1\}^{n^\alpha}$ *such that* $D(n, x, y, a, b) = 1$ *with*

$$\mathsf{TIME}_\mathtt{D}(n, x, y, a, b) \leq n^\alpha.$$

Then there exist five polynomials $\overline{\mathbf{g}}_a, \overline{\mathbf{g}}_b, \overline{\mathbf{g}}_{w_0}, \overline{\mathbf{g}}_{w_1}, \overline{\mathbf{g}}_{w_2} \in IdPoly(p, m^{ans}, p)$ *such that the following holds:*

- $\overline{\mathbf{g}}_a$ *is the Reed-Muller encoding of* $enc_\Gamma(a)$*, where* enc_Γ *is the encoding map for the "padded version" of a which maps any string with length at most* n^α *to a string with* $2^{m^{ans}}$*. The padded encoding map is given in [JNV+22a, Proposition 10.19]. The Reed-Muller encoding is defined over* $\mathbb{F}_{2^p}^{m^{PCPP}}$*.*

- $\overline{\mathbf{g}}_b$ *is the Reed-Muller encoding of* $enc_\Gamma(b)$*, defined similarly as* $\overline{\mathbf{g}}_a$*.*

- *For every* $s \in \mathbb{F}_{2^p}^{m^{PCPP}}$*, partition* $s = (s_0, \cdots, s_4, b_0, \cdots, b_4, z)$*, where for* $i \in [5]$*,* $s_i \in \mathbb{F}_{2^p}^{m^{ans}}$*,* $b_i \in \mathbb{F}_{2^p}$ *and* $z \in \mathbb{F}_{2^p}^s$*. Define the polynomial* $\mathbf{g}_{\mathtt{D}}^{Full} \in IdPoly(p, m^{PCPP}, d)$ *as the polynomial with individual degrees at most* 32 *as*

$$\mathbf{g}_{\mathtt{D}}^{Full}(s) = \mathbf{g}_{\mathtt{D}}(s)\cdot(\overline{\mathbf{g}}_a(s_0)-b_0)\cdot(\overline{\mathbf{g}}_b(s_1)-b_1)\cdot(\overline{\mathbf{g}}_{w_0}(s_2)-b_2)\cdot(\overline{\mathbf{g}}_{w_1}(s_3)-b_3)(\overline{\mathbf{g}}_{w_2}(s_4)-b_4). \quad (11.6)$$

Moreover, for every $s \in \{0,1\}^{m^{PCPP}} \subseteq \mathbb{F}_{2^p}^{m^{PCPP}}$*,* $\mathbf{g}_{\mathtt{D}}^{Full}(s) = 0$*.*

We remark that the parameter requirement for (m^{ans}, g, p) given in the above theorem follows according to [JNV+22a, Definition 10.22], and indeed the individual degree for each polynomials in the above definition is at most d. Intuitively, $\mathbf{g}_a$ and $\mathbf{g}_b$ in the above theorem encodes the answers given by the provers for the game $\mathscr{G}$, and $\overline{\mathbf{g}}_{w_0}$, $\overline{\mathbf{g}}_{w_1}$, $\overline{\mathbf{g}}_{w_2}$ is an encoding for the 3-SAT instances from the computation steps of $\mathtt{D}$ via the well known Cook-Levin encoding. In the theorem above, a polylog time bounded verifier can only compute a description of the polynomial $\mathbf{g}_{\mathtt{D}}$ given $\mathtt{D}$ and the question pair (x, y), and can only evaluate $O(\text{polylog}(n))$ points from $\mathbf{g}_{\mathtt{D}}$. The verifier must somehow query the potentially dishonest provers for the existence of $\mathbf{g}_a, \mathbf{g}_b, \mathbf{g}_{w_0}, \mathbf{g}_{w_1}, \mathbf{g}_{w_2}$ to confirm the existence of such answer pair (a, b) such that $\mathtt{D}(n, x, y, a, b) = 1$.

On a high level, in the PCPP protocol, the verifier can compute the low-individual degree polynomial $\mathbf{g}_{\mathtt{D}}$ since he has access to the description of the decider $\mathtt{D}$. Then, the verifier can, with the help of the prover, verify the existence of the five low-individual degree polynomials guaranteed by the above theorem, and as well as the resulting $\mathbf{g}_{\mathtt{D}}^{\text{Full}}$ is 0 on all the points within the "0/1 subcube". Verifying the existence of the five low-individual degree is easy via the simultaneous quantum low-individual degree test given on the last subsection. The verifier can use the following lemma to check that $\mathbf{g}_{\mathtt{D}}^{\text{Full}}$ is zero on the subcube $\{0,1\}^{m^{\text{PCPP}}}$.

Lemma 11.3.2 (Polynomial basis of zero functions, Proposition 10.21 of [JNV+22a])**.** *Let* $m, p \in \mathbb{N}$ *and let* $\mathbf{f} \in IdPoly(p, m, d)$*. Suppose* $\mathbf{f}(s) = 0$ *for all* $s \in \{0,1\}^m \subseteq \mathbb{F}_{2^p}^m$*. Then there exist* m *polynomials* $\{\mathbf{c}_i\}_{i\in[m]}$*, each* $\mathbf{c}_i \in IdPoly(p, m, d)$*, and for all* $(x_0, \cdots, x_m) \in \mathbb{F}_{2^p}^m$

$$\mathbf{f}(x_0, \cdots, x_m) = \sum_{i\in[m]} \mathbf{c}_i(x_0, \cdots, x_m)\cdot\mathbf{zero}(x_i)$$

where $\mathbf{zero} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$ *is the polynomial* $x\cdot(1-x)$*.*

Hence, instead of checking $\mathbf{g}_{\mathtt{D}}^{\text{Full}}$ directly, the verifier can ask the provers to compute each $\mathbf{c}_i$ guaranteed by the above theorem, and check their existence again via the simultaneous quantum low-individual degree test.

We now give a summary for the verification procedure for a verifier given a decider $\mathtt{D}$ and the integer n assuming the provers are honest. We first define $\mathbf{g}_a, \mathbf{g}_b, \mathbf{g}_{w_0}, \mathbf{g}_{w_1}, \mathbf{g}_{w_2} : \mathbb{F}_{2^p}^{m^{\text{PCPP}}} \to \mathbb{F}_{2^p}$ as follows: For every $s \in \mathbb{F}_{2^p}^{m^{\text{PCPP}}}$, partition $s = (s_0, \cdots, s_4, b_0, \cdots, b_4, z)$, where for $i \in [5]$, $s_i \in \mathbb{F}_{2^p}^{m^{\text{ans}}}$, $b_i \in \mathbb{F}_{2^p}$ and $z \in \mathbb{F}_{2^p}^s$.

$$\begin{aligned}
&\mathbf{g}_a(s_0, \cdots, s_4, b, w) = \overline{\mathbf{g}}_a(s_0), \qquad \mathbf{g}_b(s_0, \cdots, s_4, b, w) = \overline{\mathbf{g}}_b(s_1), \qquad \mathbf{g}_{w_0}(s_0, \cdots, s_4, b, w) = \overline{\mathbf{g}}_{w_0}(s_2)\\
&\mathbf{g}_{w_1}(s_0, \cdots, s_4, b, w) = \overline{\mathbf{g}}_{w_1}(s_3), \qquad \mathbf{g}_{w_2}(s_0, \cdots, s_4, b, w) = \overline{\mathbf{g}}_{w_2}(s_4), \qquad (11.7)
\end{aligned}$$

Note for $v \in \{a, b, w_0, w_1, w_2\}$, $\mathbf{g}_v$ and $\overline{\mathbf{g}}_v$ are almost identical, except that $\mathbf{g}_v$ is a low-degree polynomial over m^{ans} and $\overline{\mathbf{g}}_v$ is a low-degree polynomial over m^{PCPP}. We can also rewrite (11.6) without the need to partition the input s, as follows:

$$\mathbf{g}_{\mathtt{D}}^{\text{Full}}(s) = \mathbf{g}_{\mathtt{D}}(s) \cdot (\mathbf{g}_a(s) - b_0) \cdot (\mathbf{g}_b(s) - b_1) \cdot (\mathbf{g}_{w_0}(s) - b_2) \cdot (\mathbf{g}_{w_1}(s) - b_3)(\mathbf{g}_{w_2}(s) - b_4). \tag{11.8}$$

The answer reduction is a combination between the oracularization transformation, however, instead of computing the decision procedure given in Figure 11.1, the encoding from Theorem 11.3.1 are used instead to verify that the provers indeed generate the correct answers for the given question pair. To be more precise, assuming that both provers are honest, the verifier performs the following on the answer reduction protocol with the provers.

1. The verifier first samples the question pair (x, y) and (n_A, n_B) according to the sampling procedure given in Theorem 11.3.1, and send the question pair normally to the two provers.
2. Upon receiving the question label and the question pair, the prover computes the following:
   - If the question label is (Prover, A), the prover generate the answer a for the question x for $\mathscr{G}$, and then generate the polynomial $\overline{\mathbf{g}}_a$ by using the encoding map enc_Γ Theorem 11.3.1.
   - Similarly, if the question label is (Prover, B), the prover generates the answer b for the question y for $\mathscr{G}$, and then generate the polynomial $\overline{\mathbf{g}}_b$.
   - If the question label is (Oracularization), the prover generates the answer pair (a, b) for the question pair (x, y) for $\mathscr{G}$ (from which a only depends on x and b only depends on y). Then, the prover computes the following
     - The polynomial $\overline{\mathbf{g}}_a, \overline{\mathbf{g}}_b$ similarly as above.
     - $\mathbf{g}_{\mathtt{D}}(x)$ base on $\mathtt{D}$, and the question pair (x, y) given by the verifier.
     - The m^{ans}-variate polynomial $\overline{\mathbf{g}}_{w_0}, \overline{\mathbf{g}}_{w_1}, \overline{\mathbf{g}}_{w_2}$ and the m^{PCPP}-variate polynomial $\mathbf{g}_{\mathtt{D}}^{\text{Full}}(x)$ guarantee by Theorem 11.3.1.
3. Then, the verifier computes $(m^{\text{ans}}, m^{\text{PCPP}}, g, p) = \mathtt{PCPParameter}_\alpha(n)$. The verifier also computes the description of $\mathbf{g}_{\mathtt{D}}$, by running $\mathtt{ComputePCPP}_\alpha(\mathtt{D}, n, x, y)$.
4. If the question pair is (Prover, A) – (Oracularization) or (Prover, A) – (Prover, A), the verifier uses the (p, m^{ans}, p)-quantum low-individual degree test to verify that the prover shares the same low degree polynomial $\overline{\mathbf{g}}_a$.
5. If the question pair is (Prover, B) – (Oracularization) or (Prover, B) – (Prover, B), the verifier uses the (p, m^{ans}, p)-quantum low-individual degree test to verify that the prover shares the same low degree polynomial $\overline{\mathbf{g}}_b$.
6. If the question pair for both provers is (Oracularization), the verifier performs the following with some constant probability.
   (a) (Low-individual degree test on assignments) The verifier arbitrarily picks $w \in \{w_0, w_1, w_2\}$, and uses the (p, m^{ans}, p)-quantum low-individual degree test to verify that the prover shares the same low degree polynomial $\overline{\mathbf{g}}_w$.
   (b) (Simultaneous low-individual degree test) The verifier performs the $(p, m^{\text{ans}}, p, 6 + m^{\text{PCPP}})$-simultaneous low-individual degree test on the polynomials
   $$\mathbf{g}_a, \mathbf{g}_b, \mathbf{g}_{w_0}, \mathbf{g}_{w_1}, \mathbf{g}_{w_2}, \mathbf{g}_{\mathtt{D}}^{\text{Full}}, \mathbf{c}_0, \cdots, \mathbf{c}_{m^{PCPP}-1}$$
   where $c_0 \cdots c_{m^{PCPP}-1}$ are the polynomials guaranteed by Lemma 11.3.2 when applied to $\mathbf{g}_{\mathtt{D}}^{\text{Full}}(x)$ to verify that the prover actually shares these polynomials.

(c) (Evaluation test) The verifier samples $s \in \mathbb{F}_{2^p}^{m^{PCPP}}$, and ask both provers to compute

$$(u_0, \cdots, u_4) = (\mathbf{g}_a(s), \mathbf{g}_b(s), \mathbf{g}_{w_0}(s), \mathbf{g}_{w_1}(s), \mathbf{g}_{w_2}(s)), \qquad \gamma = \mathbf{g}_{\mathtt{D}}^{\text{Full}}(s),$$
$$(\beta_0, \cdots, \beta_{m^{PCPP}-1}) = (\mathbf{c}_0(s), \cdots, \mathbf{c}_{m^{PCPP}-1}(s)).$$

The verifier rejects under the following conditions.

- (Consistency check) If the two provers output different values.
- (Formula check) Parse $s = (s_0, \cdots, s_4, b_0, \cdots, b_4, z)$, where for $i \in [5]$, $s_i \in \{0,1\}^{m^{\text{ans}}}$, $b_i \in \{0,1\}$ and $z \in \{0,1\}^s$. The verifier rejects if

$$\gamma \neq \mathbf{g}_{\mathtt{D}}(s) \cdot (u_1 - b_0) \cdots (u_4 - b_4).$$

- (Zero on subcube test check) Parse $s = (s_0, \cdots, s_{m^{PCPP}-1})$, where each $s_i \in \mathbb{F}_{2^p}$ for $i \in [m^{PCPP}]$. The verifier rejects if

$$\gamma \neq \sum_{i \in [m^{PCPP}]} \beta_i \cdot \mathbf{zero}(s_i).$$

where **zero** is the polynomial defined in Lemma 11.3.2.

Since the input/output for the (evaluation test) on the prover side is the same as a "point" question on the (simultaneous low-individual degree test), the evaluation test is attached as part of the consistency test when running the simultaneous low-individual degree test in the answer reduction protocol below. As mentioned in the beginning of the section, the "oracularization" question pair can also be combined with the "simultaneous low-individual degree test" question to make the above procedure a one round interaction.

To analyze the "soundness" condition of the protocol, we have to ensure that if there is no valid answer pair (a,b) with the appropriate length such that $\mathtt{D}(n,x,y,a,b) = 1$, then the provers cannot generate valid polynomials $\mathbf{g}_v$ for $v \in \{a, b, w_0, w_1, w_2\}$ which can be used to trick the verifier in the above procedure. In order to state this more precisely, we need to first define the notion of a low-degree PCPP proof below.

Definition 11.3.3 (Low-degree PCPP proof, definition 10.23 of [JNV+22a])**.** *Given $m^{ans}, g, p \in \mathbb{N}$, let m^{PCPP} be as of the requirement given in Theorem 11.3.1. A low-degree PCPP proof is a tuple $\Pi_{m^{ans},g,p}$ of evaluation tables over polynomials*

$$(\overline{\boldsymbol{g}}_a, \overline{\boldsymbol{g}}_b, \overline{\boldsymbol{g}}_{w_0}, \overline{\boldsymbol{g}}_{w_1}, \overline{\boldsymbol{g}}_{w_2}, \boldsymbol{g}_{\mathtt{D}}^{Full}, \boldsymbol{c}_0, \cdots, \boldsymbol{c}_{m^{PCPP}-1}), \tag{11.9}$$

where $\overline{\boldsymbol{g}}_a, \overline{\boldsymbol{g}}_b, \overline{\boldsymbol{g}}_{w_0}, \overline{\boldsymbol{g}}_{w_1}, \overline{\boldsymbol{g}}_{w_2} \in IdPoly(p, m^{ans}, p)$, and $\boldsymbol{g}_{\mathtt{D}}^{Full}, \boldsymbol{c}_0, \cdots, \boldsymbol{c}_{m^{PCPP}-1} \in IdPoly(p, m^{PCPP}, p)$. The evaluation of $\Pi_{m^{ans},g,p}$ at $s \in \{\mathbb{F}_{2^p}^{m^{PCPP}}\}$ is given by

$$\boldsymbol{eval}_s(\Pi_{m^{ans},g,p}) = (\overline{\boldsymbol{g}}_a(s_0), \overline{\boldsymbol{g}}_b(s_1), \overline{\boldsymbol{g}}_{w_0}(s_2), \overline{\boldsymbol{g}}_{w_1}(s_3), \overline{\boldsymbol{g}}_{w_2}(s_4), \boldsymbol{g}_{\mathtt{D}}^{Full}(s), \boldsymbol{c}_0(s), \cdots, \boldsymbol{c}_{m^{PCPP}-1}(s)).$$

where s is parsed as $s = (s_0, \cdots, s_4, b_0, \cdots, b_4, z)$.

Intuitively, a low-degree PCPP proof is a classical proof in which the provers can generate a proof for the above verification procedure. We now state the (classical) soundness result for the PCPP procedure given in Theorem 11.3.1.

Theorem 11.3.4 (Classical soundness of the time bounded decider PCPP, theorem 10.25 of [JNV+22a])**.** *Let $n, \alpha \in \mathbb{N}$, let $(\mathtt{ComputePCPP}_\alpha, \mathtt{PCPParameter}_\alpha)$ be the two Turing machines given in Theorem 11.3.1, and let*

- $\mathtt{D}$ *be a decider for a CL verifier $\mathscr{V}$,*

- x, y *be two strings of length at most* $\log^{\alpha}(n)$,
- $(m^{ans}, m^{PCPP}, g, p)$ *be the outputs of* $\texttt{PCPPParameter}_{\alpha}(n)$,
- $\mathbf{g}_{\mathsf{D}}$ *be the outputs from* $\texttt{ComputePCPP}_{\alpha}(\mathsf{D}, n, x, y)$.

Then the following holds:

- *(Completeness): If there exist two strings* $a, b \in \{0,1\}^{n^{\alpha}}$ *such that* $D(n,x,y,a,b) = 1$ *with* $\mathsf{TIME}_{\mathsf{D}}(n,x,y,a,b) \leq n^{\alpha}$, *then there exists a low-degree PCPP proof* $\Pi_{m^{ans},g,p}$ *such that for all* $s \in \{\mathbb{F}_{2^p}^{m^{PCPP}}\}$

$$\texttt{ValidatePCPP}(\mathbf{g}_{\mathsf{D}}, m^{ans}, g, p, s, \boldsymbol{eval}_s(\Pi_{m^{ans},g,p})) = 1$$

  where the function `ValidatePCPP` *is as specified in Pseudocode 12.*

- *(Soundness): If there exists a low-degree PCPP proof* $\Pi_{m^{ans},g,p}$ *such that*

$$\Pr_{s \sim \{\mathbb{F}_{2^p}^{m^{PCPP}}\}}[\texttt{ValidatePCPP}(\mathbf{g}_{\mathsf{D}}, m^{ans}, g, p, s, \boldsymbol{eval}_s(\Pi_{m^{ans},g,p})) = 1] > \frac{1}{2}, \tag{11.10}$$

  then there exist two strings $a, b \in \{0,1\}^{n^{\alpha}}$ *such that* $\overline{\mathbf{g}_a} = enc_{\Gamma}(a)$ *and* $\overline{\mathbf{g}_b} = enc_{\Gamma}(b)$, *where* $\overline{\mathbf{g}_a}, \overline{\mathbf{g}_b}$ *are two low-individual degree polynomials used to define* $\Pi_{m^{ans},g,p}$ *and* $D(n,x,y,a,b) = 1$ *with*

$$\mathsf{TIME}_{\mathsf{D}}(n,x,y,a,b) \leq n^{\alpha}.$$

1 **Input**: Polynomial ($\mathbf{g}_{\mathsf{D}}$), parameter m^{ans}, g, p, PCPP view s, Ξ
2 Compute $m^{\text{PCPP}} = 5 \cdot m^{\text{ans}} + 5 + g$.
3 Parse $\Xi = (u_0, \cdots, u_n, \gamma, \beta_0, \cdots, \beta_{m^{\text{PCPP}}-1})$ where each variable is in $\{0,1\}^{m^{\text{PCPP}}}$, return 0 if this cannot be done.
4 Parse $s = (s_0, \cdots, s_4, b_0, \cdots, b_4, z)$, where for $i \in [5]$, $s_i \in \{0,1\}^{m^{\text{ans}}}$, $b_i \in \{0,1\}$ and $z \in \{0,1\}^s$, return 0 if this cannot be done.
5 If $\gamma \neq \mathbf{g}_{\mathsf{D}}(s) \cdot (u_1 - b_0) \cdots (u_4 - b_4)$, return 0.
6 If $\gamma \neq \sum_{i \in [m^{PCPP}]} \beta_i \cdot \mathbf{zero}(s)$, return 0
7 Return 1 if all the clause above fails.

Pseudocode 12: $\texttt{ValidatePCPP}(\mathbf{g}_{\mathsf{D}}, m^{\text{ans}}, g, p, s, \Xi)$, the validation algorithm for the PCPP procedure.

In the above theorem, the completeness clause follows directly from Theorem 11.3.1. However, the soundness clause is stated differently from the quantum soundness used in this thesis because the definition of the PCP differs from that of the classical MIP (see [ALM+98]). The soundness theorem essentially states the following: Imagine that the verifier can give a point s to one of the provers, Alice, and ask her to evaluate s on all the polynomials given in (11.9) and output the answer. Alice has to evaluate the polynomials honestly, but she does not necessarily have to generate the given polynomials according to the procedure given by the PCPP. She can instead fix *any* low-individual degree polynomials and evaluate the point s on them. What the above theorem essentially states is that, unless $\mathsf{TIME}_{\mathsf{D}}$ can be satisfied (by some a, b), Alice cannot construct **any** sets of polynomials that can cause the verifier to accept using the procedure given in Pseudocode 12 with high probability.

11.4 The answer reduction transformation

Using the PCPP procedure given in the last subsection, we give a proof of Proposition 7.4.3 below. For convenience, we restate the proposition below.

Orac Q	SLDT Q	Question content	Answer format
(Prover, A)	(Point)	$x \in \mathscr{X}$, $s^A \in \mathbb{F}_{2^p}^{m^{\text{ans}}}$	$u_A = \mathbb{F}_q$
	(Dline)	$x \in \mathscr{X}$, $(j, s_D^A) \in [m^{\text{ans}}] \times \mathbb{F}_{2^p}^{m^{\text{ans}}}$	$\mathbf{f}_A^{\text{D}} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$
	(Aline)	$x \in \mathscr{X}$, $(j, v, s_A^A) \in [m^{\text{ans}}] \times \mathbb{F}_{2^p}^{2m^{\text{ans}}}$	$\mathbf{f}_A^{\text{A}} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$
(Prover, B)	(Point)	$y \in \mathscr{X}$, $s^B \in \mathbb{F}_{2^p}^{m^{\text{ans}}}$	$u_B = \mathbb{F}_q$
	(Dline)	$y \in \mathscr{X}$, $(j, s_D^B) \in [m^{\text{ans}}] \times \mathbb{F}_{2^p}^{m^{\text{ans}}}$	$\mathbf{f}_B^{\text{D}} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$
	(Aline)	$y \in \mathscr{X}$, $(j, v, s_A^B) \in [m^{\text{ans}}] \times \mathbb{F}_{2^p}^{2m^{\text{ans}}}$	$\mathbf{f}_B^{\text{D}} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$
(Ora_o) $o \in [3]$	(Point)	$(x, y) \in \mathscr{X}^2$, $s^{w_o} \in \mathbb{F}_{2^p}^{m^{\text{ans}}}$	$u_{w_o} \in \mathbb{F}_{2^p}$
	(Dline)	$(x, y) \in \mathscr{X}^2$, $(j, s_{\text{D}}^{w_o}) \in [m^{\text{ans}}] \times \mathbb{F}_{2^p}^{m}$	$\mathbf{f}_{w_o}^{\text{D}} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$
	(Aline)	$(x, y) \in \mathscr{X}^2$, $(j, v, s_{\text{A}}^{w_o}) \in [m^{\text{ans}}] \times \mathbb{F}_{2^p}^{2m}$	$\mathbf{f}_{w_o}^{\text{D}} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$
(Ora)	(Point)	$(x, y) \in \mathscr{X}^2$, $s \in \mathbb{F}_{2^p}^{m}$	$(u_0, \cdots, u_4, \gamma, \beta_0 \cdots, \beta_{m-1}) \in \mathbb{F}_{2^p}$
	(Dline)	$(x, y) \in \mathscr{X}^2$, $(j, s^{\text{D}}) \in [m] \times \mathbb{F}_{2^p}^{m}$	$(\mathbf{f}_{U_0}^{\text{D}}, \cdots, \mathbf{f}_{U_4}^{\text{D}}, \mathbf{f}_{\Gamma}^{\text{D}}, \mathbf{f}_{B_0}^{\text{D}} \cdots, \mathbf{f}_{B_{m-1}}^{\text{D}})$ $\mathbf{f}_v^{\text{D}} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$
	(Aline)	$(x, y) \in \mathscr{X}^2$, $(j, v, s^{\text{A}}) \in [m] \times \mathbb{F}_{2^p}^{2m}$	$(\mathbf{f}_{U_0}^{\text{A}}, \cdots, \mathbf{f}_{U_4}^{\text{A}}, \mathbf{f}_{\Gamma}^{\text{A}}, \mathbf{f}_{B_0}^{\text{A}} \cdots, \mathbf{f}_{B_{m-1}}^{\text{A}})$ $\mathbf{f}_v^{\text{D}} : \mathbb{F}_{2^p} \to \mathbb{F}_{2^p}$

Figure: Q and A format for the answer reduction protocol $\texttt{AnswerReduction}_\alpha$ applied to a CL verifer $(\texttt{Q}, \texttt{D})$.

Sampling procedure

If at any point during the sampling procedure given below, the following happens:

- There is an error in running the subroutine $\texttt{Q}$ or $\texttt{PCPPameter}_\alpha(n)$.
- Or if the runtime exceed $\log^{\gamma^{\text{AR}}}(n)$ (for some γ^{AR} picked later in the proof of Proposition 7.4.3).

Then halt the sampling procedure and instead sample a question pair from $\mathscr{G}_{\text{reject}}$, the rejecting game from Definition 6.2.3.

1. Given the input n, sample a question pair $(x, y) \sim \mu_n$ by using $\texttt{Q}(n, \cdot)$. Then compute the parameter $(m^{\text{ans}}, m^{\text{PCPP}}, g, p) = \texttt{PCPPameter}_\alpha(n)$ and set $m = m^{\text{PCPP}}$.
2. Uniformly $(n_{O,0}), n_{O,1} \in \{(\text{Prover, A}), (\text{Prover, B}), (\text{Ora})\}^2$, if (Ora) is sampled for $n_{O,i}$ for $i \in \{0, 1\}$, uniformly sample another element $\{(\text{Ora})_0, (\text{Ora})_1, (\text{Ora})_2, (\text{Ora})\}$ and set $n_{O,i}$ to be that element. Sample $n_{L,0}, n_{L,1} \sim \{(\text{Point}, \text{DLine}, \text{ALine})\}$ and sample $s \sim \{0, 1\}^m$.
3. Parse $s = (s_0, \cdots, s_4, b_0, \cdots, b_4, z)$, where for $i \in [5]$, $s_i \in \{0, 1\}^{m^{\text{ans}}}$, $b_i \in \{0, 1\}$ and $z \in \{0, 1\}^s$.
4. For $i \in \{0, 1\}$ corresponds to the two provers
   (a) If $n_{O,i} = (\text{Prover, A})$, sample a question for the (p, m^{ans}, p)-quantum low-individual degree test according to the label of $n_{L,i}$, where the point question uses $s_A = s_0$. Send $((\text{Prover, A}), x)$, as well as the question label for the quantum low-individual degree test to prover i.
   (b) If $n_{O,i} = (\text{Prover, B})$, sample a question for the (p, m^{ans}, p)-quantum low-individual degree test according to the label of $n_{L,i}$, where the point question uses $s_B = s_1$. Send $((\text{Prover, B}), y)$, as well as the question label for the quantum low-individual degree test to prover i.
   (c) If $n_{O,i} = (\text{Ora})_o$ for $o \in [3]$, sample a question for the (p, m^{ans}, p)-quantum low-individual degree test according to the label of $n_{L,i}$, where the point question uses $s_{w_o} = s_{o+2}$. Send $((\text{Ora}), (x, y))$, as well as the question label for the quantum low-individual degree test to prover i.
   (d) Otherwise, sample a question for the $(p, m, p, 6 + m)$-simultaneous quantum low-individual degree test according to the label of $n_{L,i}$, where the point question uses s. Send $((\text{Ora}), (x, y))$, as well as the question label for the simultaneous quantum low-individual degree test to prover i.

Figure 11.2: The description for $\texttt{Q}^{\text{AR}}$ for the answer reduction protocol $\texttt{AnswerReduction}_\alpha$ applied to a CL verifer $(\texttt{Q}, \texttt{D})$.

Verification procedure

If at any point during the sampling procedure given below, the following happens:

- There is an error in running the subroutine $\mathtt{D}$, $\mathtt{PCPPParameter}_\alpha$ or $\mathtt{ComputePCPP}_\alpha$.
- Or if the runtime exceed $\log^{\gamma^{\mathrm{AR}}}(n)$ (for some γ^{AR} picked later in the proof of Proposition 7.4.3).

Halt the decision procedure and immediately outputs 0.

Preprocessing steps.

1. Compute the description of the polynomial $\mathbf{g}_{\mathtt{D}} = \mathtt{ComputePCPP}_\alpha(\mathtt{D}, n, x, y)$ from Theorem 11.3.1.
2. Compute the parameter $(m^{\mathrm{ans}}, m^{\mathrm{PCPP}}, g, p) = \mathtt{PCPPParameter}_\alpha(n)$ and set $m = m^{\mathrm{PCPP}}$.

The verifier then perform the following series of check in sequences:

Low-individual degree check.

1. If $n_{O,0} = n_{O,1} =$ (Prover, P) for P $\in$ {A, B} or $n_{O,0} = n_{O,1} = (\mathrm{Ora}_o)$ for $o \in [3]$, return 0 if they loses on the (p, m^{ans}, p)-quantum low-individual degree test.
2. If $n_{O,0} = n_{O,1} =$ (Ora), return 0 if they loses on the $(p, m, p, 6+m)$-simultaneous quantum low-individual degree test instances.

Prover consistency check. For $i \in \{0,1\}$

1. If $(n_{O,i}, n_{L,i}) = (n_{O,1-i}, n_{L,1-i})$, return 0 if the answer given by both provers are inconsistent with each other.
2. If $(n_{O,i}, n_{L,i}) =$ ((Prover, A), (Point)) and $(n_{O,1-i}, n_{L,1-i}) =$ ((Ora), (Point)), return 0 if $u_A \neq u_0$.
3. If $(n_{O,i}, n_{L,i}) =$ ((Prover, B), (Point)) and $(n_{O,1-i}, n_{L,1-i}) =$ ((Ora), (Point)), return 0 if $u_B \neq u_1$.
4. If $(n_{O,i}, n_{L,i}) = (\mathrm{Ora}_o, \text{(Point)})$ for $o \in [3]$ and $(n_{O,1-i}, n_{L,1-i}) =$ ((Ora), (Point)), return 0 if $u_{w_o} \neq u_{o+2}$.

PCPP proof check. For $i \in \{0,1\}$

1. If $(n_{O,i}, n_{L,i}) =$ ((Ora), (Point)), write $s = (s_0, \cdots, s_4, b_0, \cdots, b_4, z)$. Return 0 if either
   (a) $\gamma \neq \mathbf{g}_{\mathtt{D}}(s) \cdot (u_1 - b_0) \cdots (u_4 - b_4)$.
   (b) $\gamma \neq \sum_{i \in [m^{PCPP}]} \beta_i \cdot \mathbf{zero}(s_i)$

Return 1 if none of the test above fails.

Figure 11.3: The description for $\mathtt{D}^{\mathrm{AR}}$ for the answer reduction protocol $\mathtt{AnswerReduction}_\alpha$ applied to a CL verifer $(\mathtt{Q}, \mathtt{D})$.

Proposition 11.4.1 (Answer Reduction)**.** *For all constants $(\alpha, k) \in \mathbb{N}$ there exists a polynomial time algorithm $\texttt{AnswerReduction}_{\alpha,k}$ that takes, as input, a pair of Turing machines $(\mathtt{Q},\mathtt{D})$ and outputs a tuple of Turing machine machines $(\mathtt{Q}^{AR},\mathtt{D}^{AR})$ such that the following holds:*

There exist a polynomial For models $t \in \{*, co\}$, $\texttt{AnswerReduction}_{\alpha}$ outputs a pair of Turing machines $(\mathtt{Q}^{AR},\mathtt{D}^{AR})$, which defines a synchronous CL-verifier $\mathscr{V}^{AR}$ for an infinite sequence of game $\mathscr{G}^{AR} = \{\mathscr{G}_n^{AR}\}_{n\in\mathbb{N}}$ with the following properties: There exists an integer $\gamma^{AR} = O(poly(\alpha))$ such that

1. *(Runtime): $\texttt{AnswerReduction}_{\alpha}$ have runtime*

$$\mathsf{TIME}_{\texttt{AnswerReduction}_{\alpha}}(\mathrm{poly}(\alpha), |\mathtt{Q}|, |\mathtt{D}|).$$

2. *(Dependency for $\mathtt{Q}^{AR}$) The Turing machine $\mathtt{Q}^{AR}$ only depends on the input $\mathtt{Q}$ and α.*

Furthermore, if the input $\mathscr{V} = (\mathtt{Q},\mathtt{D})$ is a $\max\{4, k+2\}$ level synchronous CL verifier for the infinite sequence of synchronous game $\mathscr{G} = \{\mathscr{G}_n\}_{n\in\mathbb{N}}$ for some constant $k \in \mathbb{N}$, and there exists some constant $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, we have

- $|\mathtt{D}^{QR}| = O(\mathrm{poly}(\alpha, k))$.
- $\mathtt{Q}(n, Parameter) \leq \log^{\alpha}(n)$,
- $\mathsf{TIME}_{\mathtt{Q}}(n) \leq \log^{\alpha}(n)$,
- $\mathsf{TIME}_{\mathtt{D}} \leq n^{\alpha}$.

Then there exists some integer $n_0^{AR} = \mathrm{poly}(\gamma^{AR}, n_0)$ with $n_0 \leq n_0^{AR}$ such that for all $n \geq n_0^{AR}$

1. *(Complexity bounds for the output):*
   - $\mathsf{TIME}_{\mathtt{Q}^{AR}}(n) \leq \max\left\{\log^{\gamma^{AR}}(n), C^{trivial}\right\}$,
   - $\mathsf{TIME}_{\mathtt{D}^{AR}}(n) \leq \max\left\{\log^{\gamma^{AR}}(n), C^{trivial}\right\}$,

   for some universal constant $C^{trivial}$.
2. *(Level for the CL sampler) The Turing machine $\mathtt{Q}^{AR}$ is a level $\max\{k+2, 5\}$ CL sampler.*
3. *(Completeness) If there exists a perfect oracularizable strategy for $\mathscr{G}_n$ in model t, then there exists a perfect oracularizable strategy for $\mathscr{G}_n^{AR}$ in model t.*
4. *(Soundness) There exist a universal function $\mathbf{s}_{\alpha}^{AR}$ which depends on n and ε with $O(\mathbf{s}_{\alpha}^{AR}) = O(\mathrm{polylog}(n), \mathrm{poly}(\varepsilon))$ such that, for any polynomial $\varepsilon : \mathbb{N} \to [0,1]$*

$$\omega^t(\mathscr{G}_n) \leq 1 - \varepsilon(n) \Longrightarrow \omega^t(\mathscr{G}_n^{QR}) \leq 1 - \mathbf{s}_{\alpha}^{AR}(\varepsilon(n), n)$$

Proof. Fix the constants $\alpha, k \in \mathbb{N}$. We define the algorithm $\mathtt{AnswerReduction}_{\alpha,k}$ as follows: Given a pair of Turing machine $(\mathtt{Q},\mathtt{D})$, we specify the sampling procedure $\mathtt{Q}^{\mathrm{AR}}$ in Figure 11.2 and the decision procedure $\mathtt{D}^{\mathrm{AR}}$ by Figure 11.3. We denote $\mathscr{G}_n$ to be the n-th game generated by the original game sequence (from the input $(\mathtt{Q},\mathtt{D})$) and $\mathscr{G}_n^{\mathrm{AR}} = (\mathscr{X}_n^{\mathrm{AR}}, \mathscr{A}_n^{\mathrm{AR}}, \mu_n^{\mathrm{AR}}, D_n^{\mathrm{AR}})$ be the n-th game from the answer reduction transformation.

The "Runtime" clause of Proposition 7.4.3 follows from the description of Figure 11.2 and Figure 11.3, where computing the description of $\langle\mathtt{Q}^{\mathrm{AR}}\rangle$ and $\langle\mathtt{D}^{\mathrm{AR}}\rangle$ does not involve computing specific instances of $(\mathtt{Q},\mathtt{D})$ (and hence have no dependency on n). The "Dependency for $\mathtt{Q}^{\mathrm{AR}}$" follows from the description of Figure 11.2 (i.e. it only depends on $\mathtt{Q}$ and the constant α).

Now, assume the input for $\texttt{AnswerReduction}_{\alpha,k}$ is a game sequence $\mathscr{V} = (\texttt{Q}, \texttt{D})$ with the property given in the "furthermore" part of Proposition 7.4.3. To show the "complexity bound for the output" part, we analyze the complexity of $\texttt{Q}^{\mathrm{AR}}$ and $\texttt{D}^{\mathrm{AR}}$ without the "runtime exceeding $\log^{\gamma^{\mathrm{AR}}}(n)$" clause from Figure 11.2 and Figure 11.3. We start with $\texttt{Q}^{\mathrm{AR}}$ where, by analyzing each step of the description given in Figure 11.2, we incur the following:

1. Sampling the question pair (x, y) from the Turing machine $\texttt{Q}$ takes time $\log^{\alpha}(n)$ by definition. By Theorem 11.3.1 computing $\texttt{PCPParameter}_{\alpha}(n)$ takes time $O(\mathrm{poly}(\alpha, \log(n)))$. This implies that step 1 takes step $O(\mathrm{poly}(\alpha, \log^{\alpha}(n)))$.

2. Sampling the "oracularization" label takes constant time. By the parameter choice guaranteed by m^{ans} and $\mathbf{g}$, $m = O(\mathrm{poly}(\alpha, \log(n)))$ which means sampling s takes time $O(\mathrm{poly}(\alpha, \log(n)))$, giving step 2 of a runtime of $O(\mathrm{poly}(\alpha, \log^{\alpha}(n)))$ steps.

3. Step 3 takes time $O(\mathrm{poly}(\alpha, \log^{\alpha}(n)))$ time due to the size of m.

4. Sampling the question label for the quantum low-individual degree test/simultaneous quantum low-individual degree test takes time $O(m) = O(\mathrm{poly}(\alpha, \log^{\alpha}(n)))$ time.

Hence, since α is chosen to be a constant in the beginning, the runtime for $\texttt{Q}^{\mathrm{AR}}$, assuming it does not stop by the termination clause, takes time $O(\mathrm{polylog}(n))$ time. Now we analyze the runtime for $\texttt{D}^{\mathrm{AR}}$. By studying each line of the description given in Figure 11.3, we incur the following:

- **Preprocessing steps.** Computing $\texttt{ComputePCPP}_{\alpha}(\texttt{D}, n, x, y)$ takes time $O(\mathrm{poly}(\alpha, \log(n)))$ and computing $\texttt{PCPParameter}_{\alpha}(n)$ takes time $O(\mathrm{poly}(\alpha, \log(n), |\langle \texttt{D} \rangle|)) = O(\mathrm{poly}(\alpha, \log(n), k))$.

- **Low-individual degree check.** Verifying the quantum low-individual degree test/simultaneous quantum low-individual degree test in this instance takes time $O(\mathrm{poly}(\alpha, \log(n)))$ by the choices of parameters and Lemma 7.1.1.

- **Prover consistency check.** Comparing equality of outputs takes time $O(\mathrm{poly}(\alpha, \log(n)))$ again by the choices of parameters.

- **PCPP proof check.** Evaluating the low-individual degree polynomial and comparing the corresponding output takes time $O(\mathrm{poly}(\alpha, \log(n)))$ again by the choices of parameters.

Again, since α and k are chosen to be constants, the runtime for $\texttt{D}^{\mathrm{AR}}$, assuming it does not stop by the termination clause, takes time $O(\mathrm{polylog}(n))$. Hence, pick $\gamma^{\mathrm{AR}} \in \mathbb{N}$ to be the minimum constant such that the following holds:

$$\mathsf{TIME}_{\texttt{Q}^{\mathrm{AR}}}(n) \leq O(\log^{\gamma^{\mathrm{AR}}}(n)), \qquad \mathsf{TIME}_{\texttt{D}^{\mathrm{AR}}}(n) \leq O(\log^{\gamma^{\mathrm{AR}}}(n)). \tag{11.11}$$

and let C^{trivial} be the constant such that $\mathscr{G}^{\mathrm{reject}}$ can be both sampled and decided in time C^{trivial}. Pick $n_0^{\mathrm{AR}} \in \mathbb{N}$ be the smallest integer such that

$$\mathsf{TIME}_{\texttt{Q}^{\mathrm{AR}}}(n_0^{\mathrm{AR}}) \leq \log^{\gamma^{\mathrm{AR}}}(n_0^{\mathrm{AR}}), \qquad \mathsf{TIME}_{\texttt{D}^{\mathrm{AR}}}(n_0^{\mathrm{AR}}) \leq O(\log^{\gamma^{\mathrm{AR}}}(n_0^{\mathrm{AR}})), \tag{11.12}$$

and the "complexity bound" clause follows from the definition of the big O notation.

For the "level clause", we see that the input distribution μ_n^{AR} is essentially an instance of the oracularization transformation of $\mathscr{G}_n$ and an instance of the quantum low-individual degree test/simultaneous quantum low-individual degree test depending on the oracularization label. By combining the "sample complexity" clause on both Lemma 11.2.1 and Lemma 10.4.1, and using the series composition of CL functions given by Definition 7.2.3 (in this case, only the first label for oracularization is being used to control which of the low-individual degree/simultaneous

quantum low-individual degree test is being performed), we obtain a $\max\{k+2, 5+1\}$-CL distribution that samples μ_n^{AR}. This shows the 'level clause" of the proposition.

For the "completeness clause", we describe the perfect oracularizable strategy for $\mathcal{G}^{\mathrm{AR}}$ as follows:

1. Upon receiving the questions, the provers first perform the perfect strategy for $\mathcal{G}_n^{\mathrm{Ora}}$, the oracularized version of $\mathcal{G}_n$, guaranteed by Lemma 11.2.1 using the oracularization label and the question pair for the original game as the question label for $\mathcal{G}_n^{\mathrm{Ora}}$.

2. For $P \in \{A, B\}$, if the oracularization label is "(Label P)", the prover computes the following:

   (a) Generate the corresponding low-individual degree polynomial $\mathbf{g}_p$ according to Theorem 11.3.1.

   (b) Perform the quantum low-individual degree test using the SLDT Q label and the low-degree test question content as the question label, and using $\mathbf{g}_p$ as the "shared polynomial" between the two provers. We remark that this step is classical according to the procedure described in Section 10.2.

3. For $o \in [3]$, if the oracularization label is "Ora_o", the prover computes the following:

   (a) Compute the polynomials described in Theorem 11.3.1 using the question labels (x, y), and the answer obtained (a, b).

   (b) Perform the quantum low-individual degree test using the SLDT Q label and the low-degree test question content as the question label, and using $\mathbf{g}_{w_o}$ as the "shared polynomial" between the provers.

4. If the oracularization label is "Ora", the prover computes the following:

   (a) Compute the polynomials described in Theorem 11.3.1 using the question labels (x, y), and the answer obtained (a, b), also computes the polynomials described in Theorem 11.3.1, as well as the "zero polynomials" $\mathbf{c}_i$, $i \in [m]$ generated by Lemma 11.3.2 applied to the polynomial $\mathbf{g}_{\mathrm{D}}^{\mathrm{Full}}$.

   (b) Perform the simultaneous quantum low-individual degree test using the SLDT Q label and the low-degree test question content as the question label, and using all the polynomials computed in the previous step as the "shared polynomials" between the provers.

From a measurement perspective, this is essentially just an instance of $\mathcal{G}_n^{\mathrm{Ora}}$ for the provers. Hence, the completeness clause follows from the completeness clause of Lemma 11.2.1.

The "soundness" clause proceeds similarly to [JNV+22a, Section 10.7], except we use the notation from Table 4.1 to translate the proof from the finite-dimensional setting to the tracially embeddable strategies setting. Since the "soundness" proof is extremely technical, we choose to present it in the next section for clarity. □

11.5 Proof of the "soundness" clause for the AR transformation

Fix $\alpha, n \in \mathbb{N}$, let $\mathcal{G}_n = (\mathcal{X}_n, \mathcal{A}_n, \mu_n, D_n)$ be the nth game of $\mathcal{V}$, and let $\mathcal{G}_n^{\mathrm{AR}} = (\mathcal{X}_n^{\mathrm{AR}}, \mathcal{A}_n^{\mathrm{AR}}, \mu_n^{\mathrm{AR}}, D_n^{\mathrm{AR}})$ and $\mathcal{G}_n^{\mathrm{Ora}} = (\mathcal{X}_n^{\mathrm{Ora}}, \mathcal{A}_n^{\mathrm{Ora}}, \mu_n^{\mathrm{Ora}}, D_n^{\mathrm{Ora}})$ denote the answer reduction transformation given in Section 11.4 and Section 11.2. Given a question label $x \in \mathcal{X}_n^{\mathrm{AR}}$, we write $x = (x^{\mathrm{Ora}}, x^{\mathrm{LDL}}, (x^{\mathrm{game}}, x^{\mathrm{LDC}}))$ where each elements corresponds to the "oracularizable question label", "SLDT question label", question content for the original game $\mathcal{G}_n$, and question content for SLDT from Figure 11.2 respectively. When we fix certain question labels, we might omit that portion of the

question label for simplicity notation. Furthermore, let $(\texttt{PCPParameter}_\alpha, \texttt{ComputePCP}_\alpha)$ be the two Turing machine guaranteed by Theorem 11.3.1, and let $(m^{\text{ans}}, m, g, p) = \texttt{PCPParameter}_\alpha(n)$.

Before we start giving a proof of the "soundness clause", we first give an overview on how the proof goes. To show the "soundness" clause, it is equivalent to show that there exist a polynomial function $\mathbf{t}_\alpha^{\text{AR}}$ with $\mathbf{t}_\alpha^{\text{AR}} = O(\text{polylog}(n), \text{poly}(\varepsilon))$ such that for model $t \in \{*, co\}$

$$\omega^t(\mathcal{G}_n^{\text{AR}}) > 1-\varepsilon \Longrightarrow \omega^t(\mathcal{G}) > 1 - \mathbf{t}_\alpha^{\text{AR}}(\varepsilon, n).$$

Hence, assume that $\omega^t(\mathcal{G}_n^{\text{AR}}) > 1-\varepsilon$, and fix strategy in model t such that $\mathscr{S}$ succeed at $\mathcal{G}_n^{\text{AR}}$ with probability at most ε. We show the soundness clause by proving the following:

1. By using Theorem 10.3.1, we first show that there exist a strategy $\mathscr{S}^{\text{poly}}$ for $\mathcal{G}_n^{\text{AR}}$ which consists of the provers first performing hidden measurements and sampled $6+m$ low-individual degree polynomials, and then using these polynomials to pass all the low-individual degree polynomial/simultaneous low-individual degree polynomial test for $\mathcal{G}_n^{\text{AR}}$.

2. Then we use Theorem 11.3.4 on the polynomial generated by $\mathscr{S}^{\text{poly}}$ to construct a strategy $\mathscr{S}^{\text{Ora}}$ for $\mathcal{G}_n^{\text{Ora}}$ which succeed with probability at least $1 - O(\text{polylog}(n), \text{poly}(\varepsilon))$.

3. Finally, we conclude the proof by applying Lemma 11.2.1 to show that there exists a strategy for $\mathcal{G}_n$ which succeed with probability at least $1-O(\text{polylog}(n), \text{poly}(\varepsilon))$, thus showing the lemma.

For simplicity of notations, we work with synchronous strategies in order to show point 1 and 2. For a question pair $(x, y) = \big((x^{\text{Ora}}, x^{\text{LDL}}, (x^{\text{game}}, x^{\text{LDC}})), (y^{\text{Ora}}, y^{\text{LDL}}, (y^{\text{game}}, y^{\text{LDC}}))\big)$ for $\mathcal{G}_n^{\text{AR}}$, we observe that the synchronous question pair for $\mathcal{G}_n^{\text{AR}}$ corresponds to the case where $x^{\text{Ora}} = y^{\text{Ora}}$ and $x^{\text{LDL}} = y^{\text{LDL}}$ which occur with constant probability. This implies that the game $\mathcal{G}_n^{\text{AR}}$ is $\frac{1}{c_b}$-balanced for some constant $\frac{1}{c_b}$, and hence any strategy $\mathscr{S}$ for $\mathcal{G}_n^{\text{AR}}$ which succeed with probability $1-\varepsilon$ must be $c_b \cdot \varepsilon$-synchronous. By Lemma 5.3.1

$$\omega^t(\mathcal{G}_n^{\text{AR}}) > 1-\varepsilon \Longrightarrow \omega_s^t(\mathcal{G}_n^{\text{AR}}) > 1-\varepsilon - \mathbf{s}^{\text{Rounding}}(c_b\varepsilon) = 1-\delta_1. \tag{11.13}$$

Hence, fix a synchronous strategy $\mathscr{S} = (\mathscr{L}^2(\mathscr{A}, \tau), |\tau\rangle, \{A_a^x\})$ such that $\omega(\mathcal{G}_n^{\text{AR}}, \mathscr{S}) > 1 - c_1 \cdot \delta_1$, where c_1 is a sufficiently small constant chosen later in the proof.

Define the function $\text{eval}_s^n : \text{IdPoly}(p, m^{\text{ans}}, p)^{\times n} \to \mathbb{F}_{2^p}^{\times n}$ as $\text{eval}_s^n(\mathbf{g}_0, \cdots, \mathbf{g}_{n-1}) = (\mathbf{g}_0(s), \cdots, \mathbf{g}_{n-1}(s))$. We wish to first show the following claim:

Claim 11.5.1. *For all $(x^{game}, y^{game}) \in \mathscr{X}^2$, there exist six sets of PVM in $\mathscr{A}'$,*

$$\{G_{\overline{\mathbf{g}}}^{(Prover,\,A),x^{game}}\}_{\overline{\mathbf{g}} \in IdPoly(p,m^{ans},p)}, \quad \{G_{\overline{\mathbf{g}}}^{(Prover,\,B),y^{game}}\}_{\overline{\mathbf{g}} \in IdPoly(p,m^{ans},p)}$$
$$\{G_{\overline{\mathbf{g}}}^{(Ora)_o,(x^{game},y^{game})}\}_{\overline{\mathbf{g}} \in IdPoly(p,m^{ans},p)} \text{ for } o \in \{0,1,2\}$$
$$\{G_{\mathbf{g}_{U_0},\cdots,\mathbf{g}_{U_4},\mathbf{g}_\Gamma,\mathbf{g}_{B_0}\cdots,\mathbf{g}_{B_{m-1}}}^{(Ora),(x^{game},y^{game})}\}_{\mathbf{g}_v \in IdPoly(p,m,p)}$$

such that the following hold: For $s \in \mathbb{F}_{2^p}^{m^{ans}}$ and $n \in \mathbb{N}$,

$$A_u^{(Prover,\,A),(Point),x^{game},s} \simeq_{O(\text{poly}(\varepsilon,\log(n),\alpha))} G_{[\boldsymbol{eval}_s^1|u]}^{(Prover,\,A),x^{game}}, \tag{11.14}$$

$$A_u^{(Prover,\,B),(Point),y^{game},s} \simeq_{O(\text{poly}(\varepsilon,\log(n),\alpha))} G_{[\boldsymbol{eval}_s^1|u]}^{(Prover,\,B),y^{game}}, \tag{11.15}$$

$$A_u^{(Ora)_o,(Point),((x^{game},y^{game}),s)} \simeq_{O(\text{poly}(\varepsilon,\log(n),\alpha))} G_{[\boldsymbol{eval}_s^1|u]}^{(Ora)_o,(x^{game},y^{game})}, \quad o \in \{0,1,2\}, \tag{11.16}$$

$$A_{u_0,\cdots,u_4,\gamma,\beta_0,\cdots,\beta_{m-1}}^{(Ora),(Point),((x^{game},y^{game}),s)} \simeq_{O(\text{poly}(\varepsilon,\log(n),\alpha))} G_{[\boldsymbol{eval}_s^{6+m}|(u_0,\cdots,u_4,\gamma,\beta_0,\cdots,\beta_{m-1})]}^{Ora,(x^{game},y^{game})}, \tag{11.17}$$

where $\simeq$ for the above four equations is defined over the distribution $(x^{game}, y^{game}) \sim \mu_n$ and $s \sim \mathbb{F}_{2^p}^{m^{ans}}$ for the first 3 equation, and $s \sim \mathbb{F}_{2^p}^m$ for the last equation and over the state $|\tau\rangle$.

Proof. $\mathscr{G}^{\mathrm{AR}}$, when the question set is restricted to the case where the oracularization question label is restricted to "(Prover A)", and the question content for $\mathscr{G}_n$ is fixed to some $x^{\mathrm{game}} \in \mathscr{X}$, is precisely an instance of the (p,m,p)-low-individual degree test. Since $\mathscr{S}$ succeed with probability $1-c_1\cdot\delta_1$, and the probability that both oracularization question label in a question pair to be both "(Prover A)" being $\frac{1}{9}$. This implies that, on average over $(\mu_n)_X$, the strategy $\mathscr{S}$, when restricted to the case where $(x^{\mathrm{Ora}}, y^{\mathrm{Ora}})$ being both "(Prover A)", succeed with probability at least $1-9\cdot\delta_1$. Hence, by Theorem 10.3.1

$$A_u^{(\text{Prover, A}),(\text{Point}),x^{\mathrm{game}},s} \simeq_{\eta_{\mathrm{LD}}(p,m,p,9\cdot\delta_1)} G^{(\text{Prover, A}),x^{\mathrm{game}}}_{[\mathbf{eval}_s^1|u]},$$

where $\simeq$ for the above equations is defined over the distribution $(x^{\mathrm{game}}) \sim \mu_n$ and $s \sim \mathbb{F}_{2^p}^{m^{\mathrm{ans}}}$ By the choice of the parameter $p, m^{\mathrm{ans}} = O(\mathrm{poly}(\alpha, \log(n)))$, this immediately shows (11.14). (11.15) and (11.16) follows from a similar argument (except with the oracularization question label replaced with "(Prover B)" and "$(\mathrm{Ora})_o$", $i = \{0,1,2\}$ respectively). (11.17) follows a similar argument with the label "(Ora)" and using Lemma 10.4.1. This completes the proof. □

We wish to modify the output for the PVM associated with the "(Prover, A)", "(Prover, B)" and "$(\mathrm{Ora})_o$", $i \in \{0,1,2\}$ as PVM which outputs $\mathbf{g} \in \mathrm{IdPoly}(p,m,p)$. For $s \in \mathbb{F}_{2^p}^m$, partition $s = (s_0, \cdots, s_4, w)$ where $s_i \in \mathbb{F}_{2^p}^{m^{\mathrm{ans}}}$, $i \in [5]$ and $w \in \mathbb{F}_{2^p}^{5+g}$, we make the following post measurement processing to PVMs given in the last lemma as follows

- Treat the outputs $\mathbf{g}$ from $\{G^{(\text{Prover, A}),x^{\mathrm{game}}}_{\overline{\mathbf{g}}}\}$ as $\mathbf{g} \in \mathrm{IdPoly}(p,m,p)$ as $\mathbf{g}(s) = \overline{\mathbf{g}}(s_0)$.
- Treat the outputs $\mathbf{g}$ from $\{G^{(\text{Prover, B}),y^{\mathrm{game}}}_{\overline{\mathbf{g}}}\}$ as $\mathbf{g} \in \mathrm{IdPoly}(p,m,p)$ as $\mathbf{g}(s) = \overline{\mathbf{g}}(s_1)$.
- For $o \in \{0,1,2\}$, treat the outputs $\mathbf{g}$ from $\{G^{(\mathrm{Ora})_o,(x^{\mathrm{game}},y^{\mathrm{game}})}_{\overline{\mathbf{g}}}\}$ as $\mathbf{g} \in \mathrm{IdPoly}(p,m,p)$ with $\mathbf{g}(s) = \overline{\mathbf{g}}(s_{o+2})$.

To distinguish the two measurement outputs, we write the output polynomial as $\overline{\mathbf{g}}$ if the resulting polynomial output are from $\mathrm{IdPoly}(p, m^{\mathrm{ans}}, p)$ and $\mathbf{g}$ if the output are from $\mathrm{IdPoly}(p,m,p)$. For $i \in [5]$, define the PVM measurement

$$G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}}),U_i}_{\mathbf{g}_{U_i}} = \sum_{\substack{\mathbf{g}_{U_0},\cdots,\mathbf{g}_{U_{i-1}},\mathbf{g}_{U_{i+1}},\cdots,\mathbf{g}_{U_4} \\ \mathbf{g}_\Gamma,\mathbf{g}_{B_0},\cdots,\mathbf{g}_{B_{m-1}} \in \mathrm{IdPoly}(p,m,p)}} G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})}_{\mathbf{g}_{U_0},\cdots,\mathbf{g}_{U_4},\mathbf{g}_\Gamma,\mathbf{g}_{B_0},\cdots,\mathbf{g}_{B_{m-1}}} \tag{11.18}$$

$$G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}}),\mathrm{Full}}_{\mathbf{g}_\Gamma,\mathbf{g}_{B_0}\cdots,\mathbf{g}_{B_{m-1}}} = \sum_{\mathbf{g}_{U_0},\cdots,\mathbf{g}_{U_4} \in \mathrm{IdPoly}(p,m,p)} G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})}_{\mathbf{g}_{U_0},\cdots,\mathbf{g}_{U_4},\mathbf{g}_\Gamma,\mathbf{g}_{B_0},\cdots,\mathbf{g}_{B_{m-1}}}. \tag{11.19}$$

Since $G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})}$ is projective, for all $(x,y) \in \mathscr{X}_n^2$ and outputs $\mathbf{g}_\nu$, $\nu \in \{U_0,\cdots,U_4,\Gamma,B_0,\cdots,B_{m-1}\}$

$$G^{(\mathrm{Ora}),(x,y} = G^{(\mathrm{Ora}),(x,y),U_0}\cdots G^{(\mathrm{Ora}),(x,y),U_4} G^{(\mathrm{Ora}),(x,y),\mathrm{Full}} G^{(\mathrm{Ora}),(x,y),U_4}\cdots G^{(\mathrm{Ora}),(x,y),U_0}. \tag{11.20}$$

Base on the decision procedure for $\mathscr{G}^{\mathrm{AR}}$, we show the following claim

Claim 11.5.2. *On average over $(x,y) \sim \mu$, $s \in \mathbb{F}_{2^p}^m$ and over the state $|\tau\rangle$*

$$G^{(\textit{Prover, A}),x^{\textit{game}}}_{[\boldsymbol{eval}_s^1|u_0]} \simeq_{O(\mathrm{poly}(\varepsilon,\log(n),\alpha))} (G^{(\textit{Ora}),(x^{\textit{game}},y^{\textit{game}}),U_0}_{[\boldsymbol{eval}_s^1|u_0]})^{op} \tag{11.21}$$

$$G^{(\textit{Prover, B}),y^{\textit{game}}}_{[\boldsymbol{eval}_s^1|u_1]} \simeq_{O(\mathrm{poly}(\varepsilon,\log(n),\alpha))} (G^{(\textit{Ora}),(x^{\textit{game}},y^{\textit{game}}),U_1}_{[\boldsymbol{eval}_s^1|u_1]})^{op} \tag{11.22}$$

$$G^{(\textit{Ora})_o,(x^{\textit{game}},y^{\textit{game}})}_{[\boldsymbol{eval}_s^{o+2}|u_{o+2}]} \simeq_{O(\mathrm{poly}(\varepsilon,\log(n),\alpha))} (G^{(\textit{Ora}),(x^{\textit{game}},y^{\textit{game}}),U_{o+2}}_{[\boldsymbol{eval}_s^{o+2}|u_{o+2}]})^{op},\ o \in \{0,1,2\}. \tag{11.23}$$

Proof. We show the proof for (11.21) below, the proof for (11.22) and (11.23) follows a similar proof. Consider the question pair (x,y) for $\mathscr{G}^{\mathrm{AR}}$ where $(x^{\mathrm{Ora}}, y^{\mathrm{Ora}}) = ((\text{Prover A}), (\text{Ora}))$ and

$x^{\mathrm{LDL}} = y^{\mathrm{LDL}} = (\mathrm{Point})$, this occur with constant probability. Since $\mathscr{S}$ succeed with probability at least $1 - c_1 \cdot \delta_1$, by point 2 of the "Prover consistency check" from Figure 11.3,

$$A_{u_0}^{((\mathrm{Prover,\,A}),(\mathrm{Point}),(x^{\mathrm{game}},s))} \simeq_{O(\mathrm{poly}(\varepsilon,\log(n),\alpha))} A_{u_0,\cdots,u_4,\gamma,\beta_0,\cdots,\beta_{m-1}}^{(\mathrm{Ora}),(\mathrm{Point}),((x^{\mathrm{game}},y^{\mathrm{game}}),s)} \tag{11.24}$$

over $(x,y) \sim \mu$ and $s \sim \mathbb{F}_{2^p}^m$ and over the state $|\tau\rangle$. (11.21) then follows from Lemma 3.3.3 point 1 and 2 which translates between $\simeq$ distance to $\approx$ distance, and the triangle inequality for $\approx$ distance applied to (11.14), (11.24), and (11.17). □

For $(x^{\mathrm{game}}, y^{\mathrm{game}}) \in \mathscr{X}_n$, define the POVM measurement $M^{(x^{\mathrm{game}},y^{\mathrm{game}})}_{\mathbf{g}_{U_0},\cdots,\mathbf{g}_{U_4},\mathbf{g}_\Gamma,\mathbf{g}_{B_0}\cdots,\mathbf{g}_{B_{m-1}}}$ with outcomes $\mathbf{g}_\nu \in \mathrm{IdPoly}(p,m,p)$ as

$$\begin{aligned} M^{(x,y)}_{\mathbf{g}_{U_0},\cdots,\mathbf{g}_{U_4},\mathbf{g}_\Gamma,\mathbf{g}_{B_0}\cdots,\mathbf{g}_{B_{m-1}}} =& G^{(\mathrm{A}),x}_{\mathbf{g}_{U_0}} G^{(\mathrm{B}),y}_{\mathbf{g}_{U_1}} G^{(\mathrm{Orc})_0,(x,y)}_{\mathbf{g}_{U_2}} G^{(\mathrm{Orc})_1,(x,y)}_{\mathbf{g}_{U_3}} G^{(\mathrm{Orc})_2,(x,y)}_{\mathbf{g}_{U_4}} G^{(\mathrm{Ora}),(x,y),\mathrm{Full}}_{\mathbf{g}_\Gamma,\mathbf{g}_{B_0}\cdots,\mathbf{g}_{B_{m-1}}} \cdot \\ & G^{(\mathrm{Orc})_2,(x,y)}_{\mathbf{g}_{U_4}} G^{(\mathrm{Orc})_1,(x,y)}_{\mathbf{g}_{U_3}} G^{(\mathrm{Orc})_0,(x,y)}_{\mathbf{g}_{U_2}} G^{(\mathrm{B}),y}_{\mathbf{g}_{U_1}} G^{(\mathrm{A}),x}_{\mathbf{g}_{U_0}}, \end{aligned}$$

where for $P \in \{A,B\}$, we shorten the label (Prover, P) to (P), and remove the superscript "game" in the above equation. We remark that in contrast to $G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})}$, the output $\mathbf{g}_{U_i}$, $i \in [5]$ from $M^{(x,y)}$ are secretly polynomials in $\mathrm{IdPoly}(p, m^{\mathrm{ans}}, m)$ which is consistent with the definition for the 5 polynomials given in Theorem 11.3.1.

Furthermore, we define $M^{(x^{\mathrm{game}},y^{\mathrm{game}}),U_i}$ for $i \in [5]$, and $M^{(x^{\mathrm{game}},y^{\mathrm{game}}),\mathrm{Full}}$ in a similar manner as (11.18) and (11.19). By definition

$$\begin{aligned} M^{(x^{\mathrm{game}},y^{\mathrm{game}}),U_0} &= G^{(\mathrm{Prover,\,A}),x^{\mathrm{game}}} \\ M^{(x^{\mathrm{game}},y^{\mathrm{game}}),U_1} &= G^{(\mathrm{Prover,\,B}),y^{\mathrm{game}}} \\ M^{(x^{\mathrm{game}},y^{\mathrm{game}}),U_{o+2}} &= G^{(\mathrm{Ora})_o,(x^{\mathrm{game}},y^{\mathrm{game}})}_{\overline{\mathbf{g}}}, o \in \{0,1,2\}. \end{aligned}$$

For $(x^{\mathrm{game}}, y^{\mathrm{game}}) \in \mathscr{X}^2$ and output tuple $(\mathbf{g}_{U_0},\cdots,\mathbf{g}_{U_4},\mathbf{g}_\Gamma,\mathbf{g}_{B_0}\cdots,\mathbf{g}_{B_{m-1}})$ from $M^{(x^{\mathrm{game}},y^{\mathrm{game}})}$. We refer to the output as "good" if for a uniformly random $s = (s_0,\cdots,s_4,b_0,\cdots,b_4,z) \sim \mathbb{F}_{2^p}^m$, the following occurs with probability greater than $\frac{1}{2}$:

- $\mathbf{g}_\Gamma(s) = \mathbf{g}_{\mathtt{D}}(s)(\mathbf{g}_{U_0}(s) - b_0)(\mathbf{g}_{U_1}(s) - b_1)(\mathbf{g}_{U_2}(s) - b_2)(\mathbf{g}_{U_3}(s) - b_3)(\mathbf{g}_{U_4}(s) - b_4)$
- $\mathbf{g}_\Gamma(s) = \sum_{i\in[m]} \mathbf{g}_{B_i}(s)\mathbf{zero}(s)$,

where $\mathbf{g}_{\mathtt{D}} = \mathtt{ComputePCP}_\alpha(\langle \mathtt{D}\rangle, n, x^{\mathrm{game}}, y^{\mathrm{game}})$. By Theorem 11.3.4, if the output for $M^{(x^{\mathrm{game}},y^{\mathrm{game}})}$ is a "good" output, then there exist $a,b \in \{0,1\}^*$ with $|a|,|b| \le \log^\alpha(n)$ such that $\mathbf{g}_{U_0} = \mathrm{enc}_\Gamma(a)$ and $\mathbf{g}_{U_1} = \mathrm{enc}_\Gamma(b)$ and $D(x^{\mathrm{game}}, y^{\mathrm{game}}, a, b) = 1$

We now proof the following claim regarding the measurement $M^{(x^{\mathrm{game}},y^{\mathrm{game}})}$

Claim 11.5.3. *On average over $(x^{game}, y^{game}) \sim \mu_n$ and the state $|\tau\rangle$*

$$M^{(x^{game},y^{game})} \simeq_{O(\mathrm{poly}(\varepsilon,\log(n),\alpha))} (M^{(x^{game},y^{game})})^{op}. \tag{11.25}$$

Furthermore, on average over (x^{game}, y^{game}), the measurement output for $\langle\tau|M^{(x^{game},y^{game})}|\tau\rangle$ is "good" with probability at least $1 - O(\mathrm{poly}(\varepsilon,\log(n),\alpha))$

Proof. We first show that, on average over $(x^{\mathrm{game}}, y^{\mathrm{game}}) \sim \mu_n$ and the state $|\tau\rangle$

$$M^{(x^{\mathrm{game}},y^{\mathrm{game}})} \simeq_{O(\mathrm{poly}(\varepsilon,\log(n),\alpha))} G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})} \tag{11.26}$$

Since $G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})}$ is projective, by the definition of $M^{(x^{\mathrm{game}},y^{\mathrm{game}})}$, the last $1+m$ measurement outcome for $G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})}$ $(\mathbf{g}_\Gamma, \mathbf{g}_{B_0},\cdots,\mathbf{g}_{B_{m-1}})$ will always be the same as the measurement $M^{(x^{\mathrm{game}},y^{\mathrm{game}})}$ when the two measurements are made simultaneously (on any state). For $i \in [5]$, by Claim 11.5.2 and the Schwartz-Zippel lemma (Lemma 7.1.4)

$$M^{(x^{\mathrm{game}},y^{\mathrm{game}}),U_i}_{\mathbf{g}_{U_i}} \simeq_{\delta_2} (G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}}),U_i}_{\mathbf{g}_{U_i}})^{op}$$

for $\delta_2 = O(\mathrm{poly}(\varepsilon, \log(n), \alpha)) + \frac{m \cdot d}{2^p}$. Hence, by repeatedly applying Lemma 3.3.4, the underlying vector state is a tracial state, and using Lemma 3.3.3 to convert between $\simeq$ distance to $\approx$ distance,

$$\begin{aligned}
&G^{(\mathrm{Ora}),(x,y)} = G^{(x,y),U_0} \cdots G^{(x,y),U_4} G^{(x,y),\mathrm{Full}} G^{(x,y),U_4} \cdots G^{(x,y),U_0} \\
&\approx_{\delta_2} (M^{(x,y),U_0})^{op} G^{(x,y),U_0} \cdots G^{(x,y),U_4} G^{(x,y),\mathrm{Full}} G^{(x,y),U_4} \cdots G^{(x,y),U_1} \\
&\cdots \\
&\approx_{\delta_2} (M^{(x,y),U_4} \cdots M^{(x,y),U_0})^{op} G^{(x,y),U_0} \cdots G^{(x,y),U_4} G^{(x,y),U_4} G^{(x,y),\mathrm{Full}} \\
&= (M^{(x,y),\mathrm{Full}} M^{(x,y),U_4} \cdots M^{(x,y),U_0})^{op} G^{(x,y),U_0} \cdots G^{(x,y),U_4} \\
&\approx_{\delta_2} (M^{(x,y),U_4} M^{(x,y),\mathrm{Full}} M^{(x,y),U_4} \cdots M^{(x,y),U_0})^{op} G^{(x,y),U_0} \cdots G^{(x,y),U_3} \\
&\cdots \\
&\approx_{\delta_2} (M^{(x,y),U_0} \cdots M^{(x,y),U_4} M^{(x,y),\mathrm{Full}} M^{(x,y),U_4} \cdots M^{(x,y),U_0})^{op} = (M^{(x,y)})^{op}
\end{aligned}$$

where we remove the superscript "game" and (Ora) in the above derivation for clarity. Hence, by using the triangle inequality for $\approx$ distance and Lemma 3.3.3, this implies that

$$G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})} \simeq_{10\delta_2} (M^{(x^{\mathrm{game}},y^{\mathrm{game}})})^{op}$$

and since the underlying state is a tracial state and $\delta_2 = O(\mathrm{poly}(\varepsilon, \log(n), \alpha))$, this shows (11.26). Since the underlying state for (11.26) is the tracial state $|\tau\rangle$, we also have

$$(M^{(x^{\mathrm{game}},y^{\mathrm{game}})})^{op} \simeq_{O(\mathrm{poly}(\varepsilon,\log(n),\alpha))} (G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})})^{op} \tag{11.27}$$

For (11.25), since $G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})}$ are all projective measurements,

$$G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})} \simeq_0 (G^{(\mathrm{Ora}),(x^{\mathrm{game}},y^{\mathrm{game}})})^{op} \tag{11.28}$$

over $(x^{\mathrm{game}}, y^{\mathrm{game}}) \sim \mu_n$ and the tracial state $|\tau\rangle$. Equation (11.25) then follows from Lemma 3.3.3 and the triangle inequality of $\approx$ distance being applied to (11.26), (11.27) and (11.28).

For the second part of Claim 11.5.3, by applying the data processing inequality to (11.26),

$$M^{(x^{\mathrm{game}},y^{\mathrm{game}})}_{[\mathbf{eval}_s^{6+m}|(u_0,\cdots,u_4,\gamma,\beta_0,\cdots,\beta_{m-1})]} \simeq_{O(\mathrm{poly}(\varepsilon,\log(n),\alpha))} A^{(\mathrm{Ora}),(\mathrm{Point}),((x^{\mathrm{game}},y^{\mathrm{game}}),s)}_{[\mathbf{eval}_s^{6+m}|(u_0,\cdots,u_4,\gamma,\beta_0,\cdots,\beta_{m-1})]}. \tag{11.29}$$

Hence by applying the triangle inequality for $\approx$ distance and Lemma 3.3.3 to (11.29) and (11.17), we obtain

$$M^{(x^{\mathrm{game}},y^{\mathrm{game}})}_{[\mathbf{eval}_s^{6+m}|(u_0,\cdots,u_4,\gamma,\beta_0,\cdots,\beta_{m-1})]} \simeq_{O(\mathrm{poly}(\varepsilon,\log(n),\alpha))} A^{(\mathrm{Ora}),(\mathrm{Point}),((x^{\mathrm{game}},y^{\mathrm{game}}),s)}_{u_0,\cdots,u_4,\gamma,\beta_0,\cdots,\beta_{m-1}}. \tag{11.30}$$

Recall that $\mathscr{S}$ is a synchronous strategy which succeed at $\mathscr{G}^{\mathrm{AR}}$ with probability at least $1 - c_1 \cdot \delta_1$. Since the oracularization question label is pick with constant probability, by the "PCPP proof check" clause of Figure 11.3, on expectation over $(x^{\mathrm{game}}, y^{\mathrm{game}}) \sim \mu$ and $s = (s_0, \cdots, s_4, b_0, \cdots, b_4, z) \in \mathbb{F}_{2^p}^m$, the measurement

$$\langle \tau | A^{(\mathrm{Orc}),(\mathrm{Point}),((x^{\mathrm{game}},y^{\mathrm{game}}),s)}_{(u_0,\cdots,u_4,\gamma,\beta_0,\cdots,\beta_{m-1})} | \tau \rangle$$

outputs the answer which satisfies the properties below with probability $1 - c_2 c_1 \cdot \delta_1$

1. $\gamma = \mathbf{g}_{\mathtt{D}}(s) \cdot (u_1 - b_0) \cdots (u_4 - b_4)$,
2. $\gamma = \sum_{i \in [m]} \beta_i \cdot \mathbf{zero}(s_i)$,

where $\mathbf{g}_{\mathtt{D}} = \mathtt{ComputePCP}_\alpha - (\langle \mathtt{D} \rangle, n, x^{\mathrm{game}}, y^{\mathrm{game}})$. Pick $c_1 \in (0,1)$ used to define $\mathscr{S}$ such that $1 - c_2 c_1 \cdot \delta_1 \geq \frac{1}{2}$. Combine this with Equation (11.30), this shows that on average over $(x^{\mathrm{game}}, y^{\mathrm{game}}) \sim \mu_n$, the probability that $M^{(x^{\mathrm{game}},y^{\mathrm{game}})}$ gives an output which is "good" with probability at least $1 - = O(\mathrm{poly}(\varepsilon, \log(n), \alpha))$, thus completing the claim for the lemma. □

Base on the POVM $M^{(x^{\text{game}},y^{\text{game}})}$, we define a symmetric strategy $\mathscr{S}^{\text{Ora}} = (\mathscr{L}^2(\mathscr{A},\tau),|\tau\rangle,\{B_a^x\})$ for $\mathscr{G}^{\text{Ora}}$ as follows: Fixed $(x^{\text{game}},y^{\text{game}}) \in \mathscr{X}_n$, the measurement operator $\{B_a^{(\text{Prover, A}),x^{\text{game}}}\}$ as a data processing measurement as follows:

- Perform the measurement $M_{\mathbf{g}_{U_0}}^{x^{\text{game}},U_0}$ and obtain a polynomial $\mathbf{g}_{U_0}$.
- If there exist an $a \in \mathscr{A}_n$ such that $|a| \leq \log^{\alpha}(n)$ and $\mathbf{g}_{U_0} = \text{enc}$, output a. Otherwise, output 0.

The measurement operator $\{B^{(\text{Prover, B}),y^{\text{game}}}\}$ is defined in a similar manner as $\{B^{(\text{Prover, A}),x^{\text{game}}}\}$ with the measurement $M_{\mathbf{g}_{U_1}}^{x^{\text{game}},U_1}$. The measurement operator $\{B^{(\text{Orac}),(x^{\text{game}},y^{\text{game}})}\}$ similarly define as a data processing measurement as follows:

- Perform the measurement $M^{(x^{\text{game}},y^{\text{game}})}$ and obtain the tuple of polynomials $(\mathbf{g}_v)$.
- If the given measurement outcome is a "good" output, by Theorem 11.3.4, there exist $(a,b) \in \mathscr{A}_n^2$ such that $\mathbf{g}_{U_0} = \text{enc}_\Gamma(a)$, $\mathbf{g}_{U_1} = \text{enc}_\Gamma(b)$, output the corresponding $(a,b) \in \mathscr{A}_n^2$. Otherwise, output $(0,0)$.

We remark that the above strategy is not necessarily synchronous strategy, since $\{M^{(x^{\text{game}},y^{\text{game}})}\}$ does not necessarily have to be a PVM. We make the following claim about $\mathscr{S}^{\text{Ora}}$.

Claim 11.5.4. $\omega(\mathscr{G}^{\textbf{Orac}},\mathscr{S}^{Ora}) \geq 1 - O(\text{poly}(\varepsilon,\log(n),\alpha))$.

Proof. We consider the performance of $\mathscr{S}^{\text{Ora}}$ for different question pairs given in Figure 11.1 below:

- (Prover, P) - (Prover, P), $P \in \{A,B\}$: Since both $M^{x^{\text{game}},U_A} = G^{(\text{Prover, A}),x^{\text{game}}}$ and $M^{y^{\text{game}},U_B} = G^{(\text{Prover, B}),y^{\text{game}}}$ are both projective and $\mathscr{S}^{\text{Ora}}$ uses the tracial state as the underlying state. This implies that $\mathscr{S}^{\text{Ora}}$ always succeed on this question pair.
- (Oracularization) – (Oracularization): For the "consistency" part of this question pair, $B^{(\text{Orac}),(x^{\text{game}},y^{\text{game}})}$ is a data processed measurement of $M^{(x^{\text{game}},y^{\text{game}})}$, which by (11.25), are consistent with probability at least $1 - O(\text{poly}(\varepsilon,\log(n),\alpha))$. For the "proof checking" part of this question pair,, whenever $M^{(x^{\text{game}},y^{\text{game}})}$ returns a "good" output when performing the measurement $B^{(\text{Orac}),(x^{\text{game}},y^{\text{game}})}$, the corresponding output $(a,b) \in \mathscr{A}_n^2$ always satisfies $D_n(x,y,a,b) = 1$ by Theorem 11.3.4. By Claim 11.5.3, this occurs with $1 - O(\text{poly}(\varepsilon,\log(n),\alpha))$. Combining these two facts, this implies that $\mathscr{S}^{\text{Ora}}$ succeed on this question pair with probability at least $1 - O(\text{poly}(\varepsilon,\log(n),\alpha))$.
- (Oracularization) – (Prover, P), $P \in \{A,B\}$: Restricted to the case when the provers receiving the question label " (Oracularization)" and obtain a "good" outcome from the measurement of $M^{(x^{\text{game}},y^{\text{game}})}$, by the definition of M, the " (Prover, P)" prover would receive the same polynomial from his/her measurement output, and hence output a consistent answer label as the " (Oracularization)" prover. Since a "good" outcome occurs with probability $1 - O(\text{poly}(\varepsilon,\log(n),\alpha))$, this implies that $\mathscr{S}^{\text{Ora}}$ succeed with probability on this question pair with probability at least $1 - O(\text{poly}(\varepsilon,\log(n),\alpha))$.

By averaging out the probability given above, we see that $\omega(\mathscr{G}^{\textbf{Orac}},\mathscr{S}^{\text{Ora}}) > 1 - O(\text{poly}(\varepsilon,\log(n),\alpha))$, completing the proof of the claim. □

This shows that for model $t \in \{*,co\}$, $\omega^t(\mathscr{G}) > 1-\varepsilon$ implies that $\omega^t(\mathscr{G}^{\text{Ora}}) > 1 - O(\text{poly}(\varepsilon,\log(n),\alpha))$. The proof of Proposition 7.4.3 then follows from the "soundness" clause of Lemma 11.2.1.

CHAPTER 12

Parallel repetition

The result from this chapter is based on joint work with William Slofstra and Henry Yuen. The goal is to prove Theorem 12.2.3. The anchored parallel repetition was originally shown in [BVY21] in the tensor product model. In this chapter, we define some quantum information-theoretic tools for tracially embeddable strategies, and show how this can be used to show a parallel repetition for the commuting operator model. In addition to the parallel repetition theorem, we also introduce several generalizations of commonly used identities in quantum information theory, such as Uhlmann's theorem and mutual information, in the commuting operator model. We believe these generalizations could be of independent interest.

This chapter is organized as follows: in Section 12.1, we first establish the quantum information tools needed for this chapter. We define the anchored parallel repetition theorem in Section 12.2, and prove it in Section 12.3.

12.1 Quantum information theory for tracial von Neumann algebras

We start this section by recalling some additional theorems about von Neumann algebras necessary for establishing the quantum information tools. Let $\mathscr{A}_1 \subseteq \mathscr{B}(\mathscr{H}_1)$ and $\mathscr{A}_2 \subseteq \mathscr{B}(\mathscr{H}_2)$ be two von Neumann algebras. The von Neumann algebra tensor product $\mathscr{A}_1 \otimes \mathscr{A}_2$ is the weak operator closure of the span of $\{a \otimes b : a \in \mathscr{A}_1, b \in \mathscr{A}_2\}$ in $\mathscr{B}(\mathscr{H}_1 \otimes \mathscr{H}_2)$. For two tracial von Neumann algebras $(\mathscr{A}_1, \tau_1)$ and $(\mathscr{A}_2, \tau_2)$, the von Neumann algebra $\mathscr{A}_1 \otimes \mathscr{A}_2$ remains a tracial von Neumann algebra with the (unique) trace being $\tau_1 \otimes \tau_2$. For a state $\psi^{\mathscr{A}_1\mathscr{A}_2}$ on $\mathscr{A}_1 \otimes \mathscr{A}_2$, we let $\psi^{\mathscr{A}_1}$ denote the restriction of $\psi^{\mathscr{A}_1\mathscr{A}_2}$ to $\mathscr{A}_1 = \mathscr{A}_1 \otimes \mathbb{I}$, so $\psi^{\mathscr{A}_1}(a) = \psi^{\mathscr{A}_1\mathscr{A}_2}(a \otimes \mathbb{I})$ for all $a \in \mathscr{A}_1$. We recall the following theorem about projectors in a tracial von Neumann algebra.

Proposition 12.1.1 (Corollary 2.8 of [Tak01])**.** *Let $(\mathscr{A}, \tau)$ be a tracial von Neumann algebra and $P, Q \in \mathscr{A}$ be two projectors, then the following two conditions are equivalent:*

- $\tau(P) = \tau(Q)$.
- *P is equivalent to Q.*

12.1.1 Tomita-Takesaki construction and the positive cone

In order to discuss the relative entropy construction, we need to first give a brief summary to the well known Tomita-Takesaki construction [Tak70] in operator algebra. Intuitively, the Tomita-Takesaki construction is a way to define the commutant of a von Neumann algebra algebraically. We remark that although there is a rich theory behind the Tomita-Takesaki construction for general von Neumann algebras, we only need to focus on the tracial case for this section since we are only working with tracially embeddable strategies in this thesis.

Let $(\mathscr{A}, |\tau\rangle)$ be a tracial von Neumann algebra in standard form, and let $\mathfrak{S}_{|\tau\rangle} : \mathscr{H} \to \mathscr{H}$ be the antilinear (and potentially unbounded) map defined by

$$\mathfrak{S}_{|\tau\rangle} : a|\tau\rangle \to a^*|\tau\rangle, a \in \mathscr{A}. \tag{12.1}$$

Since $|\tau\rangle$ is cyclic, the above map is well-defined on the dense subset $\mathscr{A}|\tau\rangle$ of $\mathscr{H}$. By taking the polar decomposition, we can write $\mathfrak{S}_{|\tau\rangle}$ as

$$\mathfrak{S}_{|\tau\rangle} = \mathfrak{J}_{|\tau\rangle}\Delta_{|\tau\rangle}^{\frac{1}{2}}, \tag{12.2}$$

for some antilinear isometry $\mathfrak{J}_{|\tau\rangle}$, known as the *modular conjugation*, and some positive operator $\Delta_{|\tau\rangle}$, known as the *modular operator*. Remarkably the modular operator can be used to construct the commutant of $\mathscr{A}$ algebraically, as $\mathfrak{J}\mathscr{A}\mathfrak{J} = \mathscr{A}'$. Using the modular operator, define the *canonical positive cone* associated with $(\mathscr{A}, |\tau\rangle)$ as

$$\mathscr{H}_{|\tau\rangle}^+ = \overline{\{\Delta_{|\tau\rangle}^{\frac{1}{4}} A|\tau\rangle, A \in \mathscr{A}^+\}}, \tag{12.3}$$

where the closure is in the weak topology [Ara74]. This is a pointed closed self-dual convex cone [Ara74, Theorem 4 part 1]. The following proposition gives a bijection between states on the von Neumann algebra $\mathscr{A}$ and vectors on the canonical positive cone.

Proposition 12.1.2 (Theorem 6 of [Ara74])**.** *For every positive normal linear functional ψ on a von Neumann algebra $\mathscr{A} \subseteq \mathscr{B}(\mathscr{H})$, there exists a unique $|\psi\rangle \in \mathscr{H}_{|\tau\rangle}^+$ such that $\psi(a) = \langle\psi|a|\psi\rangle$.*

For a state ψ, we will use $|\psi\rangle$ to denote the unique vector in the positive cone above. For a state ψ acting on $\mathscr{A}$, we use $\mathrm{supp}^{\mathscr{A}}\psi$ to denote the complement of the minimal projector $P \in \mathscr{A}$ such that $\psi(P) = 0$. For any vector ψ, we have $\mathrm{supp}(\psi)|\psi\rangle = |\psi\rangle$. Intuitively, we can think of the corresponding vector in the positive cone as the "purification" of the linear functional on $\mathscr{A}$. Indeed, in the finite-dimensional case this vector is the familiar purification from quantum information theory:

Example 12.1.3. Let $\mathscr{A} = \mathbf{M}_n(\mathbb{C}) \otimes \mathbb{I}_n \subseteq \mathbf{M}_{n^2}(\mathbb{C})$, and let $|i\rangle, i = 0, \ldots, n-1$ denote the standard basis vectors for $\mathbb{C}^n$. Recall, the GNS representation using the trace $\mathrm{Tr}(\cdot)$ maps $\mathbf{M}_n(\mathbb{C})$ to $\mathbf{M}_n(\mathbb{C}) \otimes \mathbb{I}_n \in \mathbf{M}_{n^2}(\mathbb{C})$, with the linear functional $\mathrm{Tr}(\cdot)$ gets mapped to the maximally entangled state $|\tau\rangle = \frac{1}{\sqrt{n}}\sum_{i=0}^n |ii\rangle$. In this case, we see that the vector state $|\tau\rangle$ is a cyclic and separating vector for $\mathscr{A}$. As per Equation (12.1),

$$S_{|\tau\rangle}|i\rangle|j\rangle = S_{|\tau\rangle}(|i\rangle\langle j| \otimes \mathbb{I})|\tau\rangle = (|j\rangle\langle i| \otimes \mathbb{I})|\tau\rangle = |j\rangle|i\rangle,$$

and hence $S_{|\tau\rangle} = C \circ \sum_i \sum_j \frac{\lambda_i}{\lambda_j} |j,i\rangle\langle i,j|$, where C is conjugation in the standard basis. Taking the polar decomposition, the modular operator is

$$\Delta_{|\tau\rangle} = (S_{|\tau\rangle})^* S_{|\tau\rangle} = \sum_i \sum_j |i,j\rangle\langle i,j| = \mathbb{I}_4,$$

and the positive cone associated with $|\tau\rangle$ is

$$\mathscr{H}_{|\tau\rangle}^+ = \left\{(A \otimes \mathbb{I})|\tau\rangle, A \in \mathbf{M}_n(\mathbb{C})^+\right\}.$$

If ϕ is a linear functional on $\mathbf{M}_n(\mathbb{C})$, then there is a unique positive matrix σ such that $\phi(A) = \mathrm{Tr}(A\sigma) = \mathrm{Tr}(\sigma^{1/2}A\sigma^{1/2})$, and hence

$$\phi(A) = \langle\tau|\sigma^{\frac{1}{2}}A\sigma^{\frac{1}{2}} \otimes \mathbb{I}|\tau\rangle. \tag{12.4}$$

Thus the vector in $\mathscr{H}_{|\tau\rangle}^+$ associated with ϕ is $(\sigma^{\frac{1}{2}} \otimes \mathbb{I})|\tau\rangle$, which is commonly used in quantum information as a purification of the density matrix σ.

We end this subsection by reviewing some properties related to the positive cone $\mathscr{H}^+_{|\tau\rangle}$. The following proposition shows that changing the cyclic and separating vector $|\tau\rangle$ only changes the positive cone by a unitary in the commutant:

Proposition 12.1.4 (Theorem 7 part 6 of [Ara74])**.** *Let $\mathscr{A} \subseteq \mathscr{B}(\mathscr{H})$ be a tracial von Neumann algebra in standard form and let $|\tau_1\rangle$ and $|\tau_2\rangle$ be two cyclic and separating vectors for $\mathscr{A}$. Then there is a unitary $U \in \mathscr{A}'$ such that for all normal states ψ on $\mathscr{A}$, if $|\psi_1\rangle$ and $|\psi_2\rangle$ are the vectors associated to ψ in the positive cones of $|\tau_1\rangle$ and $|\tau_2\rangle$ respectively, then $|\psi_1\rangle = U|\psi_2\rangle$.*

The following proposition shows that every vector on $\mathscr{H}$ can be related to some vector within $\mathscr{H}^+_{|\tau\rangle}$ via a partial isometry.

Proposition 12.1.5. *Let $(\mathscr{A}, |\tau\rangle) \subseteq \mathscr{B}(\mathscr{H})$ be a tracial von Neumann algebra in standard form. For any $|\psi\rangle \in \mathscr{H}$, there exist a unitary $U \in \mathscr{A}'$ and a unique $|\psi^+\rangle \in \mathscr{H}^+_{|\tau\rangle}$ such that $|\psi\rangle = U|\psi^+\rangle$.*

Proof. By [Ara74, Theorem 7 part 5], there exist a partial isometry $V \in \mathscr{A}'$ and a unique $|\psi^+\rangle \in \mathscr{H}^+_{|\tau\rangle}$ such that $|\psi\rangle = V|\psi^+\rangle$, and

$$VV^* = \text{supp}^{\mathscr{A}'}(\psi), \quad V^*V = \text{supp}^{\mathscr{A}'}(\psi^+).$$

We wish to extend V into a Unitary. We first see that $\tau(VV^*) = \tau(V^*V)$, and hence $\tau(\mathbb{I} - VV^*) = \tau(I - V^*V)$. By Proposition 12.1.1, there exist a partial isometry W such that

$$WW^* = \mathbb{I} - \text{supp}^{\mathscr{A}'}(\psi), \quad W^*W = \mathbb{I} - \text{supp}^{\mathscr{A}'}(\psi^+).$$

Take $U = V + W$, we see that

$$\begin{aligned} U^*U &= (V^* + W^*)(V + W) = \mathbb{I} + W^*V + V^*W \\ &= \mathbb{I} + W^*\left(\mathbb{I} - \text{supp}^{\mathscr{A}'}(\psi)\right)\text{supp}^{\mathscr{A}'}(\psi)V + V^*\text{supp}^{\mathscr{A}'}(\psi^+)\left(\mathbb{I} - \text{supp}^{\mathscr{A}'}(\psi^+)\right)W \\ &= \mathbb{I}, \end{aligned}$$

and by a similar calculation, we have $UU^* = \mathbb{I}$ showing that U is indeed a unitary. Furthermore

$$U|\psi^+\rangle = V|\psi^+\rangle + W|\psi^+\rangle = |\psi\rangle + W\left(\mathbb{I} - \mathbb{I} - \text{supp}^{\mathscr{A}'}(\psi^+)\right)|\psi^+\rangle = |\psi\rangle,$$

and hence the proposition follows. □

The Araki-Powers-Stormer inequality relates the norm distance between states to the distance between vectors in the positive cone:

Proposition 12.1.6 (Araki-Powers-Stormer inequality, Theorem 4 part 8 of [Ara74])**.** *Let $\mathscr{A} \subseteq \mathscr{B}(\mathscr{H})$ be a tracial von Neumann algebra in standard form, and let $|\psi_1\rangle, |\psi_2\rangle \in \mathscr{H}^+_{|\tau\rangle}$. Then*

$$\|\psi_1 - \psi_2\|_{\mathscr{A}} \geq \| |\psi_1\rangle - |\psi_2\rangle \|^2.$$

We'll use the Araki-Powers-Stormer inequality in the following form:

Proposition 12.1.7. *Let $(\mathscr{A}, |\tau\rangle) \subseteq \mathscr{B}(\mathscr{H})$ be a von Neumann algebra in the standard form. For any unit vectors $|\psi_1\rangle, |\psi_2\rangle \in \mathscr{H}$, there is a unitary operator $U \in \mathscr{A}'$ such that*

$$\langle\psi_1|U|\psi_2\rangle \geq 1 - \frac{1}{2}\|\psi_1 - \psi_2\|_{\mathscr{A}}. \tag{12.5}$$

Proof. By Proposition 12.1.5, there are vectors $|\psi_1^+\rangle, |\psi_2^+\rangle \in \mathscr{H}^+_{|\tau\rangle}$ and unitaries $U_{\psi_2}, U_{\psi_1} \in \mathscr{A}'$ such that $U_{\psi_i}|\psi_i^+\rangle = |\psi_i\rangle$, $i = 1, 2$. Since $\mathscr{H}^+_{|\tau\rangle}$ is self-dual, $\langle\psi_1^+|\psi_2^+\rangle \geq 0$, so

$$\| |\psi_1^+\rangle - |\psi_2^+\rangle \| = 2 - 2\langle\psi_1^+|\psi_2^+\rangle = 2 - 2\langle\psi_1|U_{\psi_1}^*U_{\psi_2}|\psi_2\rangle,$$

and hence the proposition follows from Proposition 12.1.6 with $U = U_{\psi_1}^*U_{\psi_2}$. □

In the finite-dimensional case, we can take V to be unitary in Proposition 12.1.7. Indeed, let $\mathscr{A} = \mathbf{M}_n(\mathbb{C}) \otimes \mathbb{I} \subseteq \mathbf{M}_{n^2}(\mathbb{C})$, and let $|\psi_1\rangle^{AB}$ and $|\psi_2\rangle^{AB}$ be two unit vectors in $\mathbb{C}^{n^2}$ with reduced density matrices $\psi_i = \mathrm{Tr}_B(|\psi_i\rangle\langle\psi_i|)$, $i = 1,2$. By Uhlmann's theorem (see, e.g. [Wil13, Theorem 9.2.1]) there exist a unitary $U \in \mathbf{M}_n(\mathbb{C})$ such that

$$\langle\psi_1|\mathbb{I} \otimes U|\psi_2\rangle = \|\sqrt{\psi_1}\sqrt{\psi_2}\|_1, \tag{12.6}$$

where $\|\cdot\|_1$ is the matrix 1-norm. We remark that $\|\sqrt{\psi_1}\sqrt{\psi_2}\|_1$ is precisely the formula for quantum fidelity between two quantum states. By the Fuchs-van de Graaf inequality (see, e.g. [Wil13, Theorem 9.3.1]),

$$1 - \|\sqrt{\psi_1}\sqrt{\psi_2}\|_1 \leq \frac{1}{2}||\psi_1 - \psi_2||_1, \tag{12.7}$$

and since

$$\|\psi\|_1 = \sup\{|\mathrm{Tr}(A^*\psi)\| : A \in \mathbf{M}_n(\mathbb{C}), ||A|| \leq 1\} = \|\psi\|_{\mathbf{M}_n(\mathbb{C})},$$

Equation (12.5) follows from Equations (12.6) and (12.7), with $V = \mathbb{I} \otimes U \in \mathscr{A}'$.

12.1.2 Relative entropy

In this subsection, we review the relative entropy between two positive normal linear functionals ψ_1 and ψ_2 on a tracial von Neumann algebra $(\mathscr{A}, |\tau\rangle)$ in standard form. We follow the construction in [Ara77, Section 2]. Let $|\psi_1\rangle$ and $|\psi_2\rangle$ be the vectors corresponding to ψ_1 and ψ_2 in the positive cone $\mathscr{H}^+_{|\tau\rangle}$ of $\mathscr{A}$. Similarly to Equation (12.1), we define an anti-linear map $\mathfrak{S}^{|\tau\rangle}_{\psi_1,\psi_2} : \mathscr{A}|\psi_2\rangle \to \mathrm{supp}(\psi_1)\mathscr{H}$ as

$$\mathfrak{S}^{|\tau\rangle}_{\psi_1,\psi_2}(a|\psi_2\rangle) = \mathrm{supp}(\psi_2)a^*|\psi_1\rangle \text{ for all } a \in \mathscr{A}. \tag{12.8}$$

Extend the above (potentially unbounded) map to $\mathscr{A}|\psi_2\rangle \oplus (\mathscr{A}|\psi_2\rangle)^\perp$ by mapping $(\mathscr{A}|\psi_1\rangle)^\perp$ to zero. The extension is a well-defined antilinear operator whose domain of definition is dense in $\mathscr{H}$. Hence, $\mathfrak{S}^{|\tau\rangle}_{\psi_1,\psi_2}$ has a polar decomposition

$$\mathfrak{S}^{|\tau\rangle}_{\psi_1,\psi_2} = \mathfrak{S}^{|\tau\rangle}_{\psi_1,\psi_2}(\Delta^{|\tau\rangle}_{\psi_1,\psi_2})^{\frac{1}{2}} \tag{12.9}$$

where $\mathfrak{J}^{|\tau\rangle}_{\psi_1,\psi_2}$ is an antilinear isometry and $\Delta^{|\tau\rangle}_{\psi_1,\psi_2} \in \mathscr{A}$ is a positive operator. The relative entropy between ψ_1 and ψ_2 is defined to be

$$\mathrm{D}^{|\tau\rangle}(\psi_1\|\psi_2) = \begin{cases} \int_0^\infty \log_2 \lambda d\langle\psi_1|E_\lambda|\psi_1\rangle & \text{if } supp(\psi_1) \leq supp(\psi_2) \\ +\infty & \text{otherwise} \end{cases}, \tag{12.10}$$

where $\Delta^{|\tau\rangle}_{\psi_1,\psi_2} = \int_0^\infty \lambda dE^{|\tau\rangle}_\lambda$ is the spectral decomposition of $\Delta^{|\tau\rangle}_{\psi_1,\psi_2}$. By Proposition 12.1.4, $\mathrm{D}^{|\tau\rangle}(\psi_1\|\psi_2)$ is independent of $|\tau\rangle$, and thus we refer to this relative entropy as $\mathrm{D}(\psi_1\|\psi_2)$. To match the standard convention in quantum information, the order of arguments for D is flipped from [Ara77, Definition 3.1]. We also use $\log = \log_2$ rather than $\log = \ln$, which changes our relative entropy by a factor of $\ln(2)$. Most propositions below are unchanged, but we do get a factor of $\ln(2)$ in Proposition 12.1.12 below.

Example 12.1.8. Let $\psi_1(A) = \mathrm{Tr}(A\rho_1)$ and $\psi_2(A) = \mathrm{Tr}(A\rho_2)$ be two normal linear functionals on $\mathscr{A} = \mathbf{M}_n(\mathbb{C})$ such that $supp(\psi_1) \leq supp(\psi_2)$. Recall from Example 12.1.3 from that $\mathscr{A}$ has standard form $\mathbf{M}_n(\mathbb{C}) \otimes \mathbb{I} \subseteq \mathbf{M}_{n^2}(\mathbb{C})$, and that $|\tau\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^n |ii\rangle$ is a cyclic and separating vector for $\mathscr{A}$. The vector corresponding to ψ_i in the positive cone for $|\tau\rangle$ is $|\psi_i\rangle = \rho_i^{\frac{1}{2}} \otimes \mathbb{I}|\tau\rangle$, $i = 1,2$. Hence

$$S^{|\tau\rangle}_{\psi_1,\psi_2}(a\rho_2^{\frac{1}{2}} \otimes I)|\tau\rangle = ((\mathrm{supp}(\psi_2)a^*\rho_1^{\frac{1}{2}}) \otimes \mathbb{I})|\tau\rangle \text{ for all } a \in \mathscr{A},$$

and this implies that

$$S^{|\tau\rangle}_{\psi_1,\psi_2}(a\otimes I)|\tau\rangle = (\rho_2^{-\frac{1}{2}} a^* \rho_1^{\frac{1}{2}}\otimes \mathbb{I})|\tau\rangle = S_{|\tau\rangle}\rho_1^{1/2} a\rho_2^{-1/2}\otimes\mathbb{I}|\tau\rangle = S_{|\tau\rangle}(\rho_1^{1/2}\otimes(\rho_2^{-1/2})^T)(a\otimes\mathbb{I})|\tau\rangle$$

for all $a\in\mathscr{A}$, where $S_{|\tau\rangle}$ is the modular operator for the vector $|\tau\rangle$, and $\rho_2^{-1/2}$ is the pseudoinverse of $\rho_2^{1/2}$. Since $S^*_{|\tau\rangle}S_{|\tau\rangle}=\mathbb{I}$, the relative modular operator between ψ_1 and ψ_2 is

$$\Delta^{|\tau\rangle}_{\psi_1,\psi_2} = \rho_1\otimes(\rho_2^{-1})^T.$$

Since both ρ_2 and ρ_1 are finite-dimensional matrices, the relative entropy is

$$\mathrm{D}^{|\tau\rangle}(\psi_1\|\psi_2) = \langle\psi_1|\log\Delta^{|\tau\rangle}_{\psi_1,\psi_2}|\psi_1\rangle = \langle\tau|(\rho_1^{\frac{1}{2}}\otimes\mathbb{I})\log(\rho_1\otimes(\rho_2^{-1})^T)(\rho_1^{\frac{1}{2}}\otimes\mathbb{I})|\tau\rangle = \mathrm{Tr}(\rho_1\log(\rho_1)-\rho_1\log(\rho_2)),$$

which is the standard definition of von Neumann entropy.

We remark that the definition of relative entropy holds for general von Neumann algebras in standard form. However, we will only focus on the case where $\mathscr{A}$ is tracial, as we primarily work with tracially embeddable strategies in this thesis. We refer to [Ara77; OP04] for more details. We use the following properties of relative entropy:

Proposition 12.1.9 (Theorem 3.6 of [Ara77])**.** *If ψ_1,ψ_2, and ψ_3 are positive normal linear functionals on a von Neumann algebra $\mathscr{A}$, then:*

1. *If $\psi_1(\mathbb{I})=\psi_2(\mathbb{I})$, then $\mathrm{D}(\psi_1\|\psi_2)\geq 0$, and $\mathrm{D}(\psi_1\|\psi_2)=0$ if and only if $\psi_1=\psi_2$,*
2. *$\mathrm{D}(\alpha\psi_1\|\beta\psi_2) = \alpha\mathrm{D}(\psi_1\|\psi_2)-\alpha\cdot\psi_1(\mathbb{I})\cdot\log(\frac{\beta}{\alpha})$ for all $\alpha,\beta>0$, and*
3. *if $\psi_2\leq\psi_3$ (meaning that $\psi_2(a)\leq\psi_3(a)$ for all $a\in\mathscr{A}^+$), then $\mathrm{D}(\psi_1\|\psi_3)\leq\mathrm{D}(\psi_1\|\psi_2)$.*

A linear map $\alpha:\mathscr{A}_1\to\mathscr{A}_2$ between C^*-algebras is said to satisfy the Schwarz inequality if

$$\alpha(a^*a)\geq\alpha(a)^*\alpha(a)$$

for all $a\in\mathscr{A}_1$.

Proposition 12.1.10 (Uhlmann monotonicity theorem, Theorem 5.3 of [OP04])**.** *Let $\alpha:\mathscr{A}_1\to\mathscr{A}_2$ be a unital map (meaning $\alpha(\mathbb{I}_{\mathscr{A}_1})=\mathbb{I}_{\mathscr{A}_2}$) which satisfies the Schwarz inequality, and let $\psi_1^{\mathscr{A}_i},\psi_2^{\mathscr{A}_i}$ be positive normal linear functionals on $\mathscr{A}_i$, $i=1,2$, such that $\psi_1^{\mathscr{A}_2}\circ\alpha\leq\psi_1^{\mathscr{A}_1}$ and $\psi_2^{\mathscr{A}_2}\circ\alpha\leq\psi_2^{\mathscr{A}_1}$. Then*

$$\mathrm{D}(\psi_1^{\mathscr{A}_1}\|\psi_2^{\mathscr{A}_1})\leq\mathrm{D}(\psi_1^{\mathscr{A}_2}\|\psi_2^{\mathscr{A}_2}).$$

Proposition 12.1.11 (Additivity under direct sums)**.** *Let $\mathscr{A}_1\subseteq\mathscr{B}(\mathscr{H}_1)$ and $\mathscr{A}_2\subseteq\mathscr{B}(\mathscr{H}_2)$ be two von Neumann algebras in standard form. Let $\psi_1^{\mathscr{A}_i},\psi_2^{\mathscr{A}_i}$ be positive normal linear functionals on $\mathscr{A}_i$, $i=1,2$. Then*

$$\mathrm{D}(\psi_1^{\mathscr{A}_1}\oplus\psi_1^{\mathscr{A}_2}\|\psi_2^{\mathscr{A}_1}\oplus\psi_2^{\mathscr{A}_2}) = \mathrm{D}(\psi_1^{\mathscr{A}_1}\|\psi_1^{\mathscr{A}_2}) + \mathrm{D}(\psi_2^{\mathscr{A}_1}\|\psi_2^{\mathscr{A}_2})$$

on $\mathscr{A}_1\oplus\mathscr{A}_2\subseteq B(H_1\oplus H_2)$.

Proof. See [Hia21, Proposition 2.3] with $f(t)=t\log(t)$. □

The following proposition is an analogue of Pinsker's inequality:

Proposition 12.1.12 (Theorem 5.5 of [OP04])**.** *If ψ_1 and ψ_2 are normal states on a von Neumann algebra $\mathscr{A}$, then*

$$\|\psi_1-\psi_2\|^2_{\mathscr{A}^*}\leq 2\ln(2)\,\mathrm{D}(\psi_1\|\psi_2).$$

Proposition 12.1.13. *Let ϕ,ψ and υ be normal positive linear functionals on a von Neumann algebra $\mathscr{A}$, such that $\mathrm{D}(\phi\|\psi)\leq\lambda_1$ and $\psi\leq 2^{\lambda_2}\upsilon$ for two scalars $\lambda_1,\lambda_2\geq 0$. Then $\mathrm{D}(\phi\|\upsilon)\leq\lambda_1+\phi(\mathbb{I})\lambda_2$.*

Proof. By Proposition 12.1.9, $\mathrm{D}(\phi\|\upsilon)\leq\mathrm{D}(\phi\|2^{-\lambda_2}\psi) = \mathrm{D}(\phi\|\psi)-\phi(\mathbb{I})\log 2^{-\lambda_2}\leq\lambda_1+\phi(\mathbb{I})\lambda_2$. □

12.1.3 Mutual information

Let $\mathscr{A}_1$ and $\mathscr{A}_2$ be two von Neumann algebras in standard form. If ψ_i is a normal state on $\mathscr{A}_i$, $i = 1,2$, then there is a unique state $\psi_1 \otimes \psi_2$ on $\mathscr{A}_1 \otimes \mathscr{A}_2$ such that $\psi_1 \otimes \psi_2(a \otimes b) = \psi_1(a)\psi_2(b)$ for all $a \in \mathscr{A}_1$, $b \in \mathscr{A}_2$ (indeed, if $|\psi_i\rangle$ is the vector in the positive cone corresponding to ψ_i, then $\psi_1 \otimes \psi_2$ is the state corresponding to $|\psi_1\rangle|\psi_2\rangle$). This leads to a definition of the mutual information between two algebras:

Definition 12.1.14 (Mutual information)**.** *Suppose $\psi^{\mathscr{A}_1\mathscr{A}_2}$ is a normal state on $\mathscr{A}_1 \otimes \mathscr{A}_2$. The mutual information between $\mathscr{A}_1$ and $\mathscr{A}_2$ for the state $\psi^{\mathscr{A}_1\mathscr{A}_2}$ is*

$$\mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_\psi := \mathrm{D}(\psi^{\mathscr{A}_1\mathscr{A}_2} \| \psi^{\mathscr{A}_1} \otimes \psi^{\mathscr{A}_2}).$$

If $\psi^{\mathscr{A}_1\mathscr{A}_2\mathscr{A}_3}$ is a normal state on $\mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$, then we use the convention that $I(\mathscr{A}_1 : \mathscr{A}_2)_\psi$ denotes the relative entropy between $\mathscr{A}_1$ and $\mathscr{A}_2$ for the state $\psi^{\mathscr{A}_1\mathscr{A}_2}$.

Example 12.1.15. Let $\mathscr{A}_1 = \mathscr{A}_2 = \mathbf{M}_n(\mathbb{C})$ and let $\psi^{\mathscr{A}_1\mathscr{A}_2}(a) = \mathrm{Tr}(\sigma^{\mathscr{A}_1\mathscr{A}_2} a)$ be a state on $\mathscr{A}_1 \otimes \mathscr{A}_2$, where $\sigma^{\mathscr{A}_1\mathscr{A}_2}$ is a density matrix in $\mathbf{M}_{n^2}(\mathbb{C})$. For all $a \in \mathscr{A}_1$,

$$\psi^{\mathscr{A}_1}(a) = \mathrm{Tr}(\sigma^{\mathscr{A}_1\mathscr{A}_2}(a \otimes \mathbb{I})) = \mathrm{Tr}(\sigma_1 a),$$

where $\sigma^{\mathscr{A}_1} := \mathrm{Tr}_{\mathscr{A}_2}(\sigma^{\mathscr{A}_1\mathscr{A}_2})$ is the partial trace. Similarly, $\psi^{\mathscr{A}_1}(b) = \mathrm{Tr}((\sigma^{\mathscr{A}_2} b)$ for all $b \in \mathscr{A}_2$, where $\sigma^{\mathscr{A}_2} := \mathrm{Tr}_{\mathscr{A}_1}(\sigma^{\mathscr{A}_1\mathscr{A}_2})$. If $a, b \in \mathbf{M}_n(\mathbb{C})$, then

$$\psi_1 \otimes \psi_2(a \otimes b) = \psi_1(a)\psi_2(b) = \mathrm{Tr}(\sigma^{\mathscr{A}_1} a)\mathrm{Tr}(\sigma^{\mathscr{A}_2} b) = \mathrm{Tr}((\sigma^{\mathscr{A}_1} \otimes \sigma^{\mathscr{A}_2})(a \otimes b)).$$

Hence $\mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_\psi$ is the relative entropy between the state with density matrix $\sigma^{\mathscr{A}_1\mathscr{A}_2}$, and the state with density matrix $\sigma^{\mathscr{A}_1} \otimes \sigma^{\mathscr{A}_2}$. By Example 12.1.8, this is the usual mutual information between Alice and Bob's registers with state $\sigma^{\mathscr{A}_1\mathscr{A}_2}$.

Many of the properties of mutual information in finite dimensions extend to mutual information between von Neumann algebras.

Proposition 12.1.16 (Corollary 5.20 of [OP04])**.** *Let $\mathscr{A}_1$ and $\mathscr{A}_2$ be two von Neumann algebras and let $\phi^{\mathscr{A}_1\mathscr{A}_2}$ and $\psi^{\mathscr{A}_1} \otimes \psi^{\mathscr{A}_2}$ be two normal states on $\mathscr{A}_1 \otimes \mathscr{A}_2$. Then*

$$\mathrm{D}(\phi^{\mathscr{A}_1\mathscr{A}_2} \| \psi^{\mathscr{A}_1} \otimes \psi^{\mathscr{A}_2}) = \mathrm{D}(\phi^{\mathscr{A}_1} \| \psi^{\mathscr{A}_1}) + \mathrm{D}(\phi^{\mathscr{A}_1\mathscr{A}_2} \| \phi^{\mathscr{A}_1} \otimes \psi^{\mathscr{A}_2}).$$

Proposition 12.1.17 (Monotonicity)**.** *If $\phi^{\mathscr{A}_1\mathscr{A}_2}$ and $\psi^{\mathscr{A}_1\mathscr{A}_2}$ are two positive normal linear functionals on the von Neumann algebra $\mathscr{A}_1 \otimes \mathscr{A}_2$, then $\mathrm{D}(\phi^{\mathscr{A}_1} \| \psi^{\mathscr{A}_1}) \leq \mathrm{D}(\phi^{\mathscr{A}_1\mathscr{A}_2} \| \psi^{\mathscr{A}_1\mathscr{A}_2})$. As a result, if ψ is a state on $\mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$ then $\mathrm{I}(\mathscr{A}_1 \otimes \mathscr{A}_2 : \mathscr{A}_3)_\psi \geq \mathrm{I}(\mathscr{A}_2 : \mathscr{A}_3)_\psi$.*

Proof. The map $\alpha : \mathscr{A}_1 \to \mathscr{A}_1 \otimes \mathscr{A}_2 : a \mapsto a \otimes \mathbb{I}_{\mathscr{A}_2}$ is a unital homomorphism, and hence satisfies Schwarz's inequality. Thus

$$\mathrm{D}(\phi^{\mathscr{A}_1} \| \psi^{\mathscr{A}_1}) = \mathrm{D}(\phi^{\mathscr{A}_1\mathscr{A}_2} \circ \alpha \| \psi^{\mathscr{A}_1\mathscr{A}_2} \circ \alpha) \leq \mathrm{D}(\phi^{\mathscr{A}_1\mathscr{A}_2} \| \psi^{\mathscr{A}_1\mathscr{A}_2})$$

by Proposition 12.1.10. □

Proposition 12.1.18 (Quantum Gibbs inequality)**.** *Let $\phi^{\mathscr{A}_1\mathscr{A}_2}$ be a normal state on the von Neumann algebra $\mathscr{A}_1 \otimes \mathscr{A}_2$, and $\psi^{\mathscr{A}_1}$, $\psi^{\mathscr{A}_2}$ be two normal state on $\mathscr{A}_1$ and $\mathscr{A}_2$ respectively. Then*

$$\mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_\phi \leq \mathrm{D}(\phi^{\mathscr{A}_1\mathscr{A}_2} \| \psi^{\mathscr{A}_1} \otimes \psi^{\mathscr{A}_2}).$$

Proof. Using Proposition 12.1.16 first on $\mathscr{A}_1$ and then on $\mathscr{A}_2$, we see that

$$\begin{aligned}
&\mathrm{D}(\phi^{\mathscr{A}_1\mathscr{A}_2} \| \psi^{\mathscr{A}_1} \otimes \psi^{\mathscr{A}_2}) - \mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_\phi \\
&\quad = \mathrm{D}(\phi^{\mathscr{A}_1} \| \psi^{\mathscr{A}_1}) + \mathrm{D}(\phi^{\mathscr{A}_1\mathscr{A}_2} \| \phi^{\mathscr{A}_1} \otimes \psi^{\mathscr{A}_2}) - \mathrm{D}(\phi^{\mathscr{A}_1\mathscr{A}_2} \| \phi^{\mathscr{A}_1} \otimes \phi^{\mathscr{A}_2}) \\
&\quad = \mathrm{D}(\phi^{\mathscr{A}_1} \| \psi^{\mathscr{A}_1}) + \mathrm{D}(\phi^{\mathscr{A}_2} \| \psi^{\mathscr{A}_2}),
\end{aligned}$$

which is non-negative by Proposition 12.1.9, part (1). □

In this thesis, we only use mutual information in the very restricted context of classical-quantum states discussed below. However, we have not seen Definition 12.1.14 in the literature previously, and it's interesting to discuss other possible definitions. For instance, the double dual $\mathscr{A}^{**}$ of a C^*-algebra $\mathscr{A}$ is a von Neumann algebra containing $\mathscr{A}$, such that any state ψ on $\mathscr{A}$ extends to a normal state $\widehat{\psi}$ on $\mathscr{A}^{**}$. Thus we can define the relative entropy between two states ϕ and ψ on a C^*-algebra as the relative entropy $\mathrm{D}(\widehat{\phi},\widehat{\psi})$ between the normal states $\widehat{\phi}$ and $\widehat{\psi}$ on $\mathscr{A}^{**}$. If $\mathscr{A}$ happens to be a von Neumann algebra and ϕ and ψ are normal states, then $\mathrm{D}(\widehat{\phi},\widehat{\psi}) = \mathrm{D}(\phi,\psi)$, so this does not lead to a new notion of relative entropy.

If ψ_1 and ψ_2 are two states on C^*-algebras $\mathscr{A}_1$ and $\mathscr{A}_2$ respectively, then there is a unique state $\psi_1 \otimes_{min} \psi_2$ on the min-tensor product $\mathscr{A}_1 \otimes_{min} \mathscr{A}_2$ such that $\psi_1 \otimes \psi_2(a \otimes b) = \psi_1(a)\psi_2(b)$ for all $a \in \mathscr{A}_1$, $b \in \mathscr{A}_2$, and this pulls back to a unique state $\psi_1 \otimes_{max} \psi_2$ on $\mathscr{A}_1 \otimes_{max} \mathscr{A}_2$ with the same property (the definition for the min/max tensor product are the standard definition used for C^* algebra theory, and can be found in, e.g. [Gol21, Section 3.8]). Hence if ψ is a state on the max tensor product $\mathscr{A}_1 \otimes_{max} \mathscr{A}_2$, then we can define the mutual information between $\mathscr{A}_1$ and $\mathscr{A}_2$ for the state ψ to be

$$\mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_{\psi}^{max} := \mathrm{D}(\psi, \psi^{\mathscr{A}_1} \otimes_{max} \psi^{\mathscr{A}_2}),$$

where $\psi^{\mathscr{A}_i}$ is the restriction of ψ to $\mathscr{A}_i$ inside of $\mathscr{A}_1 \otimes_{max} \mathscr{A}_2$. If ψ is a state on $\mathscr{A}_1 \otimes_{min} \mathscr{A}_2$, then we can define $\mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_{\psi}^{min} := \mathrm{D}(\psi, \psi^{\mathscr{A}_1} \otimes_{min} \psi^{\mathscr{A}_2})$ similarly. Any state ψ on $\mathscr{A}_1 \otimes_{min} \mathscr{A}_2$ pulls back to a state $\widetilde{\psi}$ on $\mathscr{A}_1 \otimes_{max} \mathscr{A}_2$, so there are seemingly two different choices for the mutual information between $\mathscr{A}_1$ and $\mathscr{A}_2$ in this case, $\mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_{\psi}^{min}$ and $\mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_{\widetilde{\psi}}^{max}$. However, there is a surjective homomorphism from $\mathscr{A}_1 \otimes_{max} \mathscr{A}_2$ to $\mathscr{A}_1 \otimes_{min} \mathscr{A}_2$, so the following lemma shows that $\mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_{\psi}^{min} = \mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_{\widetilde{\psi}}^{max}$.

Lemma 12.1.19. *If $\alpha : \mathscr{A} \to \mathscr{B}$ is a surjective $*$-homomorphism between C^*-algebras, and ϕ and ψ are states on $\mathscr{B}$, then $D(\phi \circ \alpha, \psi \circ \alpha) = \mathrm{D}(\phi,\psi)$.*

The proof of Lemma 12.1.19 follows from [Hia21, Theorem 6.19]; since we do not make further use of this lemma, we leave the complete proof as an exercise for the reader. Similarly, if ψ is a normal state on the tensor product $\mathscr{A}_1 \otimes \mathscr{A}_2$ of two von Neumann algebras, then $\mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_{\psi} = \mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_{\psi}^{min} = \mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_{\widetilde{\psi}}^{max}$.

12.1.4 Classical-quantum states

For a discrete finite set $\mathscr{X}$, let $\mathbb{C}^{\mathscr{X}}$ denote the von Neumann algebra of functions from $\mathscr{X}$ to $\mathbb{C}$. To match the standard notation from quantum information, let $\langle x|$ denote the indicator function for $x \in X$. These functions span $\mathbb{C}^{\mathscr{X}}$, and give an isomorphism between $\mathbb{C}^{\mathscr{X}}$ and the algebra of $|\mathscr{X}| \times |\mathscr{X}|$ diagonal matrices. If $\mathscr{A}$ is another von Neumann algebra, then $\mathbb{C}^{\mathscr{X}} \otimes \mathscr{A} = \bigoplus_{x \in X} \langle x| \otimes \mathscr{A}$ is the von Neumann algebra of $|\mathscr{X}| \times |\mathscr{X}|$ diagonal matrices with coefficients from $\mathscr{A}$, and every normal state on $\mathbb{C}^{\mathscr{X}} \otimes \mathscr{A}$ is of the form

$$\phi^{\mathscr{X}\mathscr{A}} = \sum_{x \in \mathscr{X}} \mathsf{P}(x) \langle x| \otimes \phi_x$$

for some collection of normal states $\{\phi_x\}_{x \in \mathscr{X}}$ on $\mathscr{A}$ and probability measure $\mathsf{P}(x)$ on $\mathscr{X}$. Hence such states are called *classical-quantum states on $\mathscr{A}$ with classical part $\mathscr{X}$*. When dealing with multiple classical subsystems, we denote ϕ_x by $\phi^{\mathscr{X}\mathscr{A}}_{\mathscr{X}=x}$. Also, note that if $\phi^{\mathscr{X}\mathscr{A}}$ is a classical-quantum state, then $\phi^{\mathscr{X}}$ is the classical distribution P on $\mathscr{X}$. We use the following example to connect our definition with the standard definition for classical-quantum state used in quantum information.

Example 12.1.20. Let $\mathscr{X}$ be a discrete finite set, $\mathsf{P}(x)$ be a probability distribution over $\mathscr{X}$ and let $(\mathscr{A},\tau)$ be a tracial von Neumann algebra. Define the collection of normal state $\{\phi_x\}_{x \in \mathscr{X}}$

acting on $\mathscr{A}$ as $\phi_x(A) = \tau(\sigma_x A)$ for some positive element $\sigma_x \in \mathscr{A}^+$ with $\tau(\sigma_x) = 1$. We define the classical-quantum state $\phi^{\mathscr{X}\mathscr{A}} = \sum_{x\in\mathscr{X}} \mathsf{P}(x)\,\langle x| \otimes \phi_x$, and we see that for all $A \in \mathscr{M}_{|X|}(\mathbb{C}) \otimes A$

$$\phi^{\mathscr{X}\mathscr{A}}(A) = \mathrm{Tr} \otimes \tau \left(A \cdot \left(\sum_{x\in\mathscr{X}} \mathsf{P}(x)|x\rangle\langle x| \otimes \sigma_x \right) \right)$$

where Tr in the above equation is defined over $\mathscr{M}_{|X|}(\mathbb{C})$. In this case, one can intuitively think of $\sum_{x\in\mathscr{X}} \mathsf{P}(x)|x\rangle\langle x| \otimes \sigma_x$ as the "density matrix" for the state $\phi^{\mathscr{X}\mathscr{A}}$, and we see that this is consistent with the standard definition for classical-quantum state in quantum information literatures (e.g. [Wil13, Definition 4.3.5]).

We now prove some properties of classical-quantum states used in [BVY21, Section 5.1].

Proposition 12.1.21 (Chain rule for relative entropy)**.** *Let $\phi^{\mathscr{X}\mathscr{A}} = \sum_x \mathsf{P}(x)\,\langle x| \otimes \phi_x^{\mathscr{A}}$ and $\psi^{\mathscr{X}\mathscr{A}} = \sum_x \mathsf{Q}(x)\,\langle x| \otimes \psi^{\mathscr{A}}$ be two classical-quantum states on $\mathscr{A}$ with classical part $\mathscr{X}$. Then*

$$\mathrm{D}(\phi^{\mathscr{X}\mathscr{A}} \| \psi^{\mathscr{X}\mathscr{A}}) = \mathrm{D}(\mathsf{P}\|\mathsf{Q}) + \mathop{\mathbb{E}}_{x\sim\mathsf{P}} \mathrm{D}(\phi_x^{\mathscr{A}} \| \psi_x^{\mathscr{A}}),$$

where $\mathrm{D}(\mathsf{P}\|\mathsf{Q}) = \sum_x \mathsf{P}(x) \log \frac{\mathsf{P}(x)}{\mathsf{Q}(x)}$ denotes the relative entropy between two classical distributions P and Q. As a result, $\mathrm{D}(\phi\|\psi) \geq \mathbb{E}_{x\sim\mathsf{P}} \mathrm{D}(\phi_x^{\mathscr{A}} \| \psi_x^{\mathscr{A}})$.

Proof. Using Proposition 12.1.11 and part (3) of Proposition 12.1.9,

$$\begin{aligned}
\mathrm{D}(\phi^{\mathscr{X}\mathscr{A}} \| \psi^{\mathscr{X}\mathscr{A}}) &= \sum_x \mathrm{D}(\mathsf{P}(x)\phi_x^{\mathscr{A}} \| \mathsf{Q}(x)\psi_x^{\mathscr{A}}) \\
&= \sum_x \mathsf{P}(x)\mathrm{D}(\phi_x^{\mathscr{A}} \| \psi_x^{\mathscr{A}}) + \mathsf{P}(x)\log\left(\frac{\mathsf{P}(x)}{\mathsf{Q}(x)}\right) \\
&= \mathop{\mathbb{E}}_{x\sim\mathsf{P}} \mathrm{D}(\phi_x^{\mathscr{A}} \| \psi_x^{\mathscr{A}}) + \mathrm{D}(\mathsf{P}\|\mathsf{Q}).
\end{aligned}$$

Since relative entropy between classical distributions is non-negative, $\mathrm{D}(\phi\|\psi) \geq \mathbb{E}_{x\sim\mathsf{P}} \mathrm{D}(\phi_x^{\mathscr{A}} \| \psi_x^{\mathscr{A}})$. □

Proposition 12.1.22 (Conditional mutual information)**.** *Let $\phi^{\mathscr{X}\mathscr{A}_1\mathscr{A}_2} = \sum_x P(x)\,|x\rangle \otimes \phi_x$ be a classical-quantum state on $\mathscr{A}_1 \otimes \mathscr{A}_2$ with classical part $\mathscr{X}$. Then*

$$\mathrm{I}(\mathscr{X}\mathscr{A}_1 : \mathscr{A}_2)_\phi - \mathrm{I}(\mathscr{X} : \mathscr{A}_2)_\phi = \mathop{\mathbb{E}}_{x\sim\mathsf{P}} \mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_{\phi_x}.$$

Proof. The restriction of $\phi^{\mathscr{X}\mathscr{A}}$ to $\mathbb{C}^{\mathscr{X}}$ is $\phi^{\mathscr{X}} = \sum_x \mathsf{P}(x)\,\langle x|$. Since $\mathrm{D}(P,P) = 0$, Proposition 12.1.21 implies that

$$\mathrm{I}(\mathscr{X} : \mathscr{A}_2)_\phi = \mathrm{D}\left(\sum_x \mathsf{P}(x)\,|x\rangle \otimes \phi_x^{\mathscr{A}_2} \| \sum_x \mathsf{P}(x)\,|x\rangle \otimes \sum_y \mathsf{P}(y)\phi_y^{\mathscr{A}_2} \right) = \mathop{\mathbb{E}}_{x\sim\mathsf{P}} \mathrm{D}(\phi_x^{\mathscr{A}_2} \| \phi^{\mathscr{A}_2}),$$

where $\phi^{\mathscr{A}_2} = \sum_y P(y)\phi_y^{\mathscr{A}_2}$. Similarly,

$$\mathrm{I}(\mathscr{X}\mathscr{A}_1 : \mathscr{A}_2)_\phi = \mathrm{D}(\phi^{X\mathscr{A}_1\mathscr{A}_2} \| \phi^{X\mathscr{A}_1} \otimes \phi^{\mathscr{A}_2}) = \mathop{\mathbb{E}}_{x\sim\mathsf{P}} \mathrm{D}(\phi_x \| \phi_x^{\mathscr{A}_1} \otimes \phi^{\mathscr{A}_2}).$$

Applying Proposition 12.1.16 to $\mathscr{A}_2$, we see that

$$\begin{aligned}
\mathrm{I}(\mathscr{X}\mathscr{A}_1 : \mathscr{A}_2)_\phi - \mathrm{I}(\mathscr{X} : \mathscr{A}_2)_\phi &= \mathop{\mathbb{E}}_{x\sim\mathsf{P}} \mathrm{D}(\phi_x^{\mathscr{A}_1\mathscr{A}_2} \| \phi_x^{\mathscr{A}_1} \otimes \phi^{\mathscr{A}_2}) - \mathrm{D}(\phi_x^{\mathscr{A}_2} \| \phi^{\mathscr{A}_2}) \\
&= \mathop{\mathbb{E}}_{x\sim\mathsf{P}} \mathrm{D}(\phi_x^{\mathscr{A}_1\mathscr{A}_2} \| \phi_x^{\mathscr{A}_1} \otimes \phi_x^{\mathscr{A}_2}) = \mathop{\mathbb{E}}_{x\sim\mathsf{P}} \mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_{\phi_x}.
\end{aligned}$$

□

We can now prove a von Neumann algebraic version of *quantum Raz's Lemma*, which is a central tool in [BVY21]. This is a quantum analogue of Raz's lemma, which is a key part of many proofs of the classical parallel repetition theorem [Raz95; Hol09; BRR+09].

Lemma 12.1.23 (Quantum Raz's Lemma)**.** *Let $\phi^{\mathscr{X}\mathscr{A}} = \phi^{\mathscr{X}_1\mathscr{X}_2\ldots\mathscr{X}_n\mathscr{A}}$ and $\psi^{\mathscr{X}\mathscr{A}} = \psi^{\mathscr{X}_1}\otimes\psi^{\mathscr{X}_2}\otimes\ldots\otimes\psi^{\mathscr{X}_n}\otimes\psi^{\mathscr{A}}$ be two classical-quantum states with classical component $\mathscr{X} = \mathscr{X}_1\times\mathscr{X}_2\times\ldots\mathscr{X}_n$. Then*

$$\sum_{i=1}^{n} \mathrm{I}(\mathscr{X}_i : \mathscr{A})_\phi \leq \mathrm{D}(\phi^{\mathscr{X}\mathscr{A}} \| \psi^{\mathscr{X}\mathscr{A}}) . \tag{12.11}$$

Proof. Let $\mathscr{X}_{\leq i} := \mathscr{X}_1\mathscr{X}_2\cdots\mathscr{X}_i$ and $\mathscr{X}_{\geq i} := \mathscr{X}_i\mathscr{X}_{i+1}\cdots\mathscr{X}_n$. For each $2 \leq i \leq n$, Proposition 12.1.22 implies that

$$\begin{aligned}\mathrm{I}(\mathscr{X}_{\leq i-1}\mathscr{A} : \mathscr{X}_i)_{\phi^{\mathscr{X}\mathscr{A}}} - \mathrm{I}(\mathscr{X}_{\leq i-1} : \mathscr{X}_i)_{\phi^{\mathscr{X}\mathscr{A}}} &= \mathop{\mathbb{E}}_{x_{<i}\sim\phi^{\mathscr{X}_{\leq i-1}}} \mathrm{I}(\mathscr{X}_i : \mathscr{A})_{\phi^{\mathscr{X}_{\leq i}\mathscr{A}}_{x_{<i}}} \\ &= \mathrm{I}(\mathscr{X}_{\leq i} : \mathscr{A})_{\phi^{\mathscr{X}\mathscr{A}}} - \mathrm{I}(\mathscr{X}_{\leq i-1} : \mathscr{A})_{\phi^{\mathscr{X}\mathscr{A}}} . \end{aligned} \tag{12.12}$$

Repeatedly applying Proposition 12.1.16, we get that

$$\begin{aligned}\mathrm{D}(\phi^{\mathscr{X}\mathscr{A}} \| \psi^{\mathscr{X}\mathscr{A}}) &= \mathrm{D}(\phi^{\mathscr{X}_1} | \psi^{\mathscr{X}_1}) + \mathrm{D}(\phi^{\mathscr{X}\mathscr{A}} | \phi^{\mathscr{X}_1}\otimes\psi^{\mathscr{X}_{\geq 2}\mathscr{A}}) \\ &= \mathrm{D}(\phi^{\mathscr{X}_1} | \psi^{\mathscr{X}_1}) + \mathrm{D}(\phi^{\mathscr{X}_1\mathscr{X}_2} | \phi^{\mathscr{X}_1}\otimes\psi^{\mathscr{X}_2}) + \mathrm{D}(\phi^{\mathscr{X}\mathscr{A}} | \phi^{\mathscr{X}_{\leq 2}}\otimes\psi^{\mathscr{X}_{\geq 2}\mathscr{A}}) \\ &= \sum_{i=1}^{n} \mathrm{D}(\phi^{\mathscr{X}_{\leq i}} | \phi^{\mathscr{X}_{\leq i-1}}\otimes\psi^{\mathscr{X}_i}) + \mathrm{D}(\phi^{\mathscr{X}\mathscr{A}} | \phi^{\mathscr{X}}\otimes\psi^{\mathscr{A}}).\end{aligned}$$

Hence by the quantum Gibb's inequality in Proposition 12.1.18,

$$\mathrm{D}(\phi^{\mathscr{X}\mathscr{A}} \| \psi^{\mathscr{X}\mathscr{A}}) \geq \sum_{i=2}^{n} \mathrm{I}(\mathscr{X}_{\leq i-1} : \mathscr{X}_i)_{\phi^{\mathscr{X}\mathscr{A}}} + \mathrm{I}(\mathscr{X} : \mathscr{A})_{\phi^{\mathscr{X}\mathscr{A}}}.$$

Solving for $\mathrm{I}(\mathscr{X}_{\leq i-1} : \mathscr{X}_i)_{\phi^{\mathscr{X}\mathscr{A}}}$ in Equation (12.12), we get the telescoping sum

$$\begin{aligned}\sum_{i=2}^{n} \mathrm{I}(\mathscr{X}_{\leq i-1} : \mathscr{X}_i)_{\phi^{\mathscr{X}\mathscr{A}}} &= \sum_{i=2}^{n} \left(\mathrm{I}(\mathscr{X}_{\leq i-1}\mathscr{A} : \mathscr{X}_i)_{\phi^{\mathscr{X}\mathscr{A}}} - \mathrm{I}(\mathscr{X}_{\leq i} : \mathscr{A})_{\phi^{\mathscr{X}\mathscr{A}}} + \mathrm{I}(\mathscr{X}_{\leq i-1} : \mathscr{A})_{\phi^{\mathscr{X}\mathscr{A}}} \right) \\ &= \sum_{i=2}^{n} \mathrm{I}(\mathscr{X}_{\leq i-1}\mathscr{A} : \mathscr{X}_i)_{\phi^{\mathscr{X}\mathscr{A}}} + \mathrm{I}(\mathscr{X}_1 : \mathscr{A})_{\phi^{\mathscr{X}\mathscr{A}}} - \mathrm{I}(\mathscr{X} : \mathscr{A})_{\phi^{\mathscr{X}\mathscr{A}}}.\end{aligned}$$

By Proposition 12.1.17, $\mathrm{I}(\mathscr{X}_{\leq i-1}\mathscr{A} : \mathscr{X}_i)_{\phi^{\mathscr{X}\mathscr{A}}} \geq \mathrm{I}(\mathscr{A} : \mathscr{X}_i)$, so

$$\begin{aligned}\mathrm{D}(\phi^{\mathscr{X}\mathscr{A}} \| \psi^{\mathscr{X}\mathscr{A}}) \geq \sum_{i=2}^{n} \mathrm{I}(\mathscr{X}_{\leq i-1} : \mathscr{X}_i)_{\phi^{\mathscr{X}\mathscr{A}}} + \mathrm{I}(\mathscr{X} : \mathscr{A})_{\phi^{\mathscr{X}\mathscr{A}}} &= \sum_{i=2}^{n} \mathrm{I}(\mathscr{X}_{\leq i-1}\mathscr{A} : \mathscr{X}_i)_{\phi^{\mathscr{X}\mathscr{A}}} + \mathrm{I}(\mathscr{X}_1 : \mathscr{A}) \\ &\geq \sum_{i=1}^{n} \mathrm{I}(\mathscr{X}_i : \mathscr{A})_{\phi^{\mathscr{X}\mathscr{A}}}.\end{aligned}$$

□

12.2 The anchored parallel repetition theorem

We introduce the anchored parallel repetition theorem in this section. Given a non-local game $\mathscr{G} = (\mathscr{X}, \mathscr{A}, \mu, D)$ and $r \in \mathbb{N}$, we define the r-fold parallel repetition to be the game $\mathscr{G}^{\otimes r} = (\mathscr{X}^r, \mathscr{A}^r, \mu^r, D^r)$ as the game with the following question distribution and validation function:

- $\mu^n((x_0,\cdots,x_{r-1}),(y_0,\cdots y_{r-1})) = \prod_{i=0}^{r}\mu(x_i,y_i)$.
- $\mathrm{D}^r\left((x_0,\cdots,x_{r-1}),(y_0,\cdots y_{r-1}),(a_0,\cdots,a_{r-1}),(b_0,\cdots,b_{r-1})\right) = \prod_{i=0}^{r} D(x_i,.y_i,a_i,b_i)$

Intuitively, the above transformation corresponds to the verifier sampling r pairs of questions $(x_i, y_i), i \in [r]$ from the distribution μ, and sends them to the provers. The provers must respond with answers $(a_i, b_i) \in \mathscr{A}^2$ for $i \in [r]$, and the provers win iff $D(x_i, y_i, a_i, b_i) = 1$ for all $i \in [r]$. If a game $\mathscr{G}$ is synchronous, then its r-fold parallel repetition $\mathscr{G}^{\otimes r}$ is also synchronous. For clarity of notation, we use $(\vec{x}, \vec{y})$ (resp. $(\vec{a}, \vec{b})$) to emphasize that the question/answer pairs (resp. $(\vec{a}, \vec{b})$) come from the parallel-repeated game.

For a non-local game $\mathscr{G}$, we define the anchoring transformation for a game as follows:

Definition 12.2.1 (Anchoring transformation)**.** *Given a game $\mathscr{G} = (\mathscr{X}, \mathscr{A}, \mu, D)$, we define anchored transformation $\mathscr{G}_\perp = (\mathscr{X}_\perp = \mathscr{X} \cup \{\perp\}, \mathscr{A}_\perp = \mathscr{A} \cup \{\perp\}, \mu^\perp, D^\perp)$ for the game $\mathscr{G}$, where the question distribution $\mu_\perp$ is defined as*

$$\mu^\perp(x,y) = \begin{cases} \frac{1}{4}\mu(x,y) & \text{if } (x,y) \in \mathscr{X}^2 \\ \frac{1}{4}\mu_x(x) & \text{if } y = \perp \\ \frac{1}{4}\mu_y(y) & \text{if } x = \perp \\ \frac{1}{4} & \text{if } x = \perp, y = \perp \end{cases},$$

where recall μ_x and μ_y are the marginal distribution for μ on both provers' sides respectively. The evaluation $D^\perp$ is defined as

$$D^\perp(x,y,a,b) = \begin{cases} 8D(x,y,a,b) & \text{if } (x,y) \in \mathscr{X}^2 \\ 1 & \text{if } (x = a = \perp, y \in \mathscr{X}) \vee (y = b = \perp, x \in \mathscr{X}) \vee (x = y = a = b = \perp) \\ 0 & \text{otherwise} \end{cases}.$$

We remark that the above definition is a modification of the anchoring transformation from [BVY21] with α taken to be $\frac{1}{2}$, and instead of giving the provers a free win, we now expect that the provers both need to output an anchoring symbol in order to win the game. This ensures that every synchronous game remains synchronous after the transformation. As shown in [JNV+22a], this does not change the strong parallel repetition guarantee by [BVY21]. We see that the anchoring transformation only changes the value of a game by a constant factor through the following lemma.

Lemma 12.2.2 (Preservation of value of the anchoring transformation)**.** *Let $t \in \{*, co\}$, and let $\mathscr{G}$ be a non-local game. If $\omega^t(\mathscr{G}) \geq 1 - \epsilon$ for some $\epsilon \in [0,1]$, then $\omega^t(\mathscr{G}_\perp) \geq 1 - \frac{1-\epsilon}{4}$.*

The above lemma follows trivially by formulating a strategy that answers the anchoring symbol $\perp$ when $\perp$ is given as the question. It is not hard to see that, for $t \in \{*, co\}$, if there exists a perfect oracularizable synchronous strategy for $\mathscr{G}$ in model t, then there exists a perfect oracularizable synchronous strategy for $\mathscr{G}_\perp$ in model t by adding the measurement operator $A_0^\perp = \mathbb{1}_{\mathscr{A}}$ to the existing perfect strategy for $\mathscr{G}$. Furthermore, if $\omega^t(\mathscr{G}) = \epsilon$, then $\omega^t(\mathscr{G}_\perp) = \frac{3}{4} + \frac{\epsilon}{4}$.

We are now ready to state the parallel repetition theorem below.

Theorem 12.2.3 (Anchored Parallel repetition theorem)**.** *There exists a universal constant c^{para} such that, for any model $t \in \{*, co\}$, non-local games $\mathscr{G} = (\mathscr{X}, \mathscr{A}, \mu, D)$ with $\omega^t \leq 1 - \epsilon$, and $r \in \mathbb{N}$. Then*

$$\omega^t(\mathscr{G}_\perp^{\otimes r}) \leq \frac{16}{\epsilon} \cdot \exp\left(\frac{-c^{para}\epsilon^{17} r}{\log(|\mathscr{A}|+1)}\right). \tag{12.13}$$

We remark that the theorem above for the tensor product model (i.e. $t = *$) is already proven in [BVY21]. We give a proof of the above theorem for the commuting operator model in the next section. Before we begin, we first give a overview of the proof below.

12.2.1 Overview of the proof of the quantum parallel repetition theorem

We give a brief overview of the proof of the anchored parallel repetition theorem Theorem 12.2.3 in this subsection. This proof follows a similar structure as [BVY21], which itself follows a similar structure as the proof for parallel repetition theorem for *classical* values for non-local games (see, e.g. [Raz95]). These approaches argue that if a "too good to be true" strategy exists for the parallel repeated game, then, using information theoretical tools, a strategy which violates the optimal success probability for the original game can be constructed.

To be more precise, suppose for a r-fold parallel repeated game $\mathcal{G}^{\otimes r}$; there exists some strategy $\mathcal{S}^{\otimes r} = \{\mathcal{H}, |\psi\rangle, \{A^{\vec{a}}_{\vec{x}}\}, \{B^{\vec{b}}_{\vec{y}}\}\}$ which has a success rate higher than the value indicated in Theorem 12.2.3. Then, the strategy $\mathcal{S}$ must be performed in some correlated manner (i.e. the answer for the question pair $(\vec{a}_i, \vec{b}_i)$ must rely on some other question pairs which are not $(\vec{x}_i, \vec{y}_i)$). More precisely, since the strategies are correlated between the different folds of the parallel repeated game, there must exist some *critical subset* $C \subseteq [r]$ such that conditioning on the provers winning on the subset C using this strategy, the provers can win the overall parallel repeated game with high probability. If the provers were to somehow obtain an entangled state which mimics a "post-measurement state" in which they had already won on those C coordinates, then this gives them an advantage with the remaining $[r] \setminus C$ coordinates. Notably, this also gives a comparable advantage for a single instance of the game $\mathcal{G}$ by running the strategy above and embedding the coordinate $(\vec{x}_j, \vec{y}_j)$ onto one of the $\{(\vec{x}_i, \vec{y}_i)\}_{i \in [r] \setminus C}$.

Unfortunately, it is not clear how to sample such a post-measurement state locally. Since conditioning on the provers winning on coordinate C might change the input distribution on coordinate j (as an example, this could occur when the answer given to coordinate j is entirely dependent on the question on some coordinates on C), and in some cases, winning on coordinate C might depend on one of the provers getting a certain input on the jth coordinate ! This means that creating such a "post-measurement state" would often require the full question pair $\{(\vec{x}_i, \vec{y}_i)\}_{i \in [r]}$ (which includes coordinates j). This is one of the main challenges for showing the parallel repetition theorem using this approach.

One notable example in which the above approach works is the case where the classical input distribution to the non-local game is a product distribution between both provers [JMS20]; in this case, it can be shown that this "post-measurement state" is, on average, relatively uncorrelated to the provers' question pair on coordinates outside of C, and hence this state can be created locally by the provers. The anchored transformation used in [BVY21] and this thesis intuitively destroys the correlations between the provers, and by using certain conditioning (on *dependency breaking variables*), the prover's distribution can be made uncorrelated, and hence a similar argument can be made.

12.3 Proof of the anchored parallel repetition theorem

Having presented the von Neumann algebra framework for nonlocal games and some quantum information theory tools within it, we prove Theorem 12.2.3 in this subsection. Before we begin, we first give some additional background about classical probability which we need for the proof.

12.3.1 Preliminaries on classical probability theory

In this thesis, we use P, Q, S and R to denote probability distributions. Given a probability distribution P on a discrete finite set $\mathcal{X}$ and a random variable f on X, we let $\mathbb{E}_{x \sim P} f(x)$ denote the expected value of f with distribution P. We use P_X to denote the distribution of random variable X and $\mathsf{P}_X(x)$ to denote the probability that $X = x$ for some value x. For multiple random variables, e.g. X, Y, Z, $\mathsf{P}_{XYZ}(x, y, z)$ denotes their joint distribution. All random variables

are assumed to operate on the same probability space, which is usually implicit and clear from context.

We use $\mathsf{P}_{Y|X=x}(y)$ to denote the conditional distribution $\mathsf{P}_{YX}(y,x)/\mathsf{P}_X(x)$, which is defined when $\mathsf{P}_X(x) > 0$. We use the shorthand $\mathsf{P}_{X|y,z}$ to denote the distribution $\mathsf{P}_{X|Y=y,Z=z}$. For example, we may write $\mathsf{P}_{V|\omega_{-i},\vec{x}_i,\vec{y}_i}$ to denote $\mathsf{P}_{V|\Omega_{-i}=\omega_{-i},\vec{x}_i=\vec{x}_i,\vec{y}_i=\vec{y}_i}$. For an event W we let $\mathsf{P}_{XY|W}$ denote the distribution conditioned on W. We use the notation $\mathbb{E}_X f(x)$ and $\mathbb{E}_{\mathsf{P}_X} f(x)$ to denote the expectation $\sum_x \mathsf{P}_X(x) f(x)$. Let P_{XY} be a joint distribution on $\mathscr{X} \times \mathscr{Y}$ and let W denote an event. Then we define the distribution $\mathsf{P}_{X|W}\mathsf{P}_{Y|X}$ over $\mathscr{X} \times \mathscr{Y}$ as

$$(\mathsf{P}_{X|W}\mathsf{P}_{Y|X})(x,y) = \mathsf{P}_{X|W}(x) \cdot \mathsf{P}_{Y|X=x}(y)\,.$$

For distributions P_X and P_Y over the same set $\mathscr{X}$ we use $\|\mathsf{P}_{X_0} - \mathsf{P}_{X_1}\|$ to denote their total variation distance,

$$\|\mathsf{P}_{X_0} - \mathsf{P}_{X_1}\| = \frac{1}{2}\sum_{x\in\mathscr{X}} |\mathsf{P}_{X_0}(x) - \mathsf{P}_{X_1}(x)|\,. \tag{12.14}$$

We recall the following lemmas from [BVY21, Section 3.2], and we refer to Section 3.2 of the aforementioned paper for the proof.

Lemma 12.3.1. *Let Q_F and S_F be two probability distributions for random variable F and let $\mathsf{R}_{G|F}$ be a conditional probability distribution for random variable $\mathscr{G}_\perp$, conditioned on F. Then*

$$\|\mathsf{Q}_F\mathsf{R}_{G|F} - \mathsf{S}_F\mathsf{R}_{G|F}\| = \|\mathsf{Q}_F - \mathsf{S}_F\|\,.$$

Similarly, for two conditional probability distributions $\mathsf{Q}_{G|F}, \mathsf{S}_{G|F}$ and a distribution R_F,

$$\|\mathsf{R}_F\mathsf{Q}_{G|F} - \mathsf{R}_F\mathsf{S}_{G|F}\| = \mathop{\mathbb{E}}_F \|\mathsf{Q}_{G|F=f} - \mathsf{S}_{G|F=f}\|\,,$$

where $\mathbb{E}_F$ denotes the expectation over sampling f from R_F.

Lemma 12.3.2 (Data processing inequality)**.** *Let Q_{FG} and S_{FG} denote two probability distributions for random variables F, G. Then*

$$\|\mathsf{Q}_F - \mathsf{S}_F\| \le \|\mathsf{Q}_{FG} - \mathsf{S}_{FG}\|\,.$$

12.3.2 Existences of a critical subset C and dependency breaking variable

Recall from the beginning of this chapter that $\mathscr{G} = (\mathscr{X}, \mathscr{A}, \mu, D)$ denotes an anchoring game which arises from Definition 12.2.1. Let $\mathscr{G}^{\otimes r}$ denote the r-fold parallel repetition game and let $\mathscr{S}^{\otimes r} = \{\mathscr{L}^2(\mathscr{A}, \tau), |\psi\rangle = \sigma|\tau\rangle, \{A^{\vec{a}}_{\vec{x}}\}, \{B^{\vec{b}}_{\vec{y}}\}\}$ be a tracially embeddable strategy which violates the bound given in Theorem 12.2.3. We start with the following proposition which introduce the notion of the critical subset, since the proof is the same as in [BVY21], we refer to the aforementioned thesis for the proof.

Proposition 12.3.3 (Proposition 6.5 of [BVY21])**.** *Let $t \in \{*, co\}$ and let $\mathscr{G}$ be the game indicated by Theorem 12.2.3 with $\omega^t(\mathscr{G}) < 1 - \epsilon$. Let W denote the indicator for winning all r coordinates for the parallel repeated game $\mathscr{G}^{\otimes r}$. Suppose that $n \ge \frac{16}{\varepsilon}\log\frac{4}{\varepsilon\cdot\mathsf{P}(W)}$. Then there exists a set $C \subseteq [r]$ of size at most $t = \frac{8}{\varepsilon}\log\frac{4}{\varepsilon\cdot\mathsf{P}(W)}$ such that*

$$\mathop{\mathbb{E}}_I \mathsf{P}(W_i|W_C) \ge 1 - \varepsilon/2\,,$$

where $\mathbb{E}_I$ denotes the expectation over a uniformly random i chosen from $[r] \setminus C$ and $\mathsf{P}(W_i|W_C)$ denotes the probability, using the strategy $\mathscr{S}^{\otimes r}$, of winning the i-th instance of $\mathscr{G}$ conditioned on winning all instances indexed by C.

We fix a subset C promised by Proposition 12.3.3, which we call the *critical subset* for the parallel repeated game $\mathcal{G}^{\otimes r}$. We further assume without loss of generality that $C = \{r-|C|, \dots, n-1\}$.

For the remainder of this section, we reintroduce *dependency-breaking variables* from [BVY21]. These are crucial tools for controlling the correlations between the input distributions between the provers. Since most of the propositions in this section are classical in nature and are proven in [BVY21], we refer to Section 4 of the original paper for the proofs.

Recall from Section 3.4, the distributions μ_X and μ_Y denote the marginals of the game distribution μ for the anchored game $\mathcal{G}$ on the first and second coordinates respectively. We first define a "single copy" distribution $\hat{P}$ as the law of random variables $(M_{\text{player}}, M_{\text{value}}, X, Y)$, where each random variable may depend on previously defined ones. Following the same set up as [BVY21, Section 4.1], we fix a "noise" parameter

$$\eta_{\text{Anchor}} = \frac{1}{4}, \tag{12.15}$$

We define the random variable M_{player} to be a uniform distribution over the finite set $\{A, B\}$, and we define M_{value} to have the following distribution over $\mathcal{X}$: for all $(x, y) \in \mathcal{X}^2$:

$$\mathsf{P}_{M_{\text{value}}|M_{\text{player}}=A}(x) = \begin{cases} \frac{\mu_x(x)}{1-\eta_{\text{Anchor}}} & \text{if } x \neq \perp \\ \frac{\frac{1}{2}-\eta_{\text{Anchor}}}{1-\eta_{\text{Anchor}}} & \text{if } x = \perp \end{cases} \quad \text{and} \quad \mathsf{P}_{M_{\text{value}}|M_{\text{player}}=B}(y) = \begin{cases} \frac{\mu_y(y)}{1-\eta_{\text{Anchor}}} & \text{if } y \neq \perp \\ \frac{\frac{1}{2}-\eta_{\text{Anchor}}}{1-\eta_{\text{Anchor}}} & \text{if } y = \perp \end{cases}.$$

In other words, conditioned on $M_{\text{player}} = A$ (resp. $M_{\text{player}} = B$), the variable M_{value} takes on a value in $\mathcal{X}$ from a rescaled version of the distribution μ_X (resp. μ_Y) where less weight is given to the dummy question $\perp$. Finally, define the random variables (X, Y) as follows.

- If $M_{\text{player}} = A$ then X is chosen to be an "η_{Anchor}-noisy" copy of M_{value}. Precisely, $X = M_{\text{value}}$ with probability $1 - \eta_{\text{Anchor}}$ and $X = \perp$ with probability η_{Anchor}. Define Y to equal y with probability $\mu_{Y|X}(y|m)$, where m is the value of M_{value}.
- If $M_{\text{player}} = B$ then Y is chosen to be an "η_{Anchor}-noisy" copy of M_{value} and X equals x with probability $\mu_{X|Y}(x|m)$, where m is the value of M_{value}.

This specifies the distribution $\hat{P}$. We recall the following properties about $\hat{P}$ from [BVY21], which will be important for the parallel repetition theorem.

Claim 12.3.4 (Claim 4.1 of [BVY21])**.** *Conditioned on (M_{player}, M_{value}) the random variables X and Y are independent.*

Claim 12.3.5 (Claim 4.2 of [BVY21])**.** $\hat{\mathsf{P}}_{XY|M_{player}=A}(x, y) = \hat{\mathsf{P}}_{XY|M_{player}=B}(x, y) = \mu_{XY}(x, y)$ *for all $(x, y) \in \mathcal{X}^2$. In particular, the marginal distribution $\hat{\mathsf{P}}_{XY}$ is identical to the game distribution μ.*

We remark that Claim 12.3.4 and Claim 12.3.5 states that, if (x, y) is sampled according to $\hat{\mathsf{P}}$, then depending on whether $\mathbf{M}_{\text{player}}$ and $\mathbf{M}_{\text{value}}$ are conditioned, the distribution X and Y either follows the original distribution for the game, or becomes uncorrelated. We define the distribution $\mathsf{P} = (\mathbf{M}_{\text{player}}, \mathbf{M}_{\text{value}}, \mathbf{X}, \mathbf{Y}, \mathbf{A}, \mathbf{B})$ for the entirety of the input set as follows. Let $\mathbf{M}_{\text{player}} = ((\mathbf{M}_{\text{player}})_0, \dots, (\mathbf{M}_{\text{player}})_{r-1})$, $\mathbf{M}_{\text{value}} = ((\mathbf{M}_{\text{value}})_0, \dots, (\mathbf{M}_{\text{value}})_{r-1})$, $\mathbf{X} = (\mathbf{X}_0, \dots, \mathbf{X}_{r-1})$, and $\mathbf{Y} = (\mathbf{Y}_0, \dots, \mathbf{Y}_{r-1})$ be vectors of random variables, define

$$\mathsf{P}_{\mathbf{M}_{\text{player}}\mathbf{M}_{\text{value}}\mathbf{X}\mathbf{Y}} = \prod_{i=0}^{r} \hat{\mathsf{P}}_{(\mathbf{M}_{\text{player}})_i (\mathbf{M}_{\text{value}})_i \mathbf{X}_i \mathbf{Y}_i}.$$

Finally, we define random variables $\mathbf{A} = (\mathbf{A}_i)_{i\in[r]}$ and $\mathbf{B} = (\mathbf{B}_i)_{i\in[r]}$ as follows. When conditioned on $\mathbf{X}$ and $\mathbf{Y}$, the random variables $\mathbf{A}, \mathbf{B}$ are independent of $\mathbf{D}$ and $\mathbf{M}$, and for all realizations $\vec{x}, \vec{y}$ of $\vec{x}$ and $\vec{y}$, define

$$\mathsf{P}_{\mathbf{AB}|\mathbf{X}=\vec{x},\mathbf{Y}=\vec{y}}(\vec{a}, \vec{b}) = \langle\psi| A_{\vec{x}}^{\vec{a}} B_{\vec{y}}^{\vec{b}} |\psi\rangle$$

for all $\vec{a} = (\vec{a}_1, \ldots, \vec{a}_n) \in \mathscr{A}^n$ and $\vec{b} = (\vec{b}_1, \ldots, \vec{b}_n) \in \mathscr{B}^n$. The following claim connects the distribution P to the correlation set for the strategy.

Claim 12.3.6 (Claim 4.3 of [BVY21])**.** *The marginal distribution of* **XYAB** *is identical to the correlation set obtained in the repeated game $\mathscr{G}^{\otimes r}$ when the provers use the strategy $\mathscr{S}^{\otimes r}$.*

We use the probability P as the distribution which models the behavior of $\mathscr{S}^{\otimes r}$ for the proof of Theorem 12.2.3. Having define the tuple of distribution P, we are ready to define dependency breaking variables. These are crucial for controlling the correlations that arise when conditioning the distribution P on different events. Furthermore, we use $\mathbf{Q}_C = (\mathbf{X}_C, \mathbf{Y}_C)$ and $\mathbf{S}_C = (\mathbf{A}_C, \mathbf{B}_C)$ to denote random variables associated with the provers' questions and answers in the coordinates indexed by the critical set C, and we use $\mathbf{R}_C = (\mathbf{Q}_C, \mathbf{S}_C)$ to denote the correlation for $\mathscr{S}^{\otimes r}$ over the subset of coordinate C. We remark that the event $\mathbf{S}_C$ could potentially depend on the question pair from the coordinates outside of C. For $i \in [r]$ let W_i denote the indicator variable for the event that the provers win round i of the game for probability distribution P. Let $W_C = \prod_{i \in C} W_i$, we define the *dependency breaking variable* as:

Definition 12.3.7 (Dependency breaking variable)**.** *Let $C \subseteq [n]$. For all $i \in [r] \setminus C$ define the i-th* dependency-breaking variable $\mathbf{\Omega}_i$ *as*

$$\mathbf{\Omega}_i = ((\boldsymbol{M}_{player})_i, (\boldsymbol{M}_{value})_i) .$$

Furthermore we define

$$\mathbf{\Omega} = (\mathbf{\Omega}_i)_{i \in [r] \setminus C} \qquad \text{and} \qquad \mathbf{\Omega}_{-i} = (\mathbf{\Omega}_j)_{j \in [r] \setminus (C \cup \{i\})}, \quad \forall i \in [r] \setminus C . \tag{12.16}$$

We remark that when $\eta_{\text{Anchor}} = 0$ the definition of the production distribution $\mathbf{\Omega}_i \mathbf{Q}_C$ coincides with the one used by Holenstein [Hol09]; in that case, the variable $(\mathbf{M}_{\text{player}})_i$ is coupled to either $\mathbf{X}_i$ or $\mathbf{Y}_i$ exactly. Here we set η_{Anchor} to be a nonzero value that is related to the anchoring probability, as in (12.15). This "noisy coupling" between $\mathbf{\Omega}_i$ and the inputs $(\vec{x}_i, \vec{y}_i)$ is important for our analysis. Base on the above definition, we have the following claim.

Claim 12.3.8 (Claim 4.5 of [BVY21])**.** *The following properties hold for the distribution* P.

1. *The joint distribution $(\mathbf{X}, \mathbf{Y}, \mathbf{\Omega})$ is product across its r triples of coordinates. Furthermore, for any i, $\mathbf{X}_i$ and $\mathbf{Y}_i$ are independent conditioned on $\mathbf{\Omega}_i$. In particular, $\mathsf{P}_{\mathbf{\Omega}_i \mathbf{X}_i \mathbf{Y}_i} = \mathsf{P}_{\mathbf{\Omega}_i} \mathsf{P}_{\mathbf{X}_i | \mathbf{\Omega}_i} \mathsf{P}_{\mathbf{Y}_i | \mathbf{\Omega}_i}$.*
2. $\mathsf{P}_{\mathbf{S}_C | \mathbf{XY}} = \mathsf{P}_{\mathbf{S}_C | \mathbf{\Omega XY}}$.

Intuitively, this claim shows that the distribution of the answer is independent of the choices of Ω. By pre-sampling on Ω, the provers can sample the question indexed on $[r] \setminus C$ independently. This is the crucial property which allows to estimate the post measurement state of the provers conditioning on winning coordinates W_C, where recall W_C is the event where the provers win on all the coordinates C.

The following lemma shows that if the event W_C occurs with significant probability then conditioning on W_C only has a moderate effect on the distribution of $(\mathbf{X}_i, \mathbf{Y}_i)$, on average over a uniformly random choice of $i \in [r] \setminus C$. Furthermore, the distribution of $\mathbf{\Omega}_{-i} \mathbf{Q}_C \mathbf{S}_C$ is close to being independent from $(\mathbf{X}_i, \mathbf{Y}_i)$.

Lemma 12.3.9 (Lemma 4.6 of [BVY21])**.** *Let $C \subseteq [n]$ be the critical subset which arises from Proposition 12.3.3, we define the constant*

$$\eta_{PR} = \frac{1}{r - |C|} \left(\log \frac{1}{\mathsf{P}(W_C)} + |C| \log |\mathscr{A}|^2 \right) . \tag{12.17}$$

Then following inequalities hold:

1. $\frac{1}{r-|C|}\sum_{i=1}^{m}\|\mathsf{P}_{\mathbf{\Omega}_i\mathbf{X}_i\mathbf{Y}_i|W_C}-\mathsf{P}_{\mathbf{\Omega}_i\mathbf{X}_i\mathbf{Y}_i}\|\leq\sqrt{\eta_{PR}}$.
2. $\frac{1}{r-|C|}\sum_{i=1}^{m}\|\mathsf{P}_{\mathbf{\Omega}\mathbf{Q}_C\mathbf{X}_i\mathbf{Y}_i|W_C}-\mathsf{P}_{\mathbf{\Omega}\mathbf{Q}_C|W_C}\mathsf{P}_{\mathbf{X}_i\mathbf{Y}_i|\mathbf{\Omega}_i}\|\leq\sqrt{\eta_{PR}}$.
3. $\frac{1}{r-|C|}\sum_{i=1}^{m}\|\mathsf{P}_{\mathbf{\Omega}_i|W_C}\mathsf{P}_{\mathbf{\Omega}_{-i}\mathbf{Q}_C|\mathbf{X}_i=\perp,\mathbf{Y}_i=\perp,W_C}-\mathsf{P}_{\mathbf{\Omega}_i|W_C}\mathsf{P}_{\mathbf{\Omega}_{-i}\mathbf{Q}_C|\mathbf{\Omega}_i W_C}\|=O(\sqrt{\eta_{PR}})$.
4. $\frac{1}{r-|C|}\sum_{i=1}^{m}\|\mathsf{P}_{\mathbf{X}_i\mathbf{Y}_i}\mathsf{P}_{\mathbf{\Omega}_{-i}\mathbf{Q}_C|\mathbf{X}_i=\perp,\mathbf{Y}_i=\perp,W_C}-\mathsf{P}_{\mathbf{X}_i\mathbf{Y}_i}\mathsf{P}_{\mathbf{\Omega}_{-i}\mathbf{Q}_C|\mathbf{X}_i,\mathbf{Y}_i,W_C}\|=O(\sqrt{\eta_{PR}})$.

We use η_{PR} to denote the constant given in (12.17) for the critical set C for the remainder of this chapter.

12.3.3 Notation for the proof of Theorem 12.2.3

In this subsection, we set up notation used for the proof of Theorem 12.2.3. We note much of this notation is similar to notation from [BVY21, Section 4.4]. Recall from the previous subsection that $\mathscr{S}^{\otimes r}=\{\mathscr{L}^2(\mathscr{A},\tau),|\psi\rangle,\{A^{\vec{a}}_{\vec{x}}\},\{B^{\vec{b}}_{\vec{y}}\}\}$ is a tracially embeddable strategy for $\mathscr{G}^{\otimes r}$ for some anchoring game $\mathscr{G}$ which violates the bound given by Theorem 12.2.3. Let C be the subset promise by Proposition 12.3.3. For each $(\vec{a}_C,\vec{b}_C)\in(\mathscr{A}^{2C})$, $(\vec{x},\vec{y})\in\mathscr{X}^{2n}$, define

$$A^{\vec{x}}_{\vec{a}_C}=\sum_{\vec{a}|\vec{a}_C}A^{\vec{x}}_{\vec{a}}\qquad\text{and}\qquad B^{\vec{y}}_{\vec{b}_C}=\sum_{\vec{b}|\vec{b}_C}B^{\vec{y}}_{\vec{b}}\,,\tag{12.18}$$

where $\vec{a}|\vec{a}_C$ (resp. $\vec{b}|\vec{b}_C$) indicates summing over all tuples $\vec{a}$ consistent with $\vec{a}_C$ (resp. $\vec{b}$ consistent with $\vec{b}_C$). We see that the set $\{A^{\vec{x}}_{\vec{a}_C}\}$ (resp. $\{B^{\vec{y}}_{\vec{b}_C}\}$) denotes a POVM with outcomes in the set $\mathscr{A}^C$ (resp. $\mathscr{B}^C$), for all $\vec{x}$ (resp. $\vec{y}$). For all $i\in[r-|C|]$, $\omega_{-i}\in\Omega_{-i}$, $(\vec{x}_C,\vec{y}_C)\in\mathbf{Q}_C$, and $(x,y)\in\mathscr{X}^2$, define

$$A^{(\omega_{-i},\vec{x}_C),x}_{\vec{a}_C}=\mathop{\mathbb{E}}_{\vec{x}|\omega_{-i},\vec{x}_C,x}A^{\vec{x}}_{\vec{a}_C}\qquad\text{and}\qquad B^{(\omega_{-i},\vec{y}_C),y}_{\vec{b}_C}=\mathop{\mathbb{E}}_{\vec{y}|\omega_{-i},\vec{y}_C,y}B^{\vec{y}}_{\vec{b}_C}\,,\tag{12.19}$$

where $\mathbb{E}_{\vec{x}|\omega_{-i},\vec{x}_C,x}$ is shorthand for $\mathbb{E}_{\vec{x}|\mathbf{\Omega}_{-i}=\omega_{-i},\vec{x}_C=\vec{x}_C,\vec{x}_i=x}$ and similarly for $\mathbb{E}_{\vec{y}|\omega_{-i},\vec{y}_C,y}$. Let $\mathscr{X}_{/\perp}=\{\perp/x:x\in\mathscr{X}\}$ be a disjoint copy of $\mathscr{X}$. Here, for each $x\in\mathscr{X}$, "$\perp/x$" is a new symbol that is used to distinguish elements in $\mathscr{X}$ from elements in $\mathscr{X}_{/\perp}$. For all $\perp/x\in\mathscr{X}_{/\perp}$ define

$$A^{(\omega_{-i},\vec{x}_C),\perp/x}_{\vec{a}_C}=\eta_{\mathrm{Anchor}}\,A^{(\omega_{-i},\vec{x}_C),\perp}_{\vec{a}_C}+(1-\eta_{\mathrm{Anchor}})\,A^{(\omega_{-i},\vec{x}_C),x}_{\vec{a}_C}\,.\tag{12.20}$$

We remark that $A^{(\omega_{-i},\vec{x}_C),\perp/x}_{\vec{a}_C}$ can be equivalently defined as $\mathbb{E}_{\vec{x}|\mathbf{\Omega}_{-i}=\omega_{-i},\vec{x}_C=\vec{x}_C,((\mathbf{M}_{\mathrm{player}})_i,(\mathbf{M}_{\mathrm{value}})_i)=(A,x)}$, and further remark that

$$A^{(\omega_{-i},\vec{x}_C),\perp/\perp}_{\vec{a}_C}=A^{(\omega_{-i},\vec{x}_C),\perp}_{\vec{a}_C}.\tag{12.21}$$

Using that all operators are positive semidefinite, we observe for later use that

$$A^{(\omega_{-i},\vec{x}_C),\perp}_{\vec{a}_C}\leq\frac{1}{\eta_{\mathrm{Anchor}}}A^{(\omega_{-i},\vec{x}_C),\perp/x}_{\vec{a}_C}\,,\tag{12.22}$$

$$A^{(\omega_{-i},\vec{x}_C),x}_{\vec{a}_C}\leq\frac{1}{1-\eta_{\mathrm{Anchor}}}A^{(\omega_{-i},\vec{x}_C),\perp/x}_{\vec{a}_C}\,.\tag{12.23}$$

For all $i\in[n]\setminus C$, $\omega_{-i}\in\mathbf{\Omega}_{-i}$, $\vec{r}_C=(\vec{x}_C,\vec{y}_C,\vec{a}_C,\vec{b}_C)\in\mathbb{R}_C$, $x\in\mathscr{X}$, for all $s\in\{x,\perp/x\}$, and $y\in\mathscr{X}$, define the (unnormalized) state

$$|\Phi_{(\omega_{-i},\vec{r}_C),s,y}\rangle=\left(A^{(\omega_{-i},\vec{r}_C),s}_{\vec{a}_C}\right)^{1/2}\left(B^{(\omega_{-i},\vec{r}_C),y}_{\vec{b}_C}\right)^{1/2}|\psi\rangle\tag{12.24}$$

and the normalization factor

$$\gamma_{(\omega_{-i},\vec{r}_C),s,y}=\||\Phi_{(\omega_{-i},\vec{r}_C),s,y}\rangle\|\,.\tag{12.25}$$

Finally for $\gamma_{(\omega_{-i},\vec{r}_C),s,y} \neq 0$, we let

$$|\widetilde{\Phi}_{(\omega_{-i},\vec{r}_C),s,y}\rangle = \gamma^{-1}_{(\omega_{-i},\vec{r}_C),s,y} |\Phi_{(\omega_{-i},\vec{r}_C),s,y}\rangle, \tag{12.26}$$

and $|\widetilde{\Phi}_{(\omega_{-i},\vec{r}_C),s,y}\rangle = 0$ otherwise, this denotes the normalized version of the state $|\Phi_{(\omega_{-i},\vec{r}_C),s,y}\rangle$. Intuitively, for $s \in \mathscr{X}$, the state $|\widetilde{\Phi}_{(\omega_{-i},\vec{r}_C),s,y}\rangle$ corresponds to Alice and Bob first perform the following pre-sampling procedure for the strategy $\mathscr{S}^{\otimes r}$:

- For the coordinates in $[n] \setminus C$, Alice and Bob have pre-sampled $\mathbf{\Omega}_{-i} = \omega_{-i}$.
- For the coordinates in $C \cup \{i\}$, the questions are sample normally.

In this case, $\gamma_{(\omega_{-i},\vec{r}_C),s,y}$ corresponds to the expected probability in which Alice and Bob obtain the answer pair $(\vec{a}_C, \vec{b}_C)$ given the pre-sampling procedure above using the strategy $\mathscr{S}^{\otimes r}$. The state $|\widetilde{\Phi}_{(\omega_{-i},\vec{r}_C),s,y}\rangle$ correspond to the average post-measurement state for obtaining the answer pair $(\vec{a}_C, \vec{b}_C)$.

Similarly, for $s \in \mathscr{X}_\perp$, the scenario is similar as above, except for the coordinate i, Alice and Bob have pre-sampled $\mathbf{\Omega}_i = (A, s)$ as the outcome instead of sampling normally. By (12.21)

$$|\widetilde{\Phi}_{(\omega_{-i},\vec{r}_C),\perp/\perp,y}\rangle = |\widetilde{\Phi}_{(\omega_{-i},\vec{r}_C),\perp,y}\rangle \tag{12.27}$$

for all $y \in \mathscr{X}$.

For notation convenience, since many of the operator will be used in the context in which is over an expectation over the variable i, ω_{-i} and $\vec{r}_C$. For the clarity of notation, we often omit these subscripts if it is clear from context. For example, for some fix i and the operator $A_{\omega_{-i},\perp/\vec{x}_i}(\vec{a}_C)$ expected over $(\omega_{-i}, \vec{r}_C) \sim (\mathbf{\Omega}_{-i} \times \mathbf{R}_C)$, we will instead write $\mathbb{E}_{(\omega_{-i},\vec{r}_C)} A_{\perp/x}$. As another example, for the state $|\widetilde{\Phi}_{x,y}\rangle$ expected over $(\omega_{-i}, \vec{r}_C) \sim \mathbf{\Omega}_{-i} \times \mathbf{R}_C$, this is use to represent the state $|\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,x}\rangle$. To make the above intuition more concrete, we have the following proposition.

Proposition 12.3.10. *For all $s \in \mathscr{X}$,*

$$\gamma_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y} = \left(\mathsf{P}_{\mathbf{A}_C\mathbf{B}_C|\mathbf{\Omega}_{-i}=\omega_{-i},\mathbf{X}_i=x,\mathbf{Y}_i=y}(\vec{a}_C, \vec{b}_C)\right)^{1/2},$$

and for all $s = \perp/x \in \mathscr{X}_{/\perp}$,

$$\begin{aligned}\gamma_{(\omega_{-i},\vec{r}_C),s,y} &= \left(\eta_{Anchor}\mathsf{P}_{\mathbf{A}_C\mathbf{B}_C|\mathbf{\Omega}_{-i}=\omega_{-i},\mathbf{X}_i=\perp,\mathbf{Y}_i=y}(\vec{a}_C, \vec{b}_C) + (1-\eta_{Anchor})\mathsf{P}_{\mathbf{A}_C\mathbf{B}_C|\mathbf{\Omega}_{-i}=\omega_{-i},\mathbf{X}_i=x,\mathbf{Y}_i=y}(\vec{a}_C, \vec{b}_C)\right)^{1/2} \\ &= \mathsf{P}_{\mathbf{A}_C\mathbf{B}_C|\mathbf{\Omega}_{-i}=\omega_{-i},\mathbf{\Omega}_i=(A,x),\mathbf{Y}_i=y}(\vec{a}_C, \vec{b}_C)^{1/2}. \end{aligned} \tag{12.28}$$

The proof of this proposition is similar to [BVY21, Proposition 4.9] by expanding the definition of $A^{(\omega_{-i},\vec{r}_C),x}_{\vec{a}_C}$ and $B^{(\omega_{-i},\vec{r}_C),y}_{\vec{b}_C}$.

12.3.4 Proof of Theorem 12.2.3

The proof for Theorem 12.2.3 requires the following proposition:

Proposition 12.3.11. *For every $C \subseteq [r]$, $i \in [r]\setminus C$, $(\omega_{-i}, \vec{r}_C)$, $x, y \in \mathscr{X}$, there exist two unitary operator $U_{(\omega_{-i},\vec{r}_C),x} \in \mathscr{A}$ and $V_{(\omega_{-i},\vec{r}_C),y} \in \mathscr{A}'$ such that*

$$\mathop{\mathbb{E}}_{I} \mathop{\mathbb{E}}_{(\omega_{-i},\vec{r}_C)|W_C} \mathop{\mathbb{E}}_{XY} \| U_{(\omega_{-i},\vec{r}_C),x} V_{(\omega_{-i},\vec{r}_C),y} |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,\perp}\rangle - |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y}\rangle \| = O\big(\eta_{PR}^{1/16}\big),$$

where $\mathbb{E}_I$ denotes the expectation over a uniformly random $i \in [r] \setminus C$, $\mathbb{E}_{(\omega_{-i},\vec{r}_C)|W_C}$ denotes the expectation over $(\omega_{-i}, \vec{r}_C)$ sampled from $\mathsf{P}_{(\omega_{-i},\vec{r}_C)|W_C}$, and $\mathbb{E}_{XY}$ denotes the expectation over (x, y) sampled from μ.

We remark that this is an analogue of [BVY21, Proposition 5.1] for tracially embeddable strategies, and we give the proof in Section 12.3.5. Based on the above proposition, we give a proof for Theorem 12.2.3 below.

Proof. Let $\mathcal{G}$ be an anchored game, and supposed that there exist a tracially embeddable strategy $\mathcal{S}^{\otimes r} = \{\mathscr{L}^2(\mathscr{A},\tau), |\psi\rangle, \{A_{\vec{a}}^{\vec{x}}\}, \{B_{\vec{b}}^{\vec{y}}\}\}$ which violates (12.13). Let C be the critical subset C as promised by Proposition 12.3.3, and recall, without the lost of generality, we assume C is the first $|C|$ coordinates of $[r]$. Fix $i \in [r]\setminus C$, $(\omega_{-i}, \vec{r}_C = (\vec{x}_C, \vec{y}_C, \vec{a}_C, \vec{b}_C)) \in (\mathbf{\Omega}_{-i}, \mathbb{R}_C)$ and $(x,y) \in \mathscr{X}^2$. For each $\vec{a}$ such that $\pi_{\leq|C|}(\vec{a}) = \vec{a}_C$, by (12.19), we have $A_{\vec{a}}^{(\omega_{-i},\vec{x}_C),x} \leq A_{\vec{a}_C}^{(\omega_{-i},\vec{x}_C),x}$, and hence by Lemma 3.2.1, there exist a positive element $\widehat{A}_{\vec{a}}^{(\omega_{-i},\vec{x}_C),x}$ such that

$$A_{\vec{a}}^{(\omega_{-i},\vec{x}_C),x} = \left(A_{\vec{a}_C}^{(\omega_{-i},\vec{x}_C),x}\right)^{1/2} \cdot \widehat{A}_{\vec{a}}^{(\omega_{-i},\vec{x}_C),x} \cdot \left(A_{\vec{a}_C}^{(\omega_{-i},\vec{x}_C),x}\right)^{1/2}.$$

Likewise, for each $\vec{b}$ such that $\pi_{\leq|C|}(\vec{b}) = \vec{b}_C$, there exist a positive element $\widehat{B}_{\vec{b}}^{(\omega_{-i},\vec{y}_C),y}$ such that

$$B_{\vec{b}}^{(\omega_{-i},\vec{y}_C),y} = \left(B_{\vec{b}_C}^{(\omega_{-i},\vec{y}_C),y}\right)^{1/2} \cdot \widehat{B}_{\vec{b}}^{(\omega_{-i},\vec{y}_C),y} \cdot \left(B_{\vec{b}_C}^{(\omega_{-i},\vec{y}_C),y}\right)^{1/2}.$$

For fixed $(\vec{a}_C, \vec{b}_C)$, both $\{\widehat{A}^{\vec{a}}_{(\omega_{-i},\vec{x}_C),x}\}_{\vec{a}|\vec{a}_C}$ and $\{\widehat{B}^{\vec{b}}_{(\omega_{-i},\vec{y}_C),y}\}_{\vec{b}|\vec{b}_C}$ forms a POVM, where $\vec{a}|\vec{a}_C$ (resp. $\vec{b}|\vec{b}_C$) denotes summing over $\vec{a}$ such that $\pi_{\leq|C|}(\vec{a}) = \vec{a}_C$ (resp. $\pi_{\leq|C|}(\vec{b}) = \vec{b}_C$). For each $i \in [r]\setminus C$ and $(\omega_{-i}, \vec{r}_C = (\vec{x}_C, \vec{y}_C, \vec{a}_C, \vec{b}_C)) \in (\mathbf{\Omega}_{-i}, \mathbf{R}_C)$, we define a tracially embeddable strategy $\mathcal{S}_{(\omega_{-i},\vec{r}_C)} = \{\mathscr{L}^2(\mathscr{A},\tau), |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,\perp}\rangle, \{\widetilde{A}_a^{(\omega_{-i},\vec{r}_C),x}\}_{(x,a)\in\mathscr{X}\times\mathscr{A}}, \widetilde{B}_b^{(\omega_{-i},\vec{r}_C),y}\}_{(y,b)\in\mathscr{X}\times\mathscr{A}}\}$ for the game $\mathcal{G}$ as the following: The joint entanglement between the prover is $|\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,\perp}\rangle$ given in (12.26). The measurement operators $\{\widetilde{A}_a^{(\omega_{-i},\vec{r}_C),x}\}$, $\{\widetilde{B}_b^{(\omega_{-i},\vec{r}_C),y}\}$ for the strategy $\mathcal{S}_{(\omega_{-i},\vec{r}_C)}$ is defined as

$$\widetilde{A}_a^{(\omega_{-i},\vec{r}_C),x} = U^\dagger_{(\omega_{-i},\vec{r}_C),x}\left(\sum_{\vec{a}|\vec{a}_i=a,\vec{a}_C} \hat{A}_{\vec{a}}^{(\omega_{-i},\vec{x}_C),x}\right) U_{(\omega_{-i},\vec{r}_C),x}$$
$$\widetilde{B}_b^{(\omega_{-i},\vec{r}_C),y} = V^\dagger_{(\omega_{-i},\vec{r}_C),y}\left(\sum_{\vec{b}|\vec{b}_i=b,\vec{b}_C} \hat{B}_{\vec{b}}^{(\omega_{-i},\vec{y}_C),y}\right) V_{(\omega_{-i},\vec{r}_C),y},$$

where $\vec{a}|\vec{a}_i = a, \vec{a}_C$ (resp. $\vec{b}|\vec{b}_i = b, \vec{b}_C$) denotes summing over tuples $\vec{a}$ that are consistent with $\vec{a}_C$ and the ith coordinate is equal to a (resp. $\vec{b}$ that are consistent with $\vec{b}_C$ and the ith coordinate is equal to b), and $U_{(\omega_{-i},\vec{r}_C),x}$, $V_{(\omega_{-i},\vec{r}_C),y}$ are the unitary operators from Proposition 12.3.11. We remark that we choose to represent the measurement as $\{\widetilde{A}_a^{(\omega_{-i},\vec{r}_C),x}\}_{(x,a)\in\mathscr{X}\times\mathscr{A}}$ to emphasize that $\mathcal{S}_{(\omega_{-i},\vec{r}_C)}$ is a strategy for a single copy of $\mathcal{G}$ (i.e. given the question $x \in \mathscr{X}$, the first prover will measure using $\{\widetilde{A}_a^{(\omega_{-i},\vec{r}_C),x}\}_{a\in\mathscr{A}}$ on her half of the state $|\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,\perp}\rangle$).

We wish to show that on average over $(\omega_{-i}, \vec{r}_C)$ there exist a coordinate i such that the strategy $\mathcal{S}_{(\omega_{-i},\vec{r}_C)}$ violates the optimal success for the original game. Let $\mathsf{Q}_{AB|(\omega_{-i},\vec{r}_C),x,y}$ denotes the correlation $C_{x,y,a,b}$ given the strategy $\mathcal{S}_{(\omega_{-i},\vec{r}_C)}$, or

$$\begin{aligned}\mathsf{Q}_{AB|(\omega_{-i},\vec{r}_C),x,y}(a,b) &= \langle\widetilde{\Phi}_{\perp,\perp}|\tilde{A}_a^{(\omega_{-i},\vec{r}_C),x}\cdot\tilde{B}_b^{(\omega_{-i},\vec{r}_C),y}|\widetilde{\Phi}_{\perp,\perp}\rangle \\ &= \sum_{\vec{a}|\vec{a}_i=a,\vec{a}_C}\sum_{\vec{b}|\vec{b}_i=b,\vec{b}_C}\left(\langle\widetilde{\Phi}_{\perp,\perp}|U_x^\dagger V_y^\dagger\left(\widehat{A}_{\vec{a}}^{(\omega_{-i},\vec{x}_C),x}\cdot\widehat{B}_{\vec{b}}^{(\omega_{-i},\vec{x}_C),y}\right)U_xV_y|\widetilde{\Phi}_{\perp,\perp}\rangle\right).\end{aligned} \tag{12.29}$$

The following lemma shows that conditioning on $\vec{r}_C$ selected base on the question/answer pairs which wins on the critical set C, the strategy on coordinate i is closed to being independent from the other coordinates.

Lemma 12.3.12. *There exists a universal constant $\beta_{PR} \geq 1$ such that,*

$$\mathop{\mathbb{E}}_{I} \| \mathsf{P}_{(\boldsymbol{\Omega}_{-i},\mathbf{R}_C)|W_C} \cdot \mathsf{P}_{XY} \cdot \mathsf{Q}_{AB|(\omega_{-i},\vec{r}_C),x,y} - \mathsf{P}_{(\boldsymbol{\Omega}_{-i},\mathbf{R}_C)\mathbf{X}_i\mathbf{Y}_i\mathbf{A}_i\mathbf{B}_i|W_C} \| \leq \beta_{PR}\, \eta_{PR}^{1/16} ,$$

where η_{PR} is defined in (12.17) *and we identify (x,y,a,b) with $(\vec{x}_i,\vec{y}_i,\vec{a}_i,\vec{b}_i)$.*

Proof. We remark that this is essentially the same as [BVY21, Lemma 6.2]. We start with two claims, from which the proof of the lemma follows.

Claim 12.3.13. *For all $(\omega_{-i},\vec{r}_C), x, y$ and a, b,*

$$\sum_{\vec{a}|\vec{a}_i=a,\vec{a}_C} \sum_{\vec{b}|\vec{b}_i=b,\vec{b}_C} \left(\langle \widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y} | \big(\widehat{A}_{\vec{a}}^{(\omega_{-i},\vec{r}_C),x} \cdot \widehat{B}_{\vec{b}}^{(\omega_{-i},\vec{r}_C),y} \big) | \widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y} \rangle \right) = \mathsf{P}_{\vec{a}_i\vec{b}_i|(\omega_{-i},\vec{r}_C),x,y}(a,b) .$$

□

The proof of this proposition is similar to [BVY21, Claim 6.3] by expanding the definition of $\widehat{A}_{\vec{a}}^{(\omega_{-i},\vec{r}_C),x}$ and $\widehat{B}_{\vec{b}}^{(\omega_{-i},\vec{r}_C),y}$.

Claim 12.3.14. *The following holds:*

$$\mathop{\mathbb{E}}_{I} \| \mathsf{P}_{(\omega_{-i},\vec{r}_C)|W_C} \cdot \mathsf{P}_{\vec{x}_i\vec{y}_i} \cdot \mathsf{Q}_{\vec{a}_i\vec{b}_i|(\omega_{-i},\vec{r}_C)\vec{x}_i\vec{y}_i} - \mathsf{P}_{(\omega_{-i},\vec{r}_C)|W_C} \cdot \mathsf{P}_{\vec{x}_i\vec{y}_i} \cdot \mathsf{P}_{\vec{a}_i\vec{b}_i|(\omega_{-i},\vec{r}_C)\vec{x}_i\vec{y}_i} \| = O(\eta_{PR}^{1/16}) .$$

Proof. Fix $(\omega_{-i},\vec{r}_C), x, y$. We bound the total variation distance between the distribution Q and P below. For the ease of notation, we define

$$\hat{A}_a^x = \sum_{\vec{a}|\vec{a}_i=a,\vec{a}_C} \widehat{A}_{\vec{a}}^{(\omega_{-i},\vec{r}_C),x} \qquad \hat{B}_b^y = \sum_{\vec{b}|\vec{b}_i=b,\vec{b}_C} \widehat{A}_{\vec{b}}^{(\omega_{-i},\vec{r}_C),y},$$

for this proof. We see that $\{\hat{A}_a^x\}_{a\in\mathscr{A}}$ and $\{\hat{B}_b^y\}_{b\in\mathscr{A}}$ forms two sets of POVM. By Equation (12.29) and Claim 12.3.13

$$\begin{aligned}
&\| \mathsf{Q}_{\vec{a}_i\vec{b}_i|(\omega_{-i},\vec{r}_C),x,y} - \mathsf{P}_{\vec{a}_i\vec{b}_i|(\omega_{-i},\vec{r}_C),x,y} \| \\
&= \frac{1}{2} \sum_{a,b} | \langle \widetilde{\Phi}_{\perp,\perp} | U_x^\dagger V_y^\dagger \big(\hat{A}_a^x \cdot \hat{B}_b^y \big) U_x V_y | \widetilde{\Phi}_{\perp,\perp} \rangle - \langle \widetilde{\Phi}_{x,y} | \big(\hat{A}_a^x \cdot \hat{B}_b^y \big) | \widetilde{\Phi}_{x,y} \rangle | \\
&= \frac{1}{2} \sum_{a,b} | \langle \widetilde{\Phi}_{\perp,\perp} | U_x^\dagger V_y^\dagger \big(\hat{A}_a^x \cdot \hat{B}_b^y \big) U_x V_y | \widetilde{\Phi}_{\perp,\perp} \rangle - \langle \widetilde{\Phi}_{\perp,\perp} | U_x^\dagger V_y^\dagger \big(\hat{A}_a^x \cdot \hat{B}_b^y \big) | \widetilde{\Phi}_{x,y} \rangle \\
&\quad + \langle \widetilde{\Phi}_{\perp,\perp} | U_x^\dagger V_y^\dagger \big(\hat{A}_a^x \cdot \hat{B}_b^y \big) | \widetilde{\Phi}_{x,y} \rangle - \langle \widetilde{\Phi}_{x,y} | \big(\hat{A}_a^x \cdot \hat{B}_b^y \big) | \widetilde{\Phi}_{x,y} \rangle | \\
&\leq \frac{1}{2} \sum_{a,b} \Big(\| \big(\hat{A}_a^x \cdot \hat{B}_b^y \big) U_x V_y | \widetilde{\Phi}_{\perp,\perp} \rangle \| + \| \big(\hat{A}_a^x \cdot \hat{B}_b^y \big) | \widetilde{\Phi}_{x,y} \rangle \| \Big) \cdot \| U_x V_y | \widetilde{\Phi}_{\perp,\perp} \rangle - | \widetilde{\Phi}_{x,y} \rangle \| \\
&= \frac{1}{2} \left(\sum_{a,b} \| \big(\hat{A}_a^x \cdot \hat{B}_b^y \big) U_x V_y | \widetilde{\Phi}_{\perp,\perp} \rangle \| + \sum_{a,b} \| \big(\hat{A}_a^x \cdot \hat{B}_b^y \big) | \widetilde{\Phi}_{x,y} \rangle \| \right) \cdot \| U_x V_y | \widetilde{\Phi}_{\perp,\perp} \rangle - | \widetilde{\Phi}_{x,y} \rangle \| \\
&\leq \| U_x V_y | \widetilde{\Phi}_{\perp,\perp} \rangle - | \widetilde{\Phi}_{x,y} \rangle \| .
\end{aligned} \tag{12.30}$$

Where the last line follows from Jensen's inequality and $\{\hat{A}_a^x\}_{a\in\mathscr{A}}$ and $\{\hat{B}_b^y\}_{b\in\mathscr{A}}$ being both POVM. Thus, returning to (12.3.14)

$$\begin{aligned}
&\mathop{\mathbb{E}}_{I} \| \mathsf{P}_{(\omega_{-i},\vec{r}_C)|W_C} \cdot \mathsf{P}_{\vec{x}_i\vec{y}_i} \cdot \mathsf{Q}_{\vec{a}_i\vec{b}_i|(\omega_{-i},\vec{r}_C)\vec{x}_i\vec{y}_i} - \mathsf{P}_{(\omega_{-i},\vec{r}_C)|W_C} \cdot \mathsf{P}_{\vec{x}_i\vec{y}_i} \cdot \mathsf{P}_{\vec{a}_i\vec{b}_i|(\omega_{-i},\vec{r}_C)\vec{x}_i\vec{y}_i} \| \\
&\quad = \mathop{\mathbb{E}}_{I} \mathop{\mathbb{E}}_{(\omega_{-i},\vec{r}_C)|W_C} \mathop{\mathbb{E}}_{\vec{x}_i\vec{y}_i} \| \mathsf{Q}_{\vec{a}_i\vec{b}_i|(\omega_{-i},\vec{r}_C)\vec{x}_i\vec{y}_i} - \mathsf{P}_{\vec{a}_i\vec{b}_i|(\omega_{-i},\vec{r}_C)\vec{x}_i\vec{y}_i} \| \\
&\quad \leq \sqrt{2} \mathop{\mathbb{E}}_{I} \mathop{\mathbb{E}}_{(\omega_{-i},\vec{r}_C)|W_C} \mathop{\mathbb{E}}_{XY} \| U_{(\omega_{-i},\vec{r}_C),x} \otimes V_{(\omega_{-i},\vec{r}_C),y} | \widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,\perp} \rangle - | \widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y} \rangle \| \\
&\quad \leq O(\eta_{\mathrm{PR}}^{1/16}) ,
\end{aligned}$$

where the first inequality is by (12.30) and the last inequality follows from Proposition 12.3.11.

□

Based on Lemma 12.3.12, we are ready to construct a strategy for the single instance of the non-local game $\mathscr{G}$ which creates the contradiction. We remark that the remainder of this proof follows similarly as [BVY21, Section 6.2]. Recall by the definition of the critical set C in Proposition 12.3.3, we have

$$\mathop{\mathbb{E}}_{I} \mathsf{P}(W_i|W_C) \geq 1-\varepsilon/2\,,$$

and recall from Claim 12.3.8, since sampling each $\mathbf{\Omega}_i$ are independent from each coordinates, and the answers are independent from the distribution $\mathbf{\Omega}_i$ for $i \in [r]\setminus(C\cup\{i\})$. This implies by sampling a uniformly random $i \in [r]\setminus C$ and then sampling from the distribution $\mathsf{P}_{\vec{x}_i\vec{y}_i(\omega_{-i},\vec{r}_C)\vec{a}_i\vec{b}_i|W_C}$ yields a tuple $(i,\vec{x}_i,\vec{y}_i,(\omega_{-i},\vec{r}_C),\vec{a}_i,\vec{b}_i)$ such that $V(\vec{x}_i,\vec{y}_i,\vec{a}_i,\vec{b}_i)=1$ (i.e $W_i=1$) with probability at least $1-\varepsilon/2$. By Lemma 12.3.12, the distribution $\mathsf{P}_{\mathbf{X}_i\mathbf{Y}_i(\omega_{-i},\vec{r}_C)\mathbf{A}_i\mathbf{B}_i|W_C}$, is $\beta_{PR}\eta_{\mathrm{PR}}^{1/16}$ close to Alice and Bob performing the following strategy for the non-local game $\mathscr{G}$

1. Post-select the question and answer pair $r_C \sim \mathbf{R}_C$ for the critical set C in which they win on all C coordinates for the strategy $\mathscr{S}^n$, uniformly a coordinates $i \in [r]\setminus c$ and $(\omega_{-i}) \sim \Omega_{-i}$.
2. Up on receiving (x,y), perform the strategy $\mathscr{S}_{(\omega_{-i},\vec{r}_C)}$.

By convexity, there exists some coordinate i in which we can fix on step 1 of the above procedure which succeed with probability at least $1-\varepsilon/2-\beta_{PR}\eta_{\mathrm{PR}}^{1/16}$, where β_{PR} is the universal constant from Lemma 12.3.12. Let the strategy define above be $\mathscr{S}^{\mathrm{contra}}$. Set

$$c = \frac{1}{32\log(e)\,(4\beta_{PR})^{16}}.$$

By the initial contradiction assumption, we have

$$\omega(\mathscr{G}^{\otimes r},\mathscr{S}^{\otimes r}) \geq \frac{4}{\varepsilon}\exp\left(-\frac{c\,\varepsilon^{17}\,r}{\log(|\mathscr{A}|+1)}\right),$$

and by rearranging the equation for r, we have

$$\frac{\log(|\mathscr{A}|+1)}{c\cdot\varepsilon^{17}}\ln\left(\frac{4}{\varepsilon\cdot\omega(\mathscr{G}^{\otimes r},\mathscr{S}^{\otimes r})}\right) \leq n.$$

Hence, we have

$$r \geq \frac{r}{\log(|\mathscr{A}|+1)} \geq \frac{1}{c\cdot\varepsilon^{17}}\ln\left(\frac{4}{\varepsilon\cdot\omega(\mathscr{G}^{\otimes r},\mathscr{S}^{\otimes r})}\right) \geq \frac{16}{\varepsilon}\log\left(\frac{4}{\varepsilon\cdot\omega(\mathscr{G}^{\otimes r},\mathscr{S}^{\otimes r})}\right),$$

where we use that $0<\varepsilon\leq 1$, and $0<c\leq\frac{\varepsilon^{16}}{16\cdot\log(e)}$. Hence, if we consider η_{PR}, using the fact that $|C|\leq\frac{8}{\varepsilon}\log\frac{4}{\varepsilon\cdot\mathsf{P}(W)}\leq r/2$ from Proposition 12.3.3,

$$\begin{aligned}
\eta_{\mathrm{PR}} &= \frac{1}{r-|C|}\left(\log\frac{1}{\mathsf{P}(W_C)}+|C|\log|\mathscr{A}|^2\right)\\
&\leq \frac{2}{n}\left(\frac{16\cdot\log|\mathscr{A}|^2}{\varepsilon}\log\frac{4}{\varepsilon\cdot\mathsf{P}(W)}\right)\\
&\leq \frac{2}{n}\cdot\frac{16\cdot s}{\varepsilon}\cdot\frac{c\log(e)\,\varepsilon^{17}\,n}{s}\\
&= 32\,c\log(e)\,\varepsilon^{16}\,.
\end{aligned}$$

This implies that

$$\omega(\mathscr{G},\mathscr{S}^{\mathrm{contra}}) \geq 1-\varepsilon/2-\beta_{PR}\cdot\eta_{\mathrm{PR}}^{1/16} \geq 1-\varepsilon/2-\beta_{PR}\eta_{\mathrm{PR}} > 1-\varepsilon, \tag{12.31}$$

giving a strategy which wins $\mathscr{G}$ is strictly greater than $1-\varepsilon$, a contradiction. This concludes the proof for Theorem 12.2.3. □

12.3.5 Existence of the local unitary

In this section, we give a proof for Proposition 12.3.11. Similar to the proof of [BVY21, Proposition 5.1], the proof relies on two main lemmas. The first lemma, given below, guarantees the existence of a local unitary operator which allows the provers to make the local adjustment as per described in Section 12.2.1. We remark that this is a commuting operator model variant of [BVY21, Proposition 5.1], and the proof follows a similar structure.

Lemma 12.3.15. *For all i, $(\omega_{-i}, \vec{r}_C)$, x and y there exists two unitary operator $U_{(\omega_{-i},\vec{r}_C),x} \in \mathscr{A}$ and $V_{(\omega_{-i},\vec{r}_C),y} \in \mathscr{A}'$ such that with probability at least $1 - O(\eta_{PR}^{1/16})$ over the choice of a uniformly random $i \in [n] \setminus C$,*

$$\mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|W_C} \mathop{\mathbb{E}}_{\mathbf{X}} \| U_{(\omega_{-i},\vec{r}_C),x} |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,\perp}\rangle - |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,\perp}\rangle \| = O(\eta_{PR}^{1/16}) \,, \tag{12.32}$$

$$\mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|W_C} \mathop{\mathbb{E}}_{Y} \| V_{(\omega_{-i},\vec{r}_C),y} |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,\perp}\rangle - |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,y}\rangle \| = O(\eta_{PR}^{1/16}) \,, \tag{12.33}$$

$$\mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|W_C} \mathop{\mathbb{E}}_{\mathbf{XY}} \| V_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y} |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y}\rangle - |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,\perp}\rangle \| = O(\eta_{PR}^{1/16}) \,. \tag{12.34}$$

where $\mathbb{E}_{\mathbf{X}}$, $\mathbb{E}_{\mathbf{Y}}$, and $\mathbb{E}_{\mathbf{XY}}$ denote expectations under $\mu_X(x)$, $\mu_Y(y)$, and $\mu(x,y)$ respectively.

The second lemma, given below, relates the normalization factors $\gamma_{x,y}$ and $\gamma_{\perp/x,y}$ defined in (12.25). Since the second lemma is identical to [BVY21, Lemma 5.17], we instead refer the reader to the original reference for the proof.

Lemma 12.3.16. *With probability at least $1 - O(\eta_{PR}^{1/4})$ over the choice of $i \in [n] \setminus C$,*

$$\mathop{\mathbb{E}}_{\mathbf{XY}} \mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbb{R}_C|\mathbf{X}_i=\perp,\mathbf{Y}_i=\perp,W_C} \left|1 - \frac{\gamma_{(\omega_{-i},\vec{r}_C),x,y}}{\gamma_{(\omega_{-i},\vec{r}_C),\perp,\perp}}\right|^2 \leq O(\eta_{PR}^{1/4}) \,, \tag{12.35}$$

and

$$\mathop{\mathbb{E}}_{\mathbf{XY}} \mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbb{R}_C|\mathbf{X}_i=\perp,\mathbf{Y}_i=\perp,W_C} \left|1 - \frac{\gamma_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y}}{\gamma_{(\omega_{-i},\vec{r}_C),\perp,\perp}}\right|^2 \leq O(\eta_{PR}^{1/4}) \,. \tag{12.36}$$

Local operator lemma

In this subsection, we give a proof for Lemma 12.3.15. Recall from the previous section that the strategy $\mathscr{S}^{\otimes r} = \{\mathscr{L}^2(\mathscr{A},\tau), |\psi\rangle = \sigma|\tau\rangle, \{A^{\vec{a}}_{\vec{x}}\}, \{B^{\vec{b}}_{\vec{y}}\}\}$ is a tracially embeddable strategy which realizes a contradiction for Theorem 12.2.3. For all ω, $(\vec{x}_C, \vec{y}_C)$, $\vec{a}_C$, and $\mathbf{B}_C$, we define the measurement operator

$$A^{\omega,(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C} = \mathop{\mathbb{E}}_{\mathbf{X}|\boldsymbol{\Omega}=\omega,\mathbf{Q}_C=(\vec{x}_C,\vec{y}_C)} A^{\vec{x}}_{\vec{a}_C} \qquad \text{and} \qquad B^{\omega,(\vec{x}_C,\vec{y}_C)}_{\vec{b}_C} = \mathop{\mathbb{E}}_{\mathbf{Y}|\boldsymbol{\Omega}=\omega,\mathbf{Q}_C=(\vec{x}_C,\vec{y}_C)} B^{\vec{y}}_{\vec{b}_C} \tag{12.37}$$

where $A^{\vec{x}}_{\vec{a}_C}, B^{\vec{y}}_{\vec{b}_C}$ are defined in (12.18). For all $\omega, \vec{x}, \vec{y}, \vec{a}_C, \vec{b}_C$, we define the vector state

$$|\Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}\rangle = \left(A^{\omega,(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} \left(B^{\vec{y}}_{\vec{b}_C}\right)^{1/2} |\psi\rangle \tag{12.38}$$

$$|\Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}\rangle = \left(A^{\omega,(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} \left(B^{\vec{y}}_{\vec{b}_C}\right)^{1/2} |\psi\rangle \,, \tag{12.39}$$

where $|\Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}\rangle, |\Lambda_{\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C}\rangle \in \mathscr{L}^2(\mathscr{A},\tau)$. We remark that in the above definition, $A^{\omega,(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}$ and $B^{\vec{y}}_{\vec{b}_C}$ uses the same value of $\vec{y}$. Let $\Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}$ and $\Lambda_{\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C}$ denote the abstract normal states acting on $\mathscr{L}^2(\mathscr{A},\tau)$ specified by the vector $|\Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}\rangle$ and $|\Lambda_{\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C}\rangle$, respectively, i.e.

$$\Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}(A) = \langle \Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C} | A | \Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C} \rangle$$
$$\Lambda_{\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C}(A) = \langle \Lambda_{\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C} | A | \Lambda_{\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C} \rangle$$

Now define the classical-quantum states

$$\Xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}} = \sum_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C} \mathsf{P}_{\mathbf{\Omega X}_C\mathbf{Y}}(\omega,\vec{x}_C,\vec{y})\,\langle\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C| \otimes \Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}\,, \tag{12.40}$$

$$\Lambda^{\mathbf{\Omega XY}_C\mathbf{Q}_C\mathscr{A}} = \sum_{\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C} \mathsf{P}_{\mathbf{\Omega XY}_C}(\omega,\vec{x},\vec{y}_C)\,\langle\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C| \otimes \Lambda_{\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C}\,, \tag{12.41}$$

Both states are classical on the space $\mathbf{\Omega}$, $\mathbf{X}$, $\mathbf{X}_C$, $\mathbf{Y}$, $\mathbf{Y}_C$ and $\mathbf{Q}_C$ and quantum on the space $\mathscr{A}$.We remark that all classical register listed above are multiple classical registers, as each of them are probability distribution over multiple coordinates (i.e. $\mathbf{X} = (\mathbf{X}_0,\cdots,\mathbf{X}_{r-1})$). Observe that this state looks very similar – but not quite – to the one that occurs in an actual execution of the strategy $\mathscr{S}^{\otimes r}$. There are several important differences: one is that the measurements only produce answers for the coordinates indexed by C. Another difference is that the measurement operators $A^{\omega,(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}$ are not part of the strategy but instead are derived from the measurements operator. Also, note that there is no explicit register for question vector $\mathbf{X}$ (except for the $\mathbf{X}_C$ questions, which are included in the register $\mathbf{\Omega}$); instead these questions are implicitly averaged over within the $A^{\omega,(\vec{x}_C,\vec{y}_C)}$ measurement for a fix value of $(\omega,\vec{x}_C,\vec{y}_C)$.

The state $\Xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}$ is defined such that when restricted to the classical register $\Xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C}$, the resulting state representing the probability distribution $\mathsf{P}_{\mathbf{\Omega X}_C\mathbf{YA}_C\mathbf{B}_C}$. To see this, observe that for any $(\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C)$, we have

$$\begin{aligned}
\Xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C} &= \sum_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C} \mathsf{P}_{\mathbf{\Omega X}_C\mathbf{Y}}(\omega,\vec{x}_C,\vec{y})\cdot\langle\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C| \otimes \Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}(\mathbb{I}_{\mathscr{A}})\,,\\
&= \sum_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C} \left(\mathsf{P}_{\mathbf{\Omega X}_C\mathbf{Y}}(\omega,\vec{x}_C,\vec{y})\,\langle\psi|A^{\omega,(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}B^{\vec{y}}_{\vec{b}_C}|\psi\rangle\right)\cdot\langle\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C|\,,\\
&= \sum_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C} \mathsf{P}_{\mathbf{\Omega X}_C\mathbf{Y}}(\omega,\vec{x}_C,\vec{y})\left(\mathop{\mathbb{E}}_{\mathbf{X}|\mathbf{\Omega}=\omega,\mathbf{X}_C=\vec{x}_C,\mathbf{Y}=\vec{y}}\langle\psi|A^{\vec{x}}_{\vec{a}_C}B^{\vec{y}}_{\vec{b}_C}|\psi\rangle\right)\cdot\langle\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C|\,,\\
&= \sum_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C} \mathsf{P}_{\mathbf{\Omega X}_C\mathbf{Y}}(\omega,\vec{x}_C,\vec{y})\left(\mathop{\mathbb{E}}_{\mathbf{X}|\mathbf{X}_C=\vec{x}_C,\mathbf{Y}=\vec{y}}\mathsf{P}_{\mathbf{AB}|\mathbf{X}=\vec{x},\mathbf{Y}=\vec{y}}(\vec{a},\vec{b})\right)\cdot\langle\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C|\,,\\
&= \sum_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C} \left(\mathsf{P}_{\mathbf{\Omega X}_C\mathbf{YA}_C\mathbf{B}_C}(\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C)\right)\cdot\langle\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C|\,,
\end{aligned}$$

where in the third line we used Item 1 from Claim 12.3.8 and in the fourth line we used Item 2 from Claim 12.3.8 and each $\mathbf{Y}_i$ being independent of each other. By a similar calculation, the state $\Lambda^{\mathbf{\Omega XA}_C\mathbf{B}_C}$ represents the probability distribution $\mathsf{P}_{\mathbf{\Omega XY}_C\mathbf{A}_C\mathbf{B}_C}$. Since both $\mathsf{P}_{\mathbf{\Omega X}_C\mathbf{YA}_C\mathbf{B}_C}$ and $\mathsf{P}_{\mathbf{\Omega XY}_C\mathbf{A}_C\mathbf{B}_C}$ are probability distributions, the state $\Xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C}$ and $\Lambda^{\mathbf{\Omega XY}_C\mathbf{Q}_C}$ are indeed classical quantum states.

Recall, the event W_C corresponds to the event where the provers produces an winning answer given a winning question pair on all coordinates on the critical set C. Since the event W_C is determined by the random variables $(\mathbf{\Omega},\mathbf{R}_C)$ we can condition the states $\Xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}, \Lambda^{\mathbf{\Omega XY}_C\mathbf{Q}_C\mathscr{A}}$ on the event W_C to obtain states

$$\xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}} = \frac{1}{\mathsf{P}(W_C)} \sum_{\substack{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C:\\(\vec{x}_C,\vec{y}_C,\vec{a}_C,\vec{b}_C)\in W_C}} \mathsf{P}_{\mathbf{\Omega X}_C\mathbf{Y}}(\omega,\vec{x}_C,\vec{y})\cdot\langle\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C| \otimes \Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}\,,$$

$$\lambda^{\mathbf{\Omega XY}_C\mathbf{Q}_C\mathscr{A}} = \frac{1}{\mathsf{P}(W_C)} \sum_{\substack{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C:\\(\vec{x}_C,\vec{y}_C,\vec{a}_C,\vec{b}_C)\in W_C}} \mathsf{P}_{\mathbf{\Omega XY}_C}(\omega,\vec{x},\vec{y}_C)\cdot\langle\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C| \otimes \Lambda_{\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C}\,,$$

Since the event W_C is a subset of all possible coordinates in $\mathscr{X}^{2|C|}\times\mathscr{A}^{2|C|}$, by definition, we have

$$\mathsf{P}(W_C)\cdot\xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}} \le \Xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}} \qquad \mathsf{P}(W_C)\cdot\lambda^{\mathbf{\Omega XY}_C\mathbf{Q}_C\mathscr{A}} \le \Lambda^{\mathbf{\Omega XY}_C\mathbf{Q}_C\mathscr{A}}. \tag{12.42}$$

For a fix ω and $\vec{r}_C = (\vec{x}_C, \vec{y}_C, \vec{a}_C, \vec{b}_C) \in \mathscr{X}^{2|C|\mathscr{A}^{2|C|}}$, we write $\xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}$ as the (normalize) state

$$\Xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)} = \sum_{\vec{y}:\vec{y}|_C=\vec{y}_C} \mathsf{P}_{\mathbf{\Omega X}_C\mathbf{Y}}(\omega,\vec{x},\vec{y}_C)\cdot\langle\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C|\otimes\Xi_{\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C}\,,$$

In other words, the state $\Xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}$ is equivalent to the quantum-classical $\xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}$ restricted to the component where $\Omega=\omega$ and $\mathbf{R}_C=\vec{r}_C$ for all coordinates in C. We define the state $\Xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}$ as the same conditioning above, but only for fixed $\vec{x}_C,\vec{y}_C$ value, and we define the state $\Lambda^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}$ and $\Lambda^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}$ in a similar manner as above.

Likewise, for $\vec{r}_C=(\vec{x}_C,\vec{y}_C,\vec{a}_C,\vec{b}_C)\in W_C$, we define the (normalized) state $\xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}$

$$\xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)} = \frac{1}{\mathsf{P}(W_C)}\sum_{\vec{y}:\vec{y}|_C=\vec{y}_C} \mathsf{P}_{\mathbf{\Omega X}_C\mathbf{Y}}(\omega,\vec{x},\vec{y}_C)\cdot\langle\omega,\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C|\otimes\Xi_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}\,,$$

and $\xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}=0$ if $\vec{r}_C\notin W_C$. We define the state $\lambda^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}$ in a similar manner as above. By definition, for a fixed ω and $\vec{r}_C=(\vec{x}_C,\vec{y}_C,\vec{a}_C,\vec{b}_C)\in W_C$, we have $\xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}=\Xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}$. Similarly to the proof of [BVY21, Lemma 5.12], the main step of proving Lemma 12.3.15 is given by two claims which build on top of each other. The following claim is an analogue of [BVY21, Claim 5.13] for the commuting operator model. We remark that the proof is rewritten for clarity.

Claim 12.3.17.

$$\mathop{\mathbb{E}}_{i\sim[r]\setminus C}\mathop{\mathbb{E}}_{\mathbf{\Omega R}_C|W_C} I\big(\mathbf{Y}_i;\mathscr{A}\big)_{\xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}} = O(\eta_{PR})\,, \tag{12.43}$$

$$\mathop{\mathbb{E}}_{i\sim[r]\setminus C}\mathop{\mathbb{E}}_{\mathbf{\Omega R}_C|W_C} I\big(\mathbf{X}_i;\mathscr{B}\big)_{\lambda^{\mathbf{\Omega XY}_C\mathbf{Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}} = O(\eta_{PR})\,. \tag{12.44}$$

Proof. We present the proof for (12.43); the proof for (12.44) follows from a similar calculation. First, for a fixed ω and $\vec{r}_C=(\vec{x}_C,\vec{y}_C,\vec{a}_C,\vec{b}_C)\in W_C$, $\xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}=\Xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}$. Hence, by rearranging (12.43),

$$\begin{aligned}
\mathop{\mathbb{E}}_{i\sim[r]\setminus C}\mathop{\mathbb{E}}_{\mathbf{\Omega R}_C|W_C} I\big(\mathbf{Y}_i;\mathscr{A}\big)_{\xi^{\mathbf{\Omega X}_C\mathbf{Y Q}_C\mathscr{A}}_{(\omega,\vec{r}_C)}} &= \mathop{\mathbb{E}}_{i\sim[r]\setminus C}\mathop{\mathbb{E}}_{\mathbf{\Omega R}_C|W_C} \mathrm{D}\big(\xi^{\mathbf{Y}_i\mathscr{A}}_{(\omega,\vec{r}_C)}\|\Xi^{\mathbf{Y}_i}_{(\omega,\vec{r}_C)}\otimes\Xi^{\mathscr{A}}_{(\omega,\vec{r}_C)}\big)\\
&\le \mathop{\mathbb{E}}_{i\sim[r]\setminus C}\mathop{\mathbb{E}}_{\mathbf{\Omega R}_C|W_C} \mathrm{D}\big(\xi^{\mathbf{Y}\mathscr{A}}_{(\omega,\vec{r}_C)}\|\Xi^{\mathbf{Y}}_{(\omega,\vec{r}_C)}\otimes\Xi^{\mathscr{A}}_{(\omega,\vec{r}_C)}\big)\\
&\le \mathop{\mathbb{E}}_{i\sim[r]\setminus C}\mathop{\mathbb{E}}_{\mathbf{\Omega Q}_C|W_C} \mathrm{D}\big(\xi^{\mathbf{Y}\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}\|\Xi^{\mathbf{Y}}_{(\omega,\vec{r}_C)}\otimes\Xi^{\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}\big)
\end{aligned} \tag{12.45}$$

Where the second line follows from Proposition 12.1.17 and the third line follows from Proposition 12.1.21 on the distribution $\mathbf{S}_C$. We wish to use Proposition 12.1.13 to bound the inequality. For all ω and $(\vec{x}_C,\vec{y}_C)\in\mathscr{X}^{|C|^2}$

$$\begin{aligned}
\Xi^{\mathbf{Y Q}_C\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C} &= \sum_{\vec{y},\vec{a}_C,\vec{b}_C} \mathsf{P}_{\mathbf{Y}|\omega,(\vec{x}_C,\vec{y}_C)}(\vec{y})\ \langle\vec{y}|\otimes\langle\vec{a}_C\vec{b}_C|\otimes\Xi^{\mathscr{A}}_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}\\
&\le \sum_{\vec{y}}\mathsf{P}_{\mathbf{Y}|\omega}(\vec{y})\ \langle\vec{y}|\otimes\mathrm{Tr}_{|\mathscr{A}|^{2|C|}}\otimes\left(\sum_{\vec{a}_C,\vec{b}_C}\Xi^{\mathscr{A}}_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}\right)\\
&= \Xi^{\mathbf{Y}}_{\omega,\vec{x}_C,\vec{y}_C}\otimes\mathrm{Tr}_{|\mathscr{A}|^{2|C|}}\otimes\Xi^{\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}\,,
\end{aligned} \tag{12.46}$$

where the second line follows since the state $\langle\vec{y}|\le\mathrm{Tr}_{|\mathscr{A}|^{2|C|}}$, and the third line follows because $\sum_{\vec{a}_C,\vec{b}_C}\Xi^{\mathscr{A}}_{\omega,\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}(\mathbb{I}_{\mathscr{A}})=1$ (by (12.38) where both of the measurement operator A and B are POVMs). By considering the above state on the restriction of $\mathbf{M}_{|\mathscr{Y}|^r}\otimes\mathbb{I}_{|\mathscr{A}|^{2|C|}}\otimes\mathscr{A}$,

$$\Xi^{\mathbf{Y}\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}\le(|\mathscr{A}|)^{2|C|}\cdot\left(\Xi^{\mathbf{Y}}_{\omega,\vec{x}_C,\vec{y}_C}\otimes\Xi^{\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}\right).$$

Hence, by combining Proposition 12.1.13 and (12.45)

$$\begin{aligned}
\mathop{\mathbb{E}}_{i\sim[r]\setminus C}\mathop{\mathbb{E}}_{\mathbf{\Omega R}_C|W_C} I\big(\mathbf{Y}_i;\mathscr{A}\big)_{\xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}_{(\omega,\vec{r}_C)}} &\le \frac{1}{r-|C|}\left(\mathop{\mathbb{E}}_{\mathbf{\Omega Q}_C|W_C}\mathrm{D}\big(\xi^{\mathbf{Y}\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}\|\Xi^{\mathbf{Y}\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}\big)+|C|\cdot\log|\mathscr{A}|^2\right).\\
&\le \frac{1}{r-|C|}\left(\mathop{\mathbb{E}}_{\mathbf{\Omega Q}_C|W_C}\mathrm{D}\big(\xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}\|\Xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}\big)+|C|\cdot\log|\mathscr{A}|^2\right),\\
&\le \frac{1}{r-|C|}\left(\mathop{\mathbb{E}}_{\mathbf{\Omega Q}_C}\mathrm{D}\big(\xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}\|\Xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}\big)+|C|\cdot\log|\mathscr{A}|^2\right),\\
&\le \frac{1}{r-|C|}\left(\mathrm{D}\big(\xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}\|\Xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}\big)+|C|\cdot\log|\mathscr{A}|^2\right),\\
&\le \frac{1}{r-|C|}\left(\log\left(\frac{1}{P(W_C)}\right)+|C|\cdot\log|\mathscr{A}|^2\right)=\eta_{\mathrm{PR}},
\end{aligned}$$

where the first line follows from Proposition 12.1.17, the second line follows from Proposition 12.1.21 and $\xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}_{\omega,\vec{x}_C,\vec{y}_C}=0$ whenever $(\vec{x}_C,\vec{y}_C,\vec{a}_C,\vec{b}_C)\notin W_C$. The third line follows from conditioning a probability distribution will never increase the relative entropy (i.e. Proposition 12.1.13). The fifth line follows by combining Proposition 12.1.13 and (12.42). Thus showing (12.43). □

Given a fix ω_{-i} sampled from $\mathbf{\Omega}_{-i}$, $\vec{r}_C\in W_C$ and $(x,y)\in\mathscr{X}^2$, let $\omega^A_{-i,x}=(\omega_{-i},\omega_i=(A,x))$ in $\mathbf{\Omega}$. We define the (normalize) state

$$\xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}_{(\omega_{-i},\vec{r}_C),x,y}=\sum_{\vec{y}:\vec{y}|_C=\vec{y}_C,\vec{y}_i=y}\mathsf{P}_{\mathbf{\Omega X}_C\mathbf{Y}}(\omega^A_{-i,x},\vec{x},\vec{y}_C)\cdot\langle\omega^A_{-i,x},\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C|\otimes\Xi_{\omega^A_{-i,x},\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}\,. \tag{12.47}$$

Similarly, let $\omega^B_{-i,y}=(\omega_{-i},\omega_i=(B,y))$ and we define the (normalize) state

$$\lambda^{\mathbf{\Omega XY}_C\mathbf{Q}_C\mathscr{A}}_{(\omega_{-i},\vec{r}_C),x,y}=\sum_{\vec{x}:\vec{x}|_C=\vec{x}_C,\vec{x}_i=x}\mathsf{P}_{\mathbf{\Omega X}_C\mathbf{Y}}(\omega^B_{-i,y},\vec{x}_C,\vec{y})\cdot\langle\omega^B_{-i,y},\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C|\otimes\lambda_{\omega^B_{-i,y},\vec{x},\vec{y}_C,\vec{a}_C,\vec{b}_C}\,.$$

The second claim relates the states ξ and λ associated with different choices of $i,(\omega_{-i},\vec{r}_C),x,y$.

Claim 12.3.18. *The following hold:*

$$\mathop{\mathbb{E}}_{i\sim[r]\setminus C}\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\|\xi^{\mathscr{A}}_{(\omega_{-i},\vec{r}_C),x,y}-\xi^{\mathscr{A}}_{(\omega_{-i},\vec{r}_C),x,\perp}\|^2_{\mathscr{A}}=O\big(\sqrt{\eta_{PR}}\big)\,, \tag{12.48}$$

$$\mathop{\mathbb{E}}_{i\sim[r]\setminus C}\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\|\lambda^{\mathscr{A}}_{(\omega_{-i},\vec{r}_C),x,y}-\lambda^{\mathscr{A}}_{(\omega_{-i},\vec{r}_C),\perp,y}\|^2_{\mathscr{A}}=O\big(\sqrt{\eta_{PR}}\big)\,, \tag{12.49}$$

where the expectation over XY is with respect to the distribution μ_{XY}.

Proof. We show (12.48); the proof of (12.49) is similar. Fix $i\in[r]\setminus C$ and $\vec{r}_C\in W_C$, the state

$$\xi^{\mathbf{Y}_i\mathscr{A}}_{(\omega,\vec{r}_C)}=\mathop{\mathbb{E}}_{\mathbf{Y}_i|\mathbf{R}_C=\vec{r}_C,W_C}\langle\vec{y}_i|^{\mathbf{Y}_i}\otimes\xi^{\mathscr{A}}_{(\omega,\vec{r}_C),\vec{y}_i}=\xi^{\mathbf{Y}_i}_{(\omega,\vec{r}_C)}\otimes\xi^{\mathscr{A}}_{(\omega,\vec{r}_C)}$$

where the state $\xi^{\mathscr{A}}_{(\omega,\vec{r}_C),\vec{y}_i}$ is $\xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}_{(\omega,\vec{r}_C)}$ conditioning on the ith coordinates of $\vec{y}$ being $\vec{y}_I$. By applying Proposition 12.1.12

$$\begin{aligned}
\mathop{\mathbb{E}}_{i\sim[r]\setminus C}\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|W_C}\|\xi^{\mathscr{A}}_{(\omega,\vec{r}_C),\vec{y}_i}-\xi^{\mathscr{A}}_{(\omega,\vec{r}_C)}\|^2_{\mathscr{A}}&\le 2\ln 2\mathop{\mathbb{E}}_{i\sim[r]\setminus C}\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|W_C}\mathrm{D}\big(\xi^{\mathscr{A}}_{(\omega,\vec{r}_C),\vec{y}_i}\,\|\,\xi^{\mathscr{A}}_{(\omega,\vec{r}_C)}\big)\\
&=2\ln 2\mathop{\mathbb{E}}_{I}\mathop{\mathbb{E}}_{\mathbf{R}|W_C}I(\mathbf{Y}_i;\mathscr{A})_{\xi^{\mathbf{\Omega X}_C\mathbf{YQ}_C\mathscr{A}}_{(\omega,\vec{r}_C)}}\\
&=O(\eta_{\mathrm{PR}})\,,
\end{aligned} \tag{12.50}$$

where the last line follows from Claim 12.3.17. The rest of the proof follows in an identical manner to that of [BVY21, Claim 5.14]. □

We are now ready to give the proof of Lemma 12.3.15.

Proof of Lemma 12.3.15. We start by showing the existence of operators $V_{(\omega_{-i},\vec{r}_C),y}$ that satisfy (12.33). Let $i \in [r] \setminus C$, $(\omega_{-i}) \in \mathbf{\Omega}_{-i}$, $\vec{r}_C \in W_C$ and $(x,y) \in \mathscr{X}^2$. We start the proof by showing the following claim.

Claim 12.3.19. *For all $A \in \mathscr{A}$, we have*

$$\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y}(A) = \xi^{\mathscr{A}}_{(\omega_{-i},\vec{r}_C),x,y}(A)$$
$$\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,y}(A) = \xi^{\mathscr{A}}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,y}(A)$$

where $\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y}$ is the state defined by

$$\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y}(A) = \langle \widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y} | A | \widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y} \rangle$$

with $|\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y}\rangle$ being defined on (12.26).

Proof. Recall from (12.47), we can write the state $\xi^{\mathbf{\Omega}\mathbf{X}_C\mathbf{Y}\mathbf{Q}_C\mathscr{A}}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y}$ explicitly as

$$\xi^{\mathbf{\Omega}\mathbf{X}_C\mathbf{Y}\mathbf{Q}_C\mathscr{A}}_{(\omega_{-i},\vec{r}_C),x,y} = \sum_{\vec{y}:\vec{y}|_C=\vec{y}_C,\vec{y}_i=y} \frac{\mathsf{P}_{\mathbf{\Omega}\mathbf{X}_C\mathbf{Y}}(\omega^A_{-i,x},\vec{x},\vec{y}_C)}{\mathsf{P}_{\mathbf{\Omega}\mathbf{R}_C\mathbf{Y}_i}(\omega^A_{-i,x},\vec{r}_C,y)} \cdot \langle \omega^A_{-i,x},\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C| \otimes \Xi_{\omega^A_{-i,x},\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C} \,. \tag{12.51}$$

To see that the normalization is correct, we evaluate this state on the identity $\mathbb{I}_{\mathscr{A}}$ on the state $\xi^{\mathscr{A}}_{(\omega_{-i},\vec{r}_C),x,y}$ to get

$$\begin{aligned}
\xi^{\mathscr{A}}_{(\omega_{-i},\vec{r}_C),x,y}(\mathbb{I}_{\mathscr{A}}) &= \sum_{\vec{y}:\vec{y}|_C=\vec{y}_C,\vec{y}_i=y} \frac{\mathsf{P}_{\mathbf{\Omega}\mathbf{X}_C\mathbf{Y}}(\omega^A_{-i,x},\vec{x},\vec{y}_C)}{\mathsf{P}_{\mathbf{\Omega}\mathbf{R}_C\mathbf{Y}_i}(\omega^A_{-i,x},\vec{r}_C,y)} \Xi_{\omega^A_{-i,x},\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}(\mathbb{I}_{\mathscr{A}}) \,. \\
&= \sum_{\vec{y}:\vec{y}|_C=\vec{y}_C,\vec{y}_i=y} \frac{\mathsf{P}_{\mathbf{\Omega}\mathbf{X}_C\mathbf{Y}}(\omega^A_{-i,x},\vec{x},\vec{y}_C)}{\mathsf{P}_{\mathbf{\Omega}\mathbf{R}_C\mathbf{Y}_i}(\omega^A_{-i,x},\vec{r}_C,y)} \cdot \langle\psi| A^{\omega^A_{-i,x},(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C} B^{\vec{y}}_{\vec{b}_C} |\psi\rangle \\
&= \frac{1}{\mathsf{P}_{\mathbf{\Omega}\mathbf{R}_C\mathbf{Y}_i}(\omega^A_{-i,x},\vec{r}_C,y)} \sum_{\substack{\vec{x}:\vec{x}|_C=\vec{x}_C \\ \vec{y}:\vec{y}|_C=\vec{y}_C,\vec{y}_i=y}} \mathsf{P}_{\mathbf{\Omega}\mathbf{X}\mathbf{Y}\mathbf{A}_C\mathbf{B}_C}(\omega^A_{-i,x},\vec{x},\vec{y},\vec{a}_C,\vec{b}_C) \\
&= 1 \,.
\end{aligned}$$

Now we compute the restriction of $\xi^{\mathscr{A}}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y}$ to the subalgebra $\mathscr{A}$. For all $M \in \mathscr{A}$

$$\begin{aligned}
&\xi^{\mathscr{A}}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y}(M) = \sum_{\vec{y}:\vec{y}|_C=\vec{y}_C,\vec{y}_i=y} \frac{\mathsf{P}_{\mathbf{\Omega}\mathbf{X}_C\mathbf{Y}}(\omega^A_{-i,x},\vec{x},\vec{y}_C)}{\mathsf{P}_{\mathbf{\Omega}\mathbf{R}_C\mathbf{Y}_i}(\omega^A_{-i,x},\vec{r}_C,y)} \Xi_{\omega^A_{-i,x},\vec{x}_C,\vec{y},\vec{a}_C,\vec{b}_C}(M) \,. \\
&= \sum_{\vec{y}:\vec{y}|_C=\vec{y}_C,\vec{y}_i=y} \frac{\mathsf{P}_{\mathbf{\Omega}\mathbf{X}_C\mathbf{Y}}(\omega^A_{-i,x},\vec{x},\vec{y}_C)}{\mathsf{P}_{\mathbf{\Omega}\mathbf{R}_C\mathbf{Y}_i}(\omega^A_{-i,x},\vec{r}_C,y)} \cdot \langle\psi| \left(A^{\omega^A_{-i,x},(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} M \left(A^{\omega^A_{-i,x},(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} B^{\vec{y}}_{\vec{b}_C} |\psi\rangle \\
&= \frac{\mathsf{P}_{\mathbf{\Omega}\mathbf{X}_C\mathbf{Y}_C,\mathbf{Y}_i}(\omega^A_{-i,x},\vec{x}_C,\vec{y}_C,y)}{\mathsf{P}_{\mathbf{\Omega}\mathbf{R}_C\mathbf{Y}_i}(\omega^A_{-i,x},\vec{r}_C,y)} \cdot \\
&\quad \langle\psi| \left(A^{\omega^A_{-i,x},(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} M \left(A^{\omega^A_{-i,x},(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} \left(\sum_{\vec{y}:\vec{y}|_C=\vec{y}_C,\vec{y}_i=y} P_{\mathbf{Y}|\Omega=\omega_{-i}}(\vec{y}) \cdot B^{\vec{y}}_{\vec{b}_C}\right) |\psi\rangle \\
&= \frac{\langle\psi| \left(A^{\omega^A_{-i,x},(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} M \left(A^{\omega^A_{-i,x},(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} \cdot B^{(\omega_{-i},\vec{y}_C),y}_{\vec{a}_C} | \,|\psi\rangle}{\mathsf{P}_{\mathbf{A}_C\mathbf{B}_C|\mathbf{\Omega}=\omega^A_{-i,x},\mathbf{X}_C=\vec{x}_C,\mathbf{Y}_C=\vec{y}_C,\mathbf{Y}_i=y}(\vec{a}_C,\vec{b}_C)} \\
&= \gamma^{-2}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y} \langle\psi| \left(A^{\omega^A_{-i,x},(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} M \left(A^{\omega^A_{-i,x},(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} B^{(\omega_{-i},\vec{y}_C),y}_{\vec{a}_C} |\psi\rangle
\end{aligned}$$

$$= \gamma^{-2}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y} \cdot \langle\psi| \left(A^{\omega^A_{-i,x},(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} \left(B^{(\omega_{-i},\vec{y}_C),y}_{\vec{a}_C}\right)^{1/2} M \left(\left(A^{\omega^A_{-i,x},(\vec{x}_C,\vec{y}_C)}_{\vec{a}_C}\right)^{1/2} B^{(\omega_{-i},\vec{y}_C),y}_{\vec{a}_C}\right)^{1/2} |\psi\rangle\,, \tag{12.52}$$

where line 3 follows from $\mathbf{Y}_i$ being independent from all other $\mathbf{Y}_j$ and $\mathbf{\Omega}_j$ for $j \neq i$ and Claim 12.3.8, line 4 follows from the definition of $B^{(\omega_{-i},\vec{y}_C),y}$ from (12.19), line 5 follows from Proposition 12.3.10, and the last line follows from $B^{(\omega_{-i},\vec{y}_C),y}_{\vec{a}_C} \in \mathscr{A}'$. We see that the quantity given in (12.52) are precisely the definition to $\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y}(M)$ given in (12.26).

For the second part of the claim, whenever $x = \perp$, $A^{\omega^A_{-i,x}}_{\vec{a}_C}$ are the same as $A^{(\omega_{-i},\vec{x}_C),x}_{\vec{a}_C}$ given in (12.19). Hence the second part of the claim holds by (12.27). This concludes the proof of Claim 12.3.19. □

By combining Claim 12.3.19 and Claim 12.3.18, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{i\sim[r]\setminus C} \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|W_C} \mathop{\mathbb{E}}_{\mathbf{XY}} \|\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y} - \widetilde{\Phi}_{(\omega_{-i},\vec{r}_C),x,\perp}\|^2_{\mathscr{A}} &= O\big(\sqrt{\eta_{\mathrm{PR}}}\big) \\
\mathop{\mathbb{E}}_{i\sim[r]\setminus C} \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|W_C} \mathop{\mathbb{E}}_{\mathbf{XY}} \|\xi^{\mathscr{A}}_{(\omega_{-i},\vec{r}_C),x,y} - \xi^{\mathscr{A}}_{(\omega_{-i},\vec{r}_C),x,\perp}\|^2_{\mathscr{A}} &= O\big(\sqrt{\eta_{\mathrm{PR}}}\big).
\end{aligned}$$

By conditioning $\mathscr{X} = \perp$ on the third expectation (which occurs with probability $\eta_{\mathrm{Anchor}} = \frac{1}{2}$),

$$\mathop{\mathbb{E}}_{i\sim[r]\setminus C} \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|W_C} \mathop{\mathbb{E}}_{Y} \|\widetilde{\Phi}_{(\omega_{-i},\vec{r}_C),\perp,y} - \widetilde{\Phi}_{(\omega_{-i},\vec{r}_C),\perp,\perp}\|^2_{\mathscr{A}} = O\big(\sqrt{\eta_{\mathrm{PR}}}\big).$$

Now, by applying Proposition 12.1.7, there exists a collection of Unitary operator $\{V_{(\omega_{-i},\vec{r}_C),y}\}_{y\in\mathscr{Y}}$ in $\mathscr{A}'$ such that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{i\sim[r]\setminus C} \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|W_C} \mathop{\mathbb{E}}_{Y} \big\langle \widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,y} \big| V_{(\omega_{-i},\vec{r}_C),y} \big| \widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,\perp} \big\rangle \\
&\geq 1 - \frac{1}{2} \mathop{\mathbb{E}}_{i\sim[r]\setminus C} \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|W_C} \mathop{\mathbb{E}}_{Y} \mathop{\mathbb{E}}_{Y} \|\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,y} - \widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,\perp}\|_{\mathscr{A}} \\
&\geq 1 - O(\eta^{1/4}_{\mathrm{PR}})\,,
\end{aligned}$$

where the second line follows from Jensen's inequality. By translating the above equation to Euclidean distance, and then apply Jensen's inequality, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{i\sim[r]\setminus C} \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|W_C} \mathop{\mathbb{E}}_{Y} \| |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,y}\rangle - V_{(\omega_{-i},\vec{r}_C),y} |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,\perp}\rangle \| \\
&\leq \sqrt{\mathop{\mathbb{E}}_{i\sim[r]\setminus C} \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|W_C} \mathop{\mathbb{E}}_{Y} \| |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,y}\rangle - V_{(\omega_{-i},\vec{r}_C),y} |\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp,\perp}\rangle \|^2} = O\big(\eta^{1/8}_{\mathrm{PR}}\big)\,.
\end{aligned}$$

Applying Markov's inequality over the index i establishes (12.33). The argument for (12.34) proceeds similarly. We start by using Claim 12.3.19, which establish that the states $|\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y}\rangle$ and $|\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,\perp}\rangle$ are purifications of the states $\xi^{\mathscr{A}}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,y}$ and $\xi^{\mathscr{A}}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),\perp/x,\perp}$ respectively. Using Uhlmann's Theorem and Claim 12.3.18 in a similar way to how we derived (12.33) we deduce the existence of operators $V_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y}$ satisfying (12.34). To prove (12.32) we use the following claim whose proof is analogous to that of Claim 12.3.19.

Claim 12.3.20. *For all $A \in \mathscr{A}'$, we have*

$$\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,\perp}(A) = \lambda_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,\perp}(A)$$

where $|\widetilde{\Phi}_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y}\rangle$ is the vector state defined in (12.26).

Using the above claim, we can use Uhlmann's Theorem in a similar manner as above to deduce the existence of operators $U_{(\omega_{-i},\vec{r}_C),x}$ in $\mathscr{A}$ which satisfy (12.32). This concludes the proof for Lemma 12.3.15. □

12.3.6 Proof of Proposition 12.3.11

We end the section with the proof of Proposition 12.3.11. We remark that this subsection follows the structure of [BVY21, Section 5.4].

Proof. For every $i, (\omega_{-i}, \vec{r}_C)$, x and y let $U_{(\omega_{-i},\vec{r}_C),x}$, $V_{(\omega_{-i},\vec{r}_C),y}$ and $V_{(\boldsymbol{\omega}_{-i},\vec{r}_C),x,y}$ denote the unitary guarantee by Lemma 12.3.15. For notational convenience we suppress the dependence on $(i, (\omega_{-i}, \vec{r}_C))$; thus the operators $U_x, V_y, V_{x,y}$, the states $|\Phi_{x,y}\rangle$, and their normalizations $\gamma_{x,y}$ all implicitly depend on i and $(\omega_{-i}, \vec{r}_C = (\vec{x}_C, \vec{y}_C, \vec{a}_C, \vec{b}_C))$. We also write $\mathbb{E}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}$ as shorthand for $\mathbb{E}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|\mathbf{X}_i=\perp,\mathbf{Y}_i=\perp,W_C}$

For a fixed index $i \in [r] \setminus C$, we call the index to be *good* if it satisfies (i) the conclusions of Lemma 12.3.15, (ii) the conclusions of Lemma 12.3.16, and (iii) it holds that

$$\left\| \mathsf{P}_{\boldsymbol{\Omega}_{-i},\mathbf{R}_C|\mathbf{X}_i=\perp,\mathbf{Y}_i=\perp,W_C} - \mathsf{P}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|W_C} \right\| \le O(\eta_{\mathrm{PR}}^{1/4}) \,. \tag{12.53}$$

By applying the data processing inequality (Lemma 12.3.2) to Item 3 of Lemma 12.3.9 to marginalize over the random variable $\boldsymbol{\Omega}_i$, and then applying Markov's inequality over the index i, we get that (12.53) holds with probability at least $1 - O(\eta_{\mathrm{PR}}^{1/4})$ over a uniformly random choice of i. This combined with Lemma 12.3.15 and Lemma 12.3.16 implies that an index i is good with probability at least $1 - O(\eta_{\mathrm{PR}}^{1/16})$. For a good index i, by combining the bound from Lemma 12.3.15 and (12.53),

$$\mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C} \mathop{\mathbb{E}}_{\mathbf{X}} \| |\widetilde{\Phi}_{x,\perp}\rangle - U_x |\widetilde{\Phi}_{\perp,\perp}\rangle \| = O(\eta_{\mathrm{PR}}^{1/16}) \,, \tag{12.54}$$

$$\mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C} \mathop{\mathbb{E}}_{\mathbf{Y}} \| V_y |\widetilde{\Phi}_{\perp,\perp}\rangle - |\widetilde{\Phi}_{\perp,y}\rangle \| = O(\eta_{\mathrm{PR}}^{1/16}) \,, \tag{12.55}$$

$$\mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C} \mathop{\mathbb{E}}_{\mathbf{XY}} \| V_{x,y} |\widetilde{\Phi}_{\perp/x,y}\rangle - |\widetilde{\Phi}_{\perp/x,\perp}\rangle \| = O(\eta_{\mathrm{PR}}^{1/16}) \tag{12.56}$$

where we bound $O(\eta_{\mathrm{PR}}^{1/16}) + O(\eta_{\mathrm{PR}}^{1/4}) = O(\eta_{\mathrm{PR}}^{1/16})$. he main step of the proof of Proposition 12.3.11 is to combine U_x and V_y together by showing the following claim. We remark that the following claim is an commuting operator value variant of [BVY21, Claim 5.19].

Claim 12.3.21.

$$\begin{aligned}
\mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C} \mathop{\mathbb{E}}_{\mathbf{XY}} \| U_x V_y |\widetilde{\Phi}_{\perp,\perp}\rangle - |\widetilde{\Phi}_{x,y}\rangle \| \le & \mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C} \mathop{\mathbb{E}}_{\mathbf{XY}} \gamma_{\perp,\perp}^{-1} \left\| V_y |\Phi_{\perp,\perp}\rangle - |\Phi_{\perp,y}\rangle \right\| && (12.57)\\
& + 2\eta_{Anchor}^{-1/2} \gamma_{\perp,\perp}^{-1} \| V_{x,y} |\Phi_{\perp/x,y}\rangle - |\Phi_{\perp/x,\perp}\rangle \| && (12.58)\\
& + \gamma_{\perp,\perp}^{-1} \left\| U_x |\Phi_{\perp,\perp}\rangle - |\Phi_{x,\perp}\rangle \right\| + O(\eta_{PR}^{1/8}) \,, && (12.59)
\end{aligned}$$

where $\gamma_{\perp,\perp}$ is defined in (12.25).

Before proving the claim, it would be useful to work out the following two bounds. Using the definition

$$\mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C} \mathop{\mathbb{E}}_{\mathbf{XY}} \| |\widetilde{\Phi}_{x,y}\rangle - \gamma_{\perp,\perp}^{-1} |\Phi_{x,y}\rangle \| = \mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C} \mathop{\mathbb{E}}_{\mathbf{XY}} \left| 1 - \frac{\gamma_{x,y}}{\gamma_{\perp,\perp}} \right| = O(\eta_{\mathrm{PR}}^{1/8}) \,, \tag{12.60}$$

where the second line is by Jensen's inequality and (12.35) in Lemma 12.3.16. Similarly,

$$\mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C} \mathop{\mathbb{E}}_{\mathbf{XY}} \| |\widetilde{\Phi}_{\perp/x,y}\rangle - \gamma_{\perp,\perp}^{-1} |\Phi_{\perp/x,y}\rangle \| = \mathop{\mathbb{E}}_{\boldsymbol{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C} \mathop{\mathbb{E}}_{\mathbf{XY}} \left| 1 - \frac{\gamma_{\perp/x,y}}{\gamma_{\perp,\perp}} \right| = O\big(\eta_{\mathrm{PR}}^{1/8}\big) \,, \tag{12.61}$$

by (12.36). We note that in the above division by $\gamma_{\perp,\perp}$ is well-defined because $(\omega_{-i}, \vec{r}_C)$ is sampled with positive probability from the distribution $\mathsf{P}_{(\omega_{-i},\vec{r}_C)|\perp,\perp,W_C}$. We are now ready to prove Claim 12.3.21.

Proof. We start by writing

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\|U_x\,V_y\,|\widetilde{\Phi}_{\perp,\perp}\rangle-|\widetilde{\Phi}_{x,y}\rangle\,\| \\
&\leq \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\left\|U_x\,V_y\big|\widetilde{\Phi}_{\perp,\perp}\big\rangle-\frac{\gamma_{x,y}}{\gamma_{\perp,\perp}}\big|\widetilde{\Phi}_{x,y}\big\rangle\right\|+\left\|\frac{\gamma_{x,y}}{\gamma_{\perp,\perp}}\big|\widetilde{\Phi}_{x,y}\big\rangle-\big|\widetilde{\Phi}_{x,y}\big\rangle\right\| \\
&= \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\gamma_{\perp,\perp}^{-1}\left\|U_x\,V_y\big|\Phi_{\perp,\perp}\big\rangle-\big|\Phi_{x,y}\big\rangle\right\|+\left|\frac{\gamma_{x,y}}{\gamma_{\perp,\perp}}-1\right| \\
&\leq \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\gamma_{\perp,\perp}^{-1}\left\|U_x\,V_y\big|\Phi_{\perp,\perp}\big\rangle-\big|\Phi_{x,y}\big\rangle\right\|+O(\eta_{\mathrm{PR}}^{1/8}) \\
&\leq \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\gamma_{\perp,\perp}^{-1}\left(\left\|U_x\,V_y\big|\Phi_{\perp,\perp}\big\rangle-U_x\big|\Phi_{\perp,y}\big\rangle\right\|+\left\|U_x\big|\Phi_{\perp,y}\big\rangle-\big|\Phi_{x,y}\big\rangle\right\|\right)+O(\eta_{\mathrm{PR}}^{1/8}) \\
&\leq \mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\gamma_{\perp,\perp}^{-1}\left(\left\|\,V_y\big|\Phi_{\perp,\perp}\big\rangle-\big|\Phi_{\perp,y}\big\rangle\right\|+\left\|U_x\big|\Phi_{\perp,y}\big\rangle-\big|\Phi_{x,y}\big\rangle\right\|\right)+O(\eta_{\mathrm{PR}}^{1/8})\,, \qquad (12.62)
\end{aligned}$$

where the third line follows from (12.60), and the last line follows from $U_x \in \mathscr{U}(\mathscr{A})$. For all $\perp/x \in \mathscr{X}_\perp$, by Lemma 3.2.1 and (12.22) and (12.23). There exist an operator $(C^{\perp/x})^{\frac{1}{2}}$ and $(D^{\perp/x})^{\frac{1}{2}}$ such that

$$\begin{aligned}
\eta_{\mathrm{Anchor}}^{-\frac{1}{2}}(C^{\perp/x})^{\frac{1}{2}}(A^{\perp/x})^{\frac{1}{2}} &= (A^{\perp})^{\frac{1}{2}} \\
(1-\eta_{\mathrm{Anchor}})^{-\frac{1}{2}}(D^{\perp/x,x})^{\frac{1}{2}}(A^{\perp/x})^{\frac{1}{2}} &= (A^{x})^{\frac{1}{2}}.
\end{aligned}$$

By (12.24), we have

$$\begin{aligned}
|\Phi_{\perp,y}\rangle &= (B^y)^{\frac{1}{2}}(A^{\perp})^{\frac{1}{2}}\,|\psi\rangle = \eta_{\mathrm{Anchor}}^{-\frac{1}{2}}(C^{\perp/x})^{\frac{1}{2}}(B^y)^{\frac{1}{2}}(A^{\perp/x})^{\frac{1}{2}}\,|\psi\rangle = (1-\eta_{\mathrm{Anchor}})^{-\frac{1}{2}}(C^{\perp/x})^{\frac{1}{2}}\,|\Phi_{\perp/x,y}\rangle \\
|\Phi_{x,y}\rangle &= (1-\eta_{\mathrm{Anchor}})^{-\frac{1}{2}}(D^{\perp/x,x})^{\frac{1}{2}}\,|\Phi_{\perp/x,y}\rangle
\end{aligned}$$

Thus, for each $x, y \in \mathscr{X}^2$,

$$\begin{aligned}
\left\|U_x\,|\Phi_{\perp,y}\rangle-|\Phi_{x,y}\rangle\right\| &= \|\eta_{\mathrm{Anchor}}^{-\frac{1}{2}}U_x(C^{\perp/x})^{\frac{1}{2}}\,|\Phi_{\perp/x,y}\rangle-(1-\eta_{\mathrm{Anchor}})^{-\frac{1}{2}}(D^{\perp/x,x})^{\frac{1}{2}}\,|\Phi_{\perp/x,y}\rangle\,\| \\
&= \|\eta_{\mathrm{Anchor}}^{-\frac{1}{2}}U_x(C^{\perp/x})^{\frac{1}{2}}\,V_{x,y}\,|\Phi_{\perp/x,y}\rangle-(1-\eta_{\mathrm{Anchor}})^{-\frac{1}{2}}(D^{\perp/x,x})^{\frac{1}{2}}\,V_{x,y}\,|\Phi_{\perp/x,y}\rangle\,\| \\
&\leq \eta_{\mathrm{Anchor}}^{-\frac{1}{2}}\|U_x(C^{\perp/x})^{\frac{1}{2}}\,V_{x,y}\,|\Phi_{\perp/x,y}\rangle-U_x(C^{\perp/x})^{\frac{1}{2}}\,|\Phi_{\perp/x,\perp}\rangle\,\| \qquad (12.63) \\
&\quad+\|\eta_{\mathrm{Anchor}}^{-\frac{1}{2}}U_x(C^{\perp/x})^{\frac{1}{2}}\,|\Phi_{\perp/x,\perp}\rangle-(1-\eta_{\mathrm{Anchor}})^{-\frac{1}{2}}(D^{\perp/x,x})^{\frac{1}{2}}\,|\Phi_{\perp/x,\perp}\rangle\,\| \qquad (12.64) \\
&\quad+(1-\eta_{\mathrm{Anchor}})^{-\frac{1}{2}}\|(D^{\perp/x,x})^{\frac{1}{2}}\,|\Phi_{\perp/x,\perp}\rangle-(D^{\perp/x,x})^{\frac{1}{2}}\,V_{x,y}\,|\Phi_{\perp/x,y}\rangle\,\|\,, \qquad (12.65)
\end{aligned}$$

where the second line follows from $V_{x,y} \in \mathscr{U}(\mathscr{A}')$, and the third line follows from the triangle inequality. We bound each of these three terms as follows. For (12.63)

$$\eta_{\mathrm{Anchor}}^{-\frac{1}{2}}\|U_x(C^{\perp/x})^{\frac{1}{2}}\,V_{x,y}\,|\Phi_{\perp/x,y}\rangle-U_x(C^{\perp/x})^{\frac{1}{2}}\,|\Phi_{\perp/x,\perp}\rangle\,\| \leq \|\,V_{x,y}\,|\Phi_{\perp/x,y}\rangle-|\Phi_{\perp/x,\perp}\rangle\,\|.$$

Similarly, for (12.65), since $(1-\eta_{\mathrm{Anchor}})^{-\frac{1}{2}} \leq \eta_{\mathrm{Anchor}}^{-\frac{1}{2}}$ whenever $\eta_{\mathrm{Anchor}} \leq \frac{1}{2}$ ($\eta_{\mathrm{Anchor}} = \frac{1}{4}$)

$$(1-\eta_{\mathrm{Anchor}})^{-\frac{1}{2}}\|(D^{\perp/x,x})^{\frac{1}{2}}\,|\Phi_{\perp/x,\perp}\rangle-(D^{\perp/x,x})^{\frac{1}{2}}\,V_{x,y}\,|\Phi_{\perp/x,y}\rangle\,\| \leq \eta_{\mathrm{Anchor}}^{-\frac{1}{2}}\|\,|\Phi_{\perp/x,\perp}\rangle-\,V_{x,y}\,|\Phi_{\perp/x,y}\rangle\,\|$$

Finally, for (12.64)

$$\|\eta_{\mathrm{Anchor}}^{-\frac{1}{2}}U_x(C^{\perp/x})^{\frac{1}{2}}\,|\Phi_{\perp/x,\perp}\rangle-(1-\eta_{\mathrm{Anchor}})^{-\frac{1}{2}}(D^{\perp/x,x})^{\frac{1}{2}}\,|\Phi_{\perp/x,\perp}\rangle\,\| = \|U_x\,|\Phi_{\perp/x,\perp}\rangle-|\Phi_{x,\perp}\rangle\,\|$$

Putting the three bounds together, from (12.63)–(12.65) we get

$$\left\|U_x\,|\Phi_{\perp,y}\rangle-|\Phi_{x,y}\rangle\right\| \leq 2\eta_{\mathrm{anchor}}^{-1/2}\|\,V_{x,y}\,|\Phi\rangle_{\perp/x,y}-|\Phi_{\perp/x,\perp}\rangle\,\|+\left\|U_x\,|\Phi_{\perp,\perp}\rangle-|\Phi_{x,\perp}\rangle\right\|\,, \qquad (12.66)$$

from which inserting it into (12.62) proves the claim. □

To conclude the proof Proposition 12.3.11 it remains to bound each of the three terms on the right-hand side of Claim 12.3.21 by $O(\eta_{\mathrm{PR}}^{1/16})$, and then use (12.53) to exchange the expectation $\mathbb{E}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}$ with $\mathbb{E}_{(\mathbf{\Omega}_{-i},\mathbf{R}_C)|W_C}$ by introducing an additive $O(\eta_{\mathrm{PR}}^{1/4})$ error. We start with bounding (12.57):

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{Y}}\gamma_{\perp,\perp}^{-1}\left\|V_y|\Phi_{\perp,\perp}\rangle-|\Phi_{\perp,y}\rangle\right\| \\
&\qquad=\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{Y}}\left\|V_y|\widetilde{\Phi}_{\perp,\perp}\rangle-\frac{\gamma_{\perp,y}}{\gamma_{\perp,\perp}}|\widetilde{\Phi}_{\perp,y}\rangle\right\| \\
&\qquad\le\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{Y}}\|V_y|\widetilde{\Phi}_{\perp,\perp}\rangle-|\widetilde{\Phi}_{\perp,y}\rangle\|+\left\||\widetilde{\Phi}_{\perp,y}\rangle-\frac{\gamma_{\perp,y}}{\gamma_{\perp,\perp}}|\widetilde{\Phi}_{\perp,y}\rangle\right\| \\
&\qquad=O(\eta_{\mathrm{PR}}^{1/16})+\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{Y}}\left|1-\frac{\gamma_{\perp,y}}{\gamma_{\perp,\perp}}\right| \\
&\qquad=O(\eta_{\mathrm{PR}}^{1/16})+O(\eta_{\mathrm{PR}}^{1/8})=O(\eta_{\mathrm{PR}}^{1/16})\,,
\end{aligned}$$

where the third line uses (12.54) to bound the first term and the last line follows from (12.60) and conditioning on $X=\perp$ (which occurs with $\eta_{Anchor}=\frac{1}{2}$ probability), which occurs with probability $\frac{1}{2}$. We bound (12.59) in an analogous fashion. Finally, we bound (12.58) as follows:

$$\begin{aligned}
&2\eta_{\mathrm{anchor}}^{-1/2}\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\gamma_{\perp,\perp}^{-1}\left\|V_{x,y}|\Phi_{\perp/x,y}\rangle-|\Phi_{\perp/x,\perp}\rangle\right\| \\
&=2\eta_{\mathrm{anchor}}^{-1/2}\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\left\|\frac{\gamma_{\perp/x,y}}{\gamma_{\perp,\perp}}V_{x,y}|\widetilde{\Phi}_{\perp/x,y}\rangle-\frac{\gamma_{\perp/x,\perp}}{\gamma_{\perp,\perp}}|\widetilde{\Phi}_{\perp/x,\perp}\rangle\right\| \\
&\le 2\eta_{\mathrm{anchor}}^{-1/2}\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\left\|\frac{\gamma_{\perp/x,y}}{\gamma_{\perp,\perp}}|\widetilde{\Phi}_{\perp/x,y}\rangle-|\widetilde{\Phi}_{\perp/x,y}\rangle\right\|+\left\|V_{x,y}|\widetilde{\Phi}_{\perp/x,y}\rangle-|\widetilde{\Phi}_{\perp/x,\perp}\rangle\right\| \\
&\qquad\qquad+\left\||\widetilde{\Phi}_{\perp/x,\perp}\rangle-\frac{\gamma_{\perp/x,\perp}}{\gamma_{\perp,\perp}}|\widetilde{\Phi}_{\perp/x,\perp}\rangle\right\| \\
&=2\eta_{\mathrm{anchor}}^{-1/2}\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\left|1-\frac{\gamma_{\perp/x,y}}{\gamma_{\perp,\perp}}\right|+O(\eta_{\mathrm{PR}}^{1/16}/\eta_{\mathrm{anchor}}^{1/2})+2\eta_{\mathrm{anchor}}^{-1/2}\mathop{\mathbb{E}}_{\mathbf{\Omega}_{-i}\mathbf{R}_C|\perp,\perp,W_C}\mathop{\mathbb{E}}_{\mathbf{XY}}\left|1-\frac{\gamma_{\perp/x,\perp}}{\gamma_{\perp,\perp}}\right| \\
&=O(\eta_{\mathrm{PR}}^{1/8})+O(\eta_{\mathrm{PR}}^{1/16})+O(\eta_{\mathrm{PR}}^{1/8})=O(\eta_{\mathrm{PR}}^{1/16})\,.
\end{aligned}$$

The last line follows from $\eta_{\mathrm{anchor}}=\frac{1}{4}$, (12.61) to bound the first term, (12.56) to bound the second term, and (12.61) along with conditioning on $Y=\perp$ to bound the last term. □

12.4 Proof of Proposition 7.4.4

After proving the anchored parallel repetition theorem above, we are now ready to give a proof for Proposition 7.4.4. For convenience, we restate the proposition here.

Proposition 12.4.1 (Parallel repetition)**.** *For all constants $\alpha\in\mathbb{N}$, function $\mathbf{s}(n):\mathbb{N}\to[0,1]$ with $O(\mathbf{s}(n))=O(\mathrm{polylog}(n))$. Then there exists a function $\mathbf{r}(n):\mathbb{N}\to\mathbb{N}$ with $\mathbf{r}(n)=O(\alpha,\mathbf{s}(n))$ and a polynomial time algorithm $\mathtt{Parallelrep}_{\alpha,\mathbf{s}(n)}$ that takes, as input, a pair of Turing machines $(\mathtt{Q},\mathtt{D})$ and outputs a tuple of Turing machine machines $(\mathtt{Q}^{Pararep},\mathtt{D}^{Pararep})$ with*

$$\mathsf{TIME}_{\mathtt{Parallelrep}_{\alpha,\mathbf{s}(n)}}(|\mathtt{Q}|,|\mathtt{D}|)=O(\mathrm{polylog}(n))$$

such that the following holds:

1. *(Independency) $\mathtt{Q}^{Pararep}$ only dependent on $\mathtt{Q}$ and the polynomial functions $\mathbf{s}(n)$.*
2. *(Complexity bounds for the output):*
   - $\mathsf{TIME}_{\mathtt{Q}^{Pararep}}(n)\le O(\mathrm{poly}(\mathbf{r}(n),log^{\alpha}(n)))$,
   - $\mathsf{TIME}_{\mathtt{D}^{Pararep}}(n)\le O(\mathrm{poly}(\mathbf{r}(n),log^{\alpha}(n)))$,

Furthermore, if the input $\mathscr{V} = (\mathtt{Q}, \mathtt{D})$ is a k level synchronous CL verifier for the infinite sequence of synchronous game $\mathscr{G} = \{\mathscr{G}_n\}_{n\in\mathbb{N}}$ for some constant $k \in \mathbb{N}$, and there exists some constant $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, we have

$$\mathtt{Q}(n, \textit{Parameter}), \mathsf{TIME}_{\mathtt{Q}}(n), \mathsf{TIME}_{\mathtt{D}} \leq \log^{\alpha}(n).$$

Then for all $n \geq n_0$

1. *(Parameter) $\mathtt{Q}^{AR}$ is a level $k+1$ CL sampler.*
2. *(Completeness) If there exists a perfect oracularizable strategy for $\mathscr{G}_n$ in model t, then there exists a perfect oracularizable strategy for $\mathscr{G}_n^{Pararep}$ in model t.*
3. *(Soundness)*
$$\omega^t(\mathscr{G}_n) \leq 1 - \mathbf{s}(n) \Longrightarrow \omega^t(\mathscr{G}_n^{Pararep}) \leq \frac{1}{2}$$

Before giving a proof for Proposition 7.4.4, we first show the following lemma.

Lemma 12.4.2 (Properties related to the parallel repeated anchoring transformation)**.** *Let $r \in \mathbb{N}$, and let $\mathscr{G} = (\mathscr{X}, \mathscr{A}, \mu, D)$ be a CL-samplable game where the question distribution μ is a (k, m, p). Then, $\mathscr{G}_{\perp}^{\otimes r}$, the r-fold parallel repetition of the anchor transformation of $\mathscr{G}$, is samplable via a $(k+1, r\cdot(m+2), p)$ CL distribution.*

Proof. Let $\mathtt{L}^A$ and $\mathtt{L}^B$ be the CL functions acting on $V \subseteq \mathbb{F}_{2^p}^m$, which define the distribution μ. We first show that $\mathscr{G}_{\perp}$ is samplable via a $(k+1, m+2, p)$ distribution. Define $\mathtt{L}_{\perp}^A, \mathtt{L}_{\perp}^B$ as the following: $\mathtt{L}_{\perp}^A$ is defined as a series composition (Definition 7.2.3) of two CL functions, where $V^1 = \mathbb{F}_{2^p}^2$, and $V_2 = V$. For $(s_0, s_1) \in \mathbb{F}_{2^p}^2$, define the zeroth level CL function for $\mathtt{L}_{\perp}^A$ is defined as

$$(\mathtt{L}_{\perp}^A)_{0,0}(s_0, s_1) = (s_0, 0)$$

and the collection of level k CL function to be $\{(\mathtt{L}_{\perp}^A)^{(s_0,s_1)}\}_{(s_0,s_1)\in\mathbb{F}_{2^p}^2}$ to be

$$(\mathtt{L}_{\perp}^A)^{(s_0,s_1)} = \mathtt{L}^A$$

if the first bit of $\kappa(s_0)$ is 0, where recall $\kappa(s_0) \in \{0,1\}^p$ is the canonical representation of s_0 and $(\mathtt{L}_{\perp}^A)^{(s_0,s_1)} = 0$, i.e. the linear function which maps all elements in $\mathbb{F}_{2^p}^m$ to 0, otherwise. Intuitively, the first bit of $\kappa(s_0)$ being 1 indicates that the anchoring question is sampled as Alice's question. $\mathtt{L}_{\perp}^B$ is defined similarly to $\mathtt{L}_{\perp}^A$ except the zeroth level CL function is defined as

$$(\mathtt{L}_{\perp}^B)_{0,0}(s_0, s_1) = (s_1, 0).$$

We see that the CL distribution defined by $\mathtt{L}_{\perp}^A$ and $\mathtt{L}_{\perp}^B$ precisely samples $\mu_{\perp}$. The proof of the lemma then follows from applying Lemma 7.2.6 on $\mathscr{G}_{\perp}$. □

We are now ready to show Proposition 7.4.4.

Proof. Fix the constant α and a function $\mathbf{s}(n) : \mathbb{N} \to [0,1]$, with $O(\mathbf{s}(n)) = O(\text{polylog}(n))$. Let $\mathbf{r}(n)$ be a function such that

$$\frac{16}{\mathbf{s}(n)} \cdot \exp\left(\frac{-c^{\text{para}}\mathbf{s}(n)^{17}\mathbf{r}(n)}{\alpha \log(n)}\right) \leq \frac{1}{2} \tag{12.67}$$

for all $n \in \mathbb{N}$. Since $O(\mathbf{s}(n)) = O(\text{polylog}(n))$, $\mathbf{r}(n)$ can be taken to be $O(\text{polylog}(n))$.

Fix an input $(\mathtt{Q}, \mathtt{D})$ and integer n. We describe the corresponding output $(\mathtt{Q}^{\text{Pararep}}, \mathtt{D}^{\text{Pararep}})$ for $\mathtt{Parallelrep}_{\alpha,\mathbf{s}(n)}$ below: If at any point in the computation process, the computation step for running $\mathtt{Q}$ and $\mathtt{D}$ either returns an invalid output or runs for time more than n^{α}, return 0 (i.e. an invalid input). The Turing machine $\mathtt{Q}^{\text{Pararep}}$ is defined as the following: $\mathtt{Q}^{\text{Pararep}}$ reads the first input n and computes $(\mathbf{k}(n), \mathbf{m}(n), \mathbf{p}(n)) = \mathtt{Q}(n, \text{parameter})$.

- $\mathtt{Q}^{\text{Pararep}}(n, \text{parameter}) = (\mathbf{k}(n), \mathbf{r}(n) \cdot (\mathbf{m}(n)+2), \mathbf{p}(n))$.
- On input (n, Divide, s), $\mathtt{Q}^{\text{Pararep}}$ does the following:
  1. Parses $s = (s_0, \cdots, s_{\mathbf{r}(n)-1})$, where each $s_i \in \{0,1\}^{(\mathbf{m}(n)+2)\cdot\mathbf{p}(n)}$ (automatically returns 0 if the input does not match).
  2. For each $i \in [\mathbf{r}(n)]$, parse $s_i = (s_{i,0}, s_{i,>0})$, where $s_{i,0} \in \{0,1\}^{2\cdot\mathbf{p}(n)}$ and $s_{i,>0} \in \{0,1\}^{\mathbf{p}(n)\cdot\mathbf{m}(n)}$. Furthermore, parse
  $$(s_{i,1}, \cdots, s_{i,\mathbf{k}(n)}) = \mathtt{Q}(n, \text{Divide}, s_{i,>0}).$$
  3. For $j \in [\mathbf{k}(n)+1]$, let $t_j = (s_{0,j}, \cdots, s_{\mathbf{r}(n)-1,j})$, and return $(t_0, \cdots, t_{\mathbf{k}}(n))$ as the output.
- On input $(n, \text{Function}, p, j, s, x)$, $\mathtt{Q}^{\text{Pararep}}$ does the following:
  1. Parses $s = (s_0, \cdots s_{\mathbf{r}(n)-1})$, and $x = (x_0, \cdots x_{\mathbf{r}(n)-1})$., where each s_i (resp. x_i) have the same number of bits with each other (automatically returns 0 if the input does not match).
  2. If $j = 0$, apply the zeroth level linear function as specified in the proof of Lemma 12.4.2 to each of the x_i to each $i \in [\mathbf{r}(n)]$ and return the concatenated output.
  3. Otherwise, for each $i \in [\mathbf{r}(n)]$, parse $s_i = (s_{i,0}, s_{i,>0})$ where $s_{i,0} \in \{0,1\}^{2\cdot\mathbf{p}(n)}$
     - If the first bit of $s_{i,0}$ is 0 (i.e. the normal question), compute $t_i = \mathtt{Q}(n, \text{Function}, p, j-1, s_{i,>0}, x_i)$.
     - Otherwise (i.e. anchoring question), let $t_i = 0$ be the zero string that has the same length as x_i.

     Return the concatenated output of all the t_i.

The Turing machine $\mathtt{D}^{\text{Pararep}}$ is defined as the following: $\mathtt{D}^{\text{Pararep}}$ reads the first input n, and computes $(\mathbf{k}(n), \mathbf{m}(n), \mathbf{p}(n)) = \mathtt{Q}(n, \text{parameter})$.

1. On input (n, x, y, a, b), parse
$$x = (x_0, \cdots, x_{\mathbf{r}(n)-1}), \quad y = (y_0, \cdots, y_{\mathbf{r}(n)-1}), \quad a = (a_0, \cdots, a_{\mathbf{r}(n)-1}) \quad b = (b_0, \cdots, b_{\mathbf{r}(n)-1}),$$
where $x_i, y_i \in \{0,1\}^{(\mathbf{m}(n)+2),\mathbf{p}(n)}$ for all $i \in [\mathbf{r}(n)]$ and each of the a_i, b_i have the same number of bits. Output 0 (i.e. automatically reject) if this cannot be done.
2. For each $i \in [\mathbf{r}(n)]$, parse $x_i = (x_{i,0}, x_{i,>0})$, $y_i = (y_{i,0}, y_{i,>0})$ where $x_{i,0}, y_{i,0} \in \mathbb{F}_{2^p}^2$.
   - If the first bit of $x_{i,0}$ and $y_{i,0}$ are both 0 (i.e. the normal question), compute $t_i = \mathtt{D}(n, x_{i,>0}, y_{i,>0}, a_i, b_i)$.
   - Otherwise, $t_i = 1$ iff whenever x_i (resp. y_i) is the anchoring question (i.e. the first bit of $x_{i,0}$ (resp. y_i) is zero), then a_i (resp. b_i) is the zero string. $t_i = 0$ if the above condition is not met.
3. Return $\bigwedge_{i\in[\mathbf{r}(n)]} t_i$.

As seen from the description above, both $(\mathtt{Q}^{\text{Pararep}}, \mathtt{D}^{\text{Pararep}})$ can be described by using $(\mathtt{Q}, \mathtt{D})$ as a black box, and $\mathtt{Q}^{\text{Pararep}}$ depends only on $\mathtt{Q}$ and

- $\mathsf{TIME}_{\mathtt{Q}^{\text{Pararep}}}(n) \le O(\text{poly}(\mathbf{r}(n), log^{\alpha}(n)))$,
- $\mathsf{TIME}_{\mathtt{D}^{\text{Pararep}}}(n) \le O(\text{poly}(\mathbf{r}(n), log^{\alpha}(n)))$.

If $(\mathtt{Q},\mathtt{D})$ is a synchronous k-th level CL sampler for an infinite sequence of synchronous games $\{\mathcal{G}_n = (\mathcal{X}_n, \mathcal{A}_n, \mu_n, D_n)\}_{n\in\mathbb{N}}$ for some constant $k \in \mathbb{N}$,with some constant $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$\mathtt{Q}(n, \text{Parameter}), \mathsf{TIME}_{\mathtt{Q}}(n), \mathsf{TIME}_{\mathtt{D}} \leq \log^{\alpha}(n).$$

Then $(\mathtt{Q}^{\text{Pararep}}, \mathtt{D}^{\text{Pararep}})$ is a $(k+1)$-th level CL sampler for an infinite sequence of synchronous games $\{\mathcal{G}_{n,\perp}^{\otimes \mathbf{r}(n)}\}_{n\in\mathbb{N}}$, where $\mathcal{G}_{n,\perp}^{\otimes \mathbf{r}(n)}$ is the $\mathbf{r}(n)$-fold parallel repetition of the anchoring transformation for $\mathcal{G}_n$.

For completeness, note if $\mathcal{G}$ admits a perfect (synchronous) oracularizable strategy $\mathcal{S} = (\mathcal{L}^2(\mathcal{A}, \tau),$
$|\tau\rangle, \{A_a^x\})$. Then the r-fold parallel repetition game $\mathcal{G}^{\otimes r}$ admits a perfect oracularizable strategy defined on $\mathcal{L}^2(\mathcal{A}, \tau)^{\otimes r}$ because the provers can simply performing r independent instances of $\mathcal{S}$ for each question labels on the parallel repeated game. Hence, combining with the remark after Definition 12.2.1, if there exists a perfect oracularizable strategy for $\mathcal{G}_n$, then there exists a perfect oracularizable strategy for $\mathcal{G}_{n,\perp}^{\otimes \mathbf{r}(n)}$.

For soundness, since $\mathsf{TIME}_{\mathtt{D}} \leq \log^{\alpha}(n)$ for $n > n_0$, this implies that $|\mathcal{A}| = \log^{\alpha}(n)$, and hence combining (12.67) and Theorem 12.2.3 shows the soundness condition. This completes the proof of Proposition 7.4.4. □

Part IV
Implications and open problems

CHAPTER 13

Implications and open problems

We end this thesis with a few implications from the $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ theorem and results from this thesis, as well as a few open problems which could be interesting.

13.1 Implications

We start this chapter by discussing some corollaries arising from the $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ theorem.

13.1.1 Corollaries for non-local games and $\mathsf{MIP}^{\mathsf{co}}$ protocols

Computability of the commuting operator value

Recall from Chapter 6, that the NPA hierarchy is an algorithm which generates a series of upper bounds that estimates the commuting operator value of a game. As shown in [JNV+22a, Theorem 12.10], there exists a game $\mathcal{G}$ such that $\omega^{co}(\mathcal{G}) = 1$ and $\omega^{*}(\mathcal{G}) \leq \frac{1}{2}$. The best algorithm used in practice for estimating the tensor product value of the game is still the NPA hierarchy due to the inefficiency of the `Searchfrombelow` algorithm. Estimating the commuting operator value of a game is also important in the recently introduced compiled non-local game setting [KLV+22], where the optimal bound of the game is known to be bounded by the commuting operator value of the game due to a recent result by [KMP+25].

An obvious consequence of the $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ theorem is that **there is no algorithm which can estimate the commuting operator value of a game** up to any constant $c \in (0,1)$, or else one can use this algorithm to construct a halting algorithm for the *co* non-local game value problem. Our result also implies that one cannot compute the convergence rate of the NPA hierarchy in general.

A variant of $\mathsf{MIP}^{\mathsf{co}}$ called the zero-gap $\mathsf{MIP}^{\mathsf{co}}$ ($\mathsf{MIP}^{\mathsf{co}}_0$) is introduced in [MNY20]. $\mathsf{MIP}^{\mathsf{co}}_0$ is the complexity class which is complete with respect to the gapless commuting operator value problem, which is defined by the following two sets of non-local games

$$\mathsf{L}^{co}_{\text{yes}} = \left\{\mathcal{G} : \omega^{co}(\mathcal{G}) = 1\right\} \qquad \text{and} \qquad \mathsf{L}^{co}_{\text{no}} = \left\{\mathcal{G} : \omega^{co}(\mathcal{G}) < 1\right\}.$$

This class is also shown to be coRE complete by a result of [Slo19a], which implies that $\mathsf{MIP}^{\mathsf{co}} = \mathsf{MIP}^{\mathsf{co}}_0$. Combining this with our result, this shows that $\mathsf{MIP}^{\mathsf{co}}_0 = \mathsf{MIP}^{\mathsf{co}}$, or in other words, it is computationally equivalent to compute the exact commuting operator value of a game and to estimate it up to $\frac{1}{2}$ error. Interestingly, the equivalency between the gap and zero-gap versions of $\mathsf{MIP}^{\mathsf{co}}$ does not extend to MIP^{*}, as the complexity of the zero-gap MIP^{*} (MIP^{*}_0) is shown to be Π_2-complete in [MNY22], where Π_2 is defined as coRE with access to an RE oracle.

In [MS24a], it was shown that every MIP^{*} protocol has a perfect zero-knowledge proof system in the MIP^{*} model. In the same work, it was conjectured that the same would hold for the complexity class $\mathsf{MIP}^{\mathsf{co}}$ provided $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$, and a parallel repetition theorem was proved for the commuting operator model. In conjunction with the result from this thesis, one could similarly convert any $\mathsf{MIP}^{\mathsf{co}}$ protocol into a zero-knowledge $\mathsf{MIP}^{\mathsf{co}}$ protocol.

$\mathsf{MIP}^{\mathsf{co}}$ protocol with short questions and answer size.

An interesting open question left open by $\mathsf{MIP}^* = \mathsf{RE}$ is the following: what happens if we restrict the question size and answer size of a $\mathsf{MIP}^{\mathsf{co}}$ protocol while still allowing the verifier polynomial-time verification. To discuss this, we define $\mathsf{MIP}^*(q(n), a(n))$ (resp. $\mathsf{MIP}^{\mathsf{co}}(q(n), a(n))$) to be the complexity class described by a MIP^* (resp. $\mathsf{MIP}^{\mathsf{co}}$) protocol with question size $q(n)$ and answer size $a(n)$.

In the case where the answer size is restricted, it was shown in [DFN+23] that if the answer size of an MIP^* protocol is restricted to a constant, then the resulting complexity class remains RE-complete. Since the technique used to prove this theorem is essentially the same as the tools used for $\mathsf{MIP}^* = \mathsf{RE}$, we have the following corollary.

Corollary 13.1.1. $\mathsf{MIP}^{\mathsf{co}}(O(\mathrm{poly}(n)), O(1)) = \mathsf{coRE}$

The proof of the above corollary follows much of a similar intuition as the proof from Part III of this thesis: by using tracially embeddable strategies to generalize the argument holding for MIP^* to the $\mathsf{MIP}^{\mathsf{co}}$ model.

Proof. The proof follows similarly to the proof of [DFN+23, Theorem 6.10], which applies two iterations of the transformation from [NZ23b, Theorem 52] followed by a single iteration of the Hadamard test, given in [DFN+23, Figure 4]. [NZ23b] follows by applying a "scale-up" variant of the answer reduction protocol, and then the parallel repetition theorem for projective games from [DSV15]. Since the scale-up variant of the answer reduction protocol still depends on the simultaneous quantum low-individual degree test, by Lemma 10.4.1 the completeness/soundness condition for the commuting operator model follows accordingly. Interestingly, the proof of [DSV15] does not use information-theoretic tools like most other proofs of the parallel repetition theorems. However, the proof can still be extended to the commuting operator setting by working with tracially embeddable strategies (along with the translation chart from Table 4.1). The completeness/soundness condition of the Hadamard test can be extended similarly by working with tracially embeddable strategies. □

On the other hand, if we consider the case where the question size is restricted, [NZ23a] has shown that $\mathsf{MIP}^*(O(1), O(\mathrm{polylog}(n))) = \mathsf{RE}$. We can again conclude a similar corollary using similar proof techniques.

Corollary 13.1.2. $\mathsf{MIP}^{\mathsf{co}}(O(1), O(\mathrm{polylog}(n))) = \mathsf{coRE}$

Proof. The proof also follows a similar structure as the proof of [NZ23b, Theorem 54], where the authors use a different configuration of the three propositions used in the proof of the compression theorem for MIP^* in order to show the existence of a "hypercompression" scheme for $\mathsf{MIP}^* = \mathsf{RE}$. Hence, the corollary follows from applying Proposition 7.4.2, Proposition 7.4.3, Proposition 7.4.4 in the same order as the proof given in [NZ23b, Theorem 54] in order to show a hypercompression scheme for $\mathsf{MIP}^{\mathsf{co}}$. □

Finally, if both the question size and the answer size of the protocol are restricted to constant size $O(1)$, the complexity of $\mathsf{MIP}^*(O(1), O(1))$ is an open problem in itself. I personally don't have a good intuition of whether this class should be computable or not, but $\mathsf{MIP}^{\mathsf{co}}(O(1), O(1))$ should have a similar computability property as $\mathsf{MIP}^*(O(1), O(1))$.

13.1.2 Implication for operator algebras

Results related to continuous model theory.

As mentioned in the introduction, proving $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ gives many implication for continuous model theory. Since continuous model theory is not the focus of this thesis, we refer the reader to [Har23] for more in-depth introduction on this subject.

An interesting result by Jananan Arulseelan and Aareyan Manzoor establishes a stronger variant of the Connes embedding theorem [AM25]. To introduce this theorem, we need to first define the notion of a *locally universal* tracial von Neumann algebra.

Definition 13.1.3 (Locally universal tracial von Neumann algebra, Definition 4.1 of [AM25])**.** *A tracial von Neumann algebra $\mathscr{S}$ is locally universal if every separable tracial von Neumann algebra embeds in an ultrapower of $\mathscr{S}$.*

To rephrase Connes embedding theorem introduced in Section 1.1.2 using this language, the negative answer of Connes implies that the hyperfinite II_1 factor $\mathscr{R}$ is not locally universal. Arulseelan and Manzoor show that, in fact, there is no locally universal tracial von Neumann algebra which admits a computable presentation. Intuitively, one should view a separable tracial von Neumann as inherently an uncomputable object by the $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ theorem proven in this thesis. In particular, there is no computable presentation that allows one to algorithmically approximate all evaluations arising from such an algebra (in the sense that there does not exist an algorithm, when given a state ψ and an element a from the algebra, the algorithm will approximate $\psi(a)$). In the same paper, this result is also shown to extend to the II_∞ case, meaning that even separable semifinite von Neumann algebras are, by nature, inapproximable by any computational algorithm.

At the same time, by a result of Isaac Goldbring and Thomas Sinclair, it was shown the maximal tensor norm between two C^*-algebras is not explicitly definable in the framework of continuous model theory assuming $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$. Interestingly, the minimal tensor product is also shown not to be explicitly definable by using the $\mathsf{MIP}^* = \mathsf{RE}$ theorem. The C^*-algebra considered in Goldbring and Sinclair's result is specifically focus on the case where the C^*-algebra in question as the group C^*-algebra with n generators with $n \geq 2$, this also gives a negative answer to an open problem proposed by Tobias Fritz, Tim Netzer, and Andreas Thom [FNT14, Problem 4.2][1].

I believe that these two results illustrate the nature of the $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ theorem in comparison to $\mathsf{MIP}^* = \mathsf{RE}$ in the context of operator algebras. In some sense, the $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$ theorem is a general uncomputability result about tracial von Neumann algebras. In contrast, the $\mathsf{MIP}^* = \mathsf{RE}$ theorem is an uncomputability result about what is approximatable by a "limit" of infinite-dimensional matrices.

Connection to noncommutative polynomials.

Let $\mathscr{F}_n^m$ be the free group consisting of m elements of order n, and let $\mathbb{Q}(\mathscr{F}_n^m)$ be the finitely-generated $*$-algebra over $\mathbb{Q}$. [MSZ23, Theorem 1.1] showed that the $\mathbb{Q}(\mathscr{F}_n^m)$ tensor product positivity problem is coRE-hard in the case where $n, m \geq 2$, $(n,m) \neq (2,2)$, where the $\mathbb{Q}(\mathscr{F}_n^m)$ tensor product positivity problem is defined as follows: Given an element $g \in \mathbb{Q}(\mathscr{F}_n^m) \otimes \mathbb{Q}(\mathscr{F}_n^m)$, deciding whether the element g is positive in $(\mathbb{Q}(\mathscr{F}_n^m) \otimes \mathbb{C}) \otimes (\mathbb{Q}(\mathscr{F}_n^m) \otimes \mathbb{C})$? Since one can also view elements of $(\mathbb{Q}(\mathscr{F}_n^m) \otimes \mathbb{C})$ as an m-variate noncommutative polynomial over elements of order n, the above problem can also be formulated as determining the positivity for two noncommutative polynomials tensored together. The $\mathbb{Q}(\mathscr{F}_n^m)$ tensor product positivity problem is conjectured to be Π_0^2-complete.

By considering the game polynomial introduced in [WHK23] and its connection to the commuting operator strategies, the complexity class $\mathsf{MIP}^{\mathsf{co}}$ is related to the "gap" version of the $\mathbb{Q}(\mathscr{F}_n^m)$ tensor product positivity problem, where the gap $\mathbb{Q}(\mathscr{F}_n^m)$ tensor product positivity problem is defined as the following decision problem: given $g \in \mathbb{Q}(\mathscr{F}_n^m) \otimes \mathbb{Q}(\mathscr{F}_n^m)$, decide either $g - \mathbb{I}$ is positive or g is not positive, where $\mathbb{I}$ is the identity element in $\mathbb{Q}(\mathscr{F}_n^m) \otimes \mathbb{Q}(\mathscr{F}_n^m)$. [MSZ23] shows that our main theorem, $\mathsf{MIP}^{\mathsf{co}} = \mathsf{coRE}$, implies that the gap $\mathbb{Q}(\mathscr{F}_n^m)$ tensor product positivity problem is also RE-complete. This, in turn, also implies that the $\mathbb{Q}(\mathscr{F}_n^m)$ tensor

[1] Thus showing that "no, you cannot compute the operator norm"! Thereby answering the question proposed by the title of this paper "Can you compute the operator norm?"

product positivity problem is at least RE-hard, giving stronger evidence that this problem is Π_0^2-complete.

13.2 Open problems

I will end this thesis with several open problems. These problem are not necessarily mine, but these are the problems that I encountered during my Ph.D. which I found to be interesting. Hopefully putting these problems in Chapter 13 of this thesis won't bring it bad luck for solving them in the future.

13.2.1 Open problems about the compression paradigm

In Chapter 2, we introduced the generalized compression paradigm for showing RE-completeness or coRE-completeness of a given decision problem, and we showed that the gap compression theorem for non-local games gives a way to use the generalized compression framework for showing that the tensor product/commuting operator value problem is RE/coRE-complete. As mentioned in Chapter 2, the compression paradigm requires one to first show that one can simulate the uniform succinct problem instance efficiently, and both examples of RE/coRE-completeness are based on the underlying problem being NP-hard in a specific manner. It would be interesting to see what other uncomputable problems satisfy the compression paradigm "naturally" (i.e. a compression map can be defined on a succinct instance of said problem without going through a reduction to another problem) can one use these techniques to show other problems to be uncomputable. Alternatively, can one build another uncomputable problem based on another PCP construction (for example, Dinur's graph expansion based PCP [DV07]).

Another natural follow-up question is whether one can simplify other proofs of RE/coRE-completeness using the compression paradigm. For example, one such problem is the spectral gap problem for many-body systems [CPW15], whose proof is based on a reduction to the Wang tiling problem. Perhaps one can apply the compression directly in order to show uncomputability.

In the compression paradigm proof, there is an assumption that L_{no} for the decision problem needs to be in coRE for this argument to go through, and we point out some of the potential pitfalls if this assumption is not satisfied. An interesting open problem is whether one can formulate a notion of the compression technique for higher levels of the arithmetic hierarchy? In particular, [MNY22] defines a notion of super compression which is used to show that MIP_0^* is Π_2-complete. Is there a way to formulate this for Π_2-completeness in general? Can this general formulation be made to hold for higher levels of the arithmetic hierarchy?

13.2.2 Open problems about non-local correlations and non-local games.

Open problems related to MIP^{co}

As mentioned earlier in Section 13.1, $\mathsf{MIP}^{\text{co}} = \mathsf{MIP}_0^{\text{co}}$, where $\mathsf{MIP}_0^{\text{co}}$ is the zero-gap variant of MIP^{co}. This implies that there must exist a direct reduction from the commuting operator value problem to the gapless commuting operator value problem. Thus, an interesting open problem is whether there exists a natural reduction between these two problems without going through the Non-Halting problem?

Another open problem is whether there exists a more reasonable experimental realization between the tensor product model and the commuting operator model of entanglement. Although we show that no algorithm can find such a separation, this does not eliminate other mathematical techniques which can be used to find a game that realizes said separation. Another interesting question is whether one can find the smallest question/answer size for which

a separation can be found. This game would also provide a more reasonable way to construct a counterexample to Connes' embedding problem.

Does there exist a more general uniform sequence of games that admits a gap compression theorem, or are conditionally linear verifiers the most general class of uniform games that can be compressed? As discussed in the technical overview, conditionally linear verifiers are tailor-made to take advantage of self-testing from non-local games literature and PCP constructions from the interactive proof systems literature, so we suspect any general uniform game sequence that admits a gap compression theorem must also be defined in a way that enables both techniques.

In this thesis, we show that the commuting operator value problem with parameters $(1, 1-\varepsilon)$ for conditional linear sampler games is coRE-complete for all constant $\varepsilon > 0$. One natural question is whether this result holds for other classes of games. [MSS+25] shows that there exists a constant $c_{\text{Inde}} > 0$ such that the tensor product value problem for independent set games with parameters $(1, 1-c_{\text{Inde}})$ is RE-complete, and assuming the main result of our work, the commuting operator value problem for independent set games with parameters $(1, 1-c_{\text{Inde}})$ would be coRE-complete. Similar results have been derived for the constraint satisfaction problem (CSP) games and 3-colouring games by [CM25]. An interesting open problem is whether these types of results hold for other classes of games for both the tensor product value and the commuting operator value problems.

At the other extreme, are there parameters in which the commuting operator value problem is "easy" (i.e. computable)? In [CMS24], it was shown that there exists a parameter d_{col} such that for all $d' < d_{\text{col}}$, the tensor product value problem with parameters $(1, 1-d')$ for 3-colouring games is decidable in polynomial time! Does the same phenomenon hold for commuting operator values? Does there exist a class of games such that the tensor product value problem with parameters $(1, 1-c)$ is computable, but uncomputable for the commuting operator value problem with the same parameter (or vice versa)? Does there exist a variant of the "unique games conjecture" analogous to classical MIP [Kho02] or MIP^* [KRT09; MS24b] for MIP^{co}?

Lastly, another interesting question is what it means to have noisy entanglement in the MIP^{co} model. It was shown that if the provers' measurements were restricted to noisy EPR pairs, the complexity corresponds inversely with the presence of noise [DFN+23]. However, since there is no notion of a tensor factor within the commuting operator model, there could be MIP^{co} protocols with more exotic noise models than in MIP^*. Is there a more interesting noise model for the commuting operator model that one could consider? Does noise reduce the computational power of MIP^{co}?

13.2.3 Open problems related to tracially embeddable strategies and correlations

Exact tracially embeddable strategies. In Theorem 4.1.3, we only show that the *closure* of the set of correlations generated by a tracially embeddable strategy $\mathscr{C}_{qc}^{Tr}$ is equal to the set of commuting operator correlations. Thus, a natural question would be whether the correlation set $\mathscr{C}_{qc}^{Tr}$ is closed, in the ℓ_1 sense, at all. We suspect the answer is no since to get the positive σ in the proof of Theorem 4.1.3, an approximation theorem must be applied to the states in $\mathscr{L}^1(\mathscr{A}, \tau)$ (Proposition 4.2.3). Since the set of commuting operator correlations is closed, a negative answer to the above question would imply that there exist some quantum commuting operator correlations which a tracially embeddable strategy cannot realize.

Another interesting question is whether the same question posed in the last paragraph is true when we relax the definition of tracially embeddable strategy and allow the state to be any vector state $|\psi\rangle \in \mathscr{H}$ instead, where $\mathscr{H}$ is the Hilbert space where the standard form of $\mathscr{A}$ is represented in. Intuitively, the formulation above is a tracially embeddable strategy variant of "C_q" in the tensor product model while the variant given can be intuitively thought of as C_{qa}.

Implementation of a tracial state within the NPA hierarchy. As mentioned in Chapter 6, in [NPA08], a hierarchy of semidefinite programs is used in order to determine the optimal success probability (i.e. the commuting operator value) of a non-local game. Roughly speaking, level n of the hierarchy optimizes over the Gram matrix $\{S|\psi\rangle\}$ where S consists of products of observables of length n from Alice or Bob. By using Theorem 4.1.3, it is possible to define a variant of the NPA hierarchy where σ is considered to be an additional variable that can be added to each level n, as well as an additional constraint to be added using the cyclicity property of the trace. Does this variant of the hierarchy require fewer levels for convergence, or can it be potentially more efficient when the hierarchy does converge?

Operator algebraic formulation of self-testing. Roughly speaking, a self-test refers to the uniqueness of quantum strategies given a specific correlation set, or correlation sets close to it in the approximate case. It was recently shown in [PSZ+24] that having an exact self-test is equivalent to having a unique abstract state acting on the game algebra. However, it is unclear what the definition of robust self-testing should be in the operator algebraic framework.

In this thesis, we take advantage of the structure of tracially embeddable strategies and define a notion of robust self-testing of any approximate correlation for the commuting operator model in Theorem 8.3.2. Informally, in our definition, a game has the rigidity property if, for all tracially embeddable strategies, there exists a pair of partial isometries which commute with each other (this corresponds to Alice and Bob applying the partial isometries locally), and map the state used in the approximate strategy to a vector state that is close to the state arising from the ideal tracially embeddable strategy. We remark that this notion of self-testing is similar to the notion of self-testing defined in [NV17; CS19].

The proof of the main theorem could also give some potential insight into defining robust self-testing in the operator algebraic framework. The proof mainly relies on Haagerup's reduction [HJX10, Theorem 2.1], which, roughly speaking, approximates the crossed product of the game algebra by a sequence of increasing tracial von Neumann algebras. A potential definition for robust self-testing could be: for any other correlation close to the given correlation, the approximate correlation must be realized by a state on a tracial von Neumann algebra, and this state must also be close to the "unique state" from the definition of exact self-testing. Since this is beyond the scope of this thesis, we will instead leave it as an open problem about the proper definition of robust self-testing in the algebraic framework and whether it is equivalent to the definition we have given in Theorem 8.3.2.

13.2.4 Open problems related to operator algebras

Additional insight into the Connes embedding problem. Our proof techniques for $\mathsf{MIP}^* = \mathsf{RE}$ and $\mathsf{MIP}^{\mathrm{co}} = \mathsf{coRE}$ can potentially give an alternative perspective into the counter-example for the Connes embedding problem. If we view a non-local game as a functional which maps correlations into $(0,1)$, a norm on this set of functionals would correspond to their optimal success rate on a given correlation set. Theorem 1.2.1 can intuitively be seen as a map that maps functionals acting on a correlation to one that acts on a smaller correlation set (in terms of input/output), while maintaining the norm to some degree. The fact that Theorem 1.2.1 preserves both tensor product and commuting operator values means that the difference between the tensor product value and the commuting operator value only lies in how the compression is being used. This intuition might be useful in constructing a counterexample for the Connes embedding problem using operator algebraic techniques.

Due to the characterization by [Fri12], the tensor product model corresponds to the "max" tensor product between two free algebras, whereas the commuting operator model corresponds to the "min" tensor product between two free algebras. Thus, $\mathsf{MIP}^{\mathrm{co}} = \mathsf{coRE}$ allows one to work with the "max" tensor product when considering operator algebraic results which rely on the

Connes embedding problem being false. Can this additional insight be helpful for the operator algebra community?

Lastly, does the undefinability result for locally universal von Neumann algebras discussed in Section 13.1.2 extend to type III von Neumann algebras? This is an open question asked by [AM25], which I found to be interesting. The reason why this result can be extended easily from the II_1 setting to the II_∞ setting is due to the following fact: If $\mathscr{A}$ is a type II_1 von Neumann algebra, the algebra $\mathscr{A} \otimes \mathscr{B}(\mathscr{H})$ for some infinite-dimensional, separable Hilbert space $\mathscr{H}$ is type II_∞. As mentioned in [AM25], extending this to the type III setting will likely require some heavy lifting from the Tomita-Takesaki construction [Tak01].

Application to other operator algebra problems. The complexity of approximating the values of non-local games has a natural connection to the study of operator algebra. As mentioned above, the $\mathsf{MIP}^* = \mathsf{RE}$ theorem gives a negative answer to the Connes embedding problem in the study of tracial von Neumann algebras. [BCL+24; BCV24] gives a negative answer to the Aldous-Lyons problem [AL18] in probability theory by showing that $\mathsf{TailoredMIP}^*$, MIP^* with a more restricted class of strategies, is RE-complete. An important open problem in group theory is whether a non-hyperlinear group exists. It is shown in [PS25] that LinMIP^*, or MIP^* protocols being restricted to Linear Constraint System games [CM14; KPS18], being computable is equivalent to the existence of a non-hyperlinear group. The class of Linear Constraint System games does not fall into the conditionally linear verifier framework, and it would be interesting to see if a similar compression technique can be used to resolve this problem. Interestingly, resolving this problem would also imply the existence of a non-sofic group, which is one of the biggest open problems in modern group theory.

Another interesting set of strategies is the set of invariant random subgroup (IRS) strategies introduced in [Man25b]. Intuitively, this can be seen as the "commuting operator variant" of the "Z-aligned permutation strategies" introduced in [BCL+24], used to disprove the Aldous-Lyons problem. It is conjectured that $\mathsf{MIP}^{\mathrm{IRS}}$, MIP with access to IRS strategies, is coRE-complete. Showing this complexity-theoretical result would also give a similar "undefinability" statement for invariant random subgroups. Showing this conjecture would also have interesting implications for the Ergodic theory community, and we refer to [Man25a] for more details.

A more general question, proposed by Andrew Marks at the UT Austin Graduate Mini-School on Sofic Groups and the Connes Embedding Problem, is whether one can phrase the Connes embedding problem and other types of "approximation" type problem in terms of a decision problem, and potentially apply the compression paradigm onto these problems? To me, the advantage of non-local games seems to be that one can use both self-testing techniques and the quantum low-individual degree test in order to define a compression map. For the aforementioned LinMIP^* problem, the bottleneck seems to be the lack of a perfect completeness linear PCP, for which there is not even a natural classical analogue! For solving operator-algebraic inspired problems, if one were to try to use the techniques from this thesis, perhaps the right approach is to move away from non-local games, and explore other ways to use the compression paradigm.

13.3 Outro

I personally do not want to end the technical portion of this thesis on a negative note about the role of non-local games on operator algebraic problems for the future. Since the majority of this thesis is about the complexity of non-local games, my heart still wishes that these techniques can be pushed further. This is why I am adding an outro section.

To be fair, I don't have a good intuition about whether continuously exploring the connection between operator algebra and quantum complexity would lead to more groundbreaking research in operator algebra. However, at the time of writing this thesis, I am optimistic that

the compression paradigm will bring many more interesting results for both the complexity theory, computability theory and operator algebra communities, and I hope to realize this in my academic career. Hopefully in 10 years, when I look back on this thesis, I will have a better understanding of the potential and limitations of using non-local games for operator algebraic problems!

Bibliography

[AL18] David Aldous and Russell Lyons. *Processes on Unimodular Random Networks*. 2018. DOI: 10.48550/arXiv.math/0603062. arXiv: math/0603062 (cit. on p. 199).

[AR94] Noga Alon and Yuval Roichman. *Random Cayley Graphs and Expanders*. Random Structures & Algorithms **5**, 2 (1994), pp. 271–284. DOI: 10.1002/rsa.3240050203 (cit. on p. 91).

[AP10] Claire Anantharaman and Sorin Popa. *An Introduction to II1 Factors* (2010) (cit. on pp. 35, 47).

[Ara74] Huzihiro Araki. *Some Properties of Modular Conjugation Operator of von Neumann Algebras and a Non-Commutative Radon-Nikodym Theorem with a Chain Rule*. Pacific Journal of Mathematics **50**, 2 (1974), pp. 309–354 (cit. on pp. 45, 48, 160, 161).

[Ara77] Huzihiro Araki. *Relative Entropy for States of von Neumann Algebras. II*. Publications of the Research Institute for Mathematical Sciences **13**, 1 (1977), pp. 173–192. DOI: 10.2977/prims/1195190105 (cit. on pp. 13, 162, 163).

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. *Proof Verification and the Hardness of Approximation Problems*. Journal of the ACM **45**, 3 (1998), pp. 501–555. DOI: 10.1145/278298.278306 (cit. on pp. 7, 128, 139, 148).

[AS98] Sanjeev Arora and Shmuel Safra. *Probabilistic Checking of Proofs: A New Characterization of NP*. Journal of the ACM **45**, 1 (1998), pp. 70–122. DOI: 10.1145/273865.273901 (cit. on pp. 7, 128).

[AM25] Jananan Arulseelan and Aareyan Manzoor. *The Universal Theory of Locally Universal Tracial von Neumann Algebras Is Not Computable*. 2025. DOI: 10.48550/arXiv.2508.21709. arXiv: 2508.21709 [math] (cit. on pp. 195, 199).

[AGR82] Alain Aspect, Philippe Grangier, and Gérard Roger. *Experimental Realization of Einstein-Podolsky-Rosen-Bohm* Gedankenexperiment *: A New Violation of Bell's Inequalities*. Physical Review Letters **49**, 2 (1982), pp. 91–94. DOI: 10.1103/PhysRevLett.49.91 (cit. on p. 2).

[BFL91] László Babai, Lance Fortnow, and Carsten Lund. *Non-Deterministic Exponential Time Has Two-Prover Interactive Protocols*. computational complexity **1**, 1 (1991), pp. 3–40. DOI: 10.1007/BF01200056 (cit. on pp. 7, 16, 65, 73, 86, 127, 128, 132).

[BRR+09] Boaz Barak, Anup Rao, Ran Raz, Ricky Rosen, and Ronen Shaltiel. *Strong Parallel Repetition Theorem for Free Projection Games*. Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques. Ed. by Irit Dinur, Klaus Jansen, Joseph Naor, and José Rolim. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2009, pp. 352–365. ISBN: 978-3-642-03685-9. DOI: 10.1007/978-3-642-03685-9_27 (cit. on p. 167).

[BVY21] Mohammad Bavarian, Thomas Vidick, and Henry Yuen. *Anchored Parallel Repetition for Nonlocal Games*. arXiv:1509.07466 [quant-ph] (2021). arXiv: 1509.07466 [quant-ph] (cit. on pp. 13, 17, 87, 159, 166–178, 180, 181, 184).

[Bel64] J. S. Bell. *On the Einstein Podolsky Rosen Paradox*. Physics Physique Fizika **1**, 3 (1964), pp. 195–200. DOI: 10.1103/PhysicsPhysiqueFizika.1.195 (cit. on p. 1).

[Bla06] Bruce Blackadar. *Operator Algebras*. Ed. by Joachim Cuntz and Vaughan F.R. Jones. Vol. 122. Encyclopaedia of Mathematical Sciences. Berlin, Heidelberg: Springer, 2006. ISBN: 978-3-540-28486-4 978-3-540-28517-5. DOI: 10.1007/3-540-28517-2 (cit. on pp. 5, 36).

[BGM+93] Ian F. Blake, XuHong Gao, Ronald C. Mullin, Scott A. Vanstone, and Tomik Yaghoobian. *Applications of Finite Fields*. Ed. by Alfred J. Menezes. Boston, MA: Springer US, 1993. ISBN: 978-1-4419-5130-4 978-1-4757-2226-0. DOI: 10.1007/978-1-4757-2226-0 (cit. on p. 73).

[BCL+24] Lewis Bowen, Michael Chapman, Alexander Lubotzky, and Thomas Vidick. *The Aldous–Lyons Conjecture I: Subgroup Tests*. 2024. DOI: 10.48550/arXiv.2408.00110. arXiv: 2408.00110 [math] (cit. on pp. 18, 199).

[BCV24] Lewis Bowen, Michael Chapman, and Thomas Vidick. *The Aldous–Lyons Conjecture II: Undecidability*. 2024. DOI: 10.48550/arXiv.2501.00173. arXiv: 2501.00173 [quant-ph] (cit. on pp. 18, 199).

[CHS+69] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. *Proposed Experiment to Test Local Hidden-Variable Theories*. Physical Review Letters **23**, 15 (1969), pp. 880–884. DOI: 10.1103/PhysRevLett.23.880 (cit. on p. 2).

[CHT+04] R. Cleve, P. Hoyer, B. Toner, and J. Watrous. *Consequences and Limits of Nonlocal Strategies*. Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004. 2004, pp. 236–249. DOI: 10.1109/CCC.2004.1313847 (cit. on pp. 8, 66).

[CM14] Richard Cleve and Rajat Mittal. *Characterization of Binary Constraint System Games*. Automata, Languages, and Programming. Ed. by Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias. Berlin, Heidelberg: Springer, 2014, pp. 320–331. ISBN: 978-3-662-43948-7. DOI: 10.1007/978-3-662-43948-7_27 (cit. on pp. 92, 199).

[CS19] Andrea Coladangelo and Jalex Stark. *Robust Self-Testing for Linear Constraint System Games*. 2019. arXiv: 1709.09267 [quant-ph] (cit. on p. 198).

[Con76] A. Connes. *Classification of Injective Factors Cases*. The Annals of Mathematics **104**, 1 (1976), p. 73. DOI: 10.2307/1971057. JSTOR: 1971057 (cit. on pp. 4, 56).

[CPW15] Toby Cubitt, David Perez-Garcia, and Michael M. Wolf. *Undecidability of the Spectral Gap (Full Version)*. 2015. DOI: 10.1017/fmp.2021.15 (cit. on p. 196).

[CM25] Eric Culf and Kieran Mastel. *RE-completeness of Entangled Constraint Satisfaction Problems*. 2025. DOI: 10.48550/arXiv.2410.21223. arXiv: 2410.21223 [quant-ph] (cit. on p. 197).

[CMS24] Eric Culf, Hamoon Mousavi, and Taro Spirig. *Approximation Algorithms for Noncommutative CSPs*. 2024. DOI: 10.48550/arXiv.2312.16765. arXiv: 2312.16765 [quant-ph] (cit. on p. 197).

[dlS22a] Mikael de la Salle. *Orthogonalization of Positive Operator Valued Measures*. 2022. DOI: 10.48550/arXiv.2103.14126. arXiv: 2103.14126 [quant-ph] (cit. on p. 36).

[dlS22b] Mikael de la Salle. *Spectral Gap and Stability for Groups and Non-Local Games*. 2022. arXiv: 2204.07084 [math] (cit. on pp. 15, 90, 91, 96, 102).

[dlS23] Mikael de la Salle. *Algèbras de von Neumann, Produits Tensoriels, Corrélations Quantiques et Calculabilité [d'après Ji, Natarajan, Vidick, Wright et Yuen]*. 2023 (cit. on p. 96).

[dlSM23] Mikael de La Salle and Amine Marrakchi. *Almost Synchronous Correlations and Tomita-Takesaki Theory*. 2023. DOI: 10.48550/arXiv.2307.08129 (cit. on pp. 55, 56).

[DSV15] Irit Dinur, David Steurer, and Thomas Vidick. *A Parallel Repetition Theorem for Entangled Projection Games*. 2015. DOI: 10.48550/arXiv.1310.4113. arXiv: 1310.4113 [quant-ph] (cit. on p. 194).

[DV07] Irit Dinur and View Profile. *The PCP Theorem by Gap Amplification*. Journal of the ACM **54**, 3 (2007), 12–es. DOI: 10.1145/1236457.1236459 (cit. on pp. 16, 196).

[DFN+23] Yangjing Dong, Honghao Fu, Anand Natarajan, Minglong Qin, Haochen Xu, and Penghui Yao. *The Computational Advantage of MIP* Vanishes in the Presence of Noise*. 2023. arXiv: 2312.04360 [quant-ph] (cit. on pp. 107, 194, 197).

[DRS08] Bruno Durand, Andrei Romashchenko, and Alexander Shen. *Fixed Point and Aperiodic Tilings*. Developments in Language Theory. Ed. by Masami Ito and Masafumi Toyama. Berlin, Heidelberg: Springer, 2008, pp. 276–288. ISBN: 978-3-540-85780-8. DOI: 10.1007/978-3-540-85780-8_22 (cit. on pp. 18, 28).

[FJV+19] Joseph Fitzsimons, Zhengfeng Ji, Thomas Vidick, and Henry Yuen. *Quantum Proof Systems for Iterated Exponential Time, and Beyond*. Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing. STOC 2019. New York, NY, USA: Association for Computing Machinery, 2019, pp. 473–480. ISBN: 978-1-4503-6705-9. DOI: 10.1145/3313276.3316343 (cit. on p. 9).

[Fri12] Tobias Fritz. *Tsirelson's Problem and Kirchberg's Conjecture*. Reviews in Mathematical Physics **24**, 05 (2012), p. 1250012. DOI: 10.1142/S0129055X12500122. arXiv: 1008.1168 [math-ph, physics:quant-ph] (cit. on pp. 6, 47, 198).

[FNT14] Tobias Fritz, Tim Netzer, and Andreas Thom. *Can You Compute the Operator Norm?* Proceedings of the American Mathematical Society **142**, 12 (2014), pp. 4265–4276. DOI: 10.1090/S0002-9939-2014-12170-8. arXiv: 1207.0975 [math] (cit. on p. 195).

[Gol21] Isaac Goldbring. *The Connes Embedding Problem: A Guided Tour*. 2021. arXiv: 2109.12682 [quant-ph] (cit. on pp. 6, 165).

[GMR85] S Goldwasser, S Micali, and C Rackoff. *The Knowledge Complexity of Interactive Proof-Systems*. Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing. STOC '85. New York, NY, USA: Association for Computing Machinery, 1985, pp. 291–304. ISBN: 978-0-89791-151-1. DOI: 10.1145/22145.22178 (cit. on p. 6).

[GH17] W T Gowers and O Hatami. *Inverse and Stability Theorems for Approximate Representations of Finite Groups*. Sbornik: Mathematics **208**, 12 (2017), pp. 1784–1817. DOI: 10.1070/SM8872 (cit. on p. 89).

[Gri20] Alex B. Grilo. *A Simple Protocol for Verifiable Delegation of Quantum Computation in One Round*. 2020. DOI: 10.48550/arXiv.1711.09585. arXiv: 1711.09585 [quant-ph] (cit. on pp. 9, 16, 85).

[HK64] Rudolf Haag and Daniel Kastler. *An Algebraic Approach to Quantum Field Theory*. Journal of Mathematical Physics **5**, 7 (1964), pp. 848–861. DOI: 10.1063/1.1704187 (cit. on p. 3).

[HJX10] Uffe Haagerup, Marius Junge, and Quanhua Xu. *A Reduction Method for Noncommutative Lp-spaces and Applications*. Transactions of the American Mathematical Society **362**, 4 (2010), pp. 2125–2165. JSTOR: 40590903 (cit. on pp. 11, 45, 46, 198).

[Har04] Prahladh Harsha. *Robust PCPs of Proximity and Shorter PCPs*. Thesis. Massachusetts Institute of Technology, 2004 (cit. on pp. 16, 144).

[Har23] Bradd Hart. *An Introduction To Continuous Model Theory*. 2023. DOI: 10.48550/arXiv.2303.03969. arXiv: 2303.03969 [math.OA] (cit. on p. 194).

[Hia21] Fumio Hiai. *Quantum F-Divergences in von Neumann Algebras: Reversibility of Quantum Operations*. Mathematical Physics Studies. Singapore: Springer, 2021. ISBN: 978-981-334-198-2 978-981-334-199-9. DOI: 10.1007/978-981-33-4199-9 (cit. on pp. 163, 165).

[Hol09] Thomas Holenstein. *Parallel Repetition: Simplifications and the No-Signaling Case*. Theory of Computing **5**, 1 (2009), pp. 141–172. DOI: 10.4086/toc.2009.v005a008. arXiv: cs/0607139 (cit. on pp. 17, 167, 172).

[IV12] Tsuyoshi Ito and Thomas Vidick. *A Multi-Prover Interactive Proof for NEXP Sound against Entangled Provers*. 2012. arXiv: 1207.0550 [quant-ph] (cit. on p. 9).

[JJU+11] Rahul Jain, Zhengfeng Ji, Sarvagya Upadhyay, and John Watrous. *QIP = PSPACE*. Journal of the ACM (JACM) **58**, 6 (2011), 30:1–30:27. DOI: 10.1145/2049697.2049704 (cit. on p. 8).

[JMS20] Rahul Jain, Carl A. Miller, and Yaoyun Shi. *Parallel Device-Independent Quantum Key Distribution*. IEEE Transactions on Information Theory **66**, 9 (2020), pp. 5567–5584. DOI: 10.1109/TIT.2020.2986740. arXiv: 1703.05426 [quant-ph] (cit. on p. 169).

[JPY14] Rahul Jain, Attila Pereszlényi, and Penghui Yao. *A Parallel Repetition Theorem for Entangled Two-Player One-Round Games under Product Distributions*. 2014 IEEE 29th Conference on Computational Complexity (CCC). 2014, pp. 209–216. DOI: 10.1109/CCC.2014.29 (cit. on p. 17).

[JNV+22a] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. *MIP*=RE*. 2022. DOI: 10.48550/arXiv.2001.04383. arXiv: 2001.04383 [quant-ph] (cit. on pp. 9, 11–13, 15–17, 37, 39, 53, 66, 71–79, 82, 85, 86, 89, 92, 98, 99, 104, 107, 110, 112–115, 118, 119, 129, 132, 139, 140, 144, 145, 147, 153, 168, 193).

[JNV+22b] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. *Quantum Soundness of Testing Tensor Codes*. 2022. arXiv: 2111.08131 [quant-ph] (cit. on pp. 9, 16, 127–129, 131, 132).

[Jus72] J. Justesen. *Class of Constructive Asymptotically Good Algebraic Codes*. IEEE Transactions on Information Theory **18**, 5 (1972), pp. 652–656. DOI: 10.1109/TIT.1972.1054893 (cit. on p. 96).

[KR97a] Richard V. Kadison and John R. Ringrose. *Fundamentals of the Theory of Operator Algebras. Volume I*. American Mathematical Soc., 1997. ISBN: 978-0-8218-0819-1 (cit. on pp. 34, 46, 47).

[KR97b] Richard V. Kadison and John R. Ringrose. *Fundamentals of the Theory of Operator Algebras. Volume II*. American Mathematical Soc., 1997. ISBN: 978-0-8218-0820-7 (cit. on p. 46).

[KLV+22] Yael Kalai, Alex Lombardi, Vinod Vaikuntanathan, and Lisa Yang. *Quantum Advantage from Any Non-Local Game*. 2022. DOI: 10.48550/arXiv.2203.15877. arXiv: 2203.15877 [quant-ph] (cit. on p. 193).

[KRT09] Julia Kempe, Oded Regev, and Ben Toner. *Unique Games with Entangled Provers Are Easy*. 2009. DOI: 10.48550/arXiv.0710.0655. arXiv: 0710.0655 [quant-ph] (cit. on pp. 53, 197).

[Kho02] S. Khot. *On the Power of Unique 2-Prover 1-Round Games*. Proceedings 17th IEEE Annual Conference on Computational Complexity. 2002, pp. 25–. DOI: 10.1109/CCC.2002.1004334 (cit. on pp. 7, 53, 197).

[KPS18] Se-Jin Kim, Vern I. Paulsen, and Christopher Schafhauser. *A Synchronous Game for Binary Constraint Systems*. Journal of Mathematical Physics **59**, 3 (2018), p. 032201. DOI: 10.1063/1.4996867. arXiv: 1707.01016 [quant-ph] (cit. on p. 199).

[Kir93] Eberhard Kirchberg. *On non-semisplit extensions, tensor products and exactness of group C*-algebras*. Inventiones mathematicae **112**, 3 (1993), pp. 449–490 (cit. on p. 5).

[KM02] Hirotada Kobayashi and Keiji Matsumoto. *Quantum Multi-prover Interactive Proof Systems with Limited Prior Entanglement*. Algorithms and Computation. Ed. by Prosenjit Bose and Pat Morin. Berlin, Heidelberg: Springer, 2002, pp. 115–127. ISBN: 978-3-540-36136-7. DOI: 10.1007/3-540-36136-7_11 (cit. on p. 8).

[KMP+25] Alexander Kulpe, Giulio Malavolta, Connor Paddock, Simon Schmidt, and Michael Walter. *A Bound on the Quantum Value of All Compiled Nonlocal Games*. 2025. DOI: 10.48550/arXiv.2408.06711. arXiv: 2408.06711 [quant-ph] (cit. on p. 193).

[Lin22] Junqiao Lin. *Almost Synchronous Correlations Defined within Tracial von Neumann Algebras* (2022). HDL: 10012/18629 (cit. on pp. 9, 15, 44).

[Lin24] Junqiao Lin. *Tracial Embeddable Strategies: Lifting MIP* Tricks to MIPco*. 2024. DOI: 10.48550/arXiv.2304.01940. arXiv: 2304.01940 [quant-ph] (cit. on pp. v, 96).

[Lin26] Junqiao (Randy) Lin. *MIP^{co}=coRE*. Proceedings of the 58th Annual ACM Symposium on Theory of Computing. STOC '26. New York, NY, USA: Association for Computing Machinery, 2026, pp. 314–322. ISBN: 979-8-4007-2536-4. DOI: 10.1145/3798129.3800751 (cit. on p. v).

[Lui25] Lauritz van Luijk. *Entanglement in von Neumann Algebraic Quantum Information Theory*. PhD thesis. 2025. DOI: 10.15488/19655. arXiv: 2510.07563 [quant-ph] (cit. on p. 13).

[MSS+25] Laura Mančinska, Pieter Spaas, Taro Spirig, and Matthijs Vernooij. *Gap-Preserving Reductions and RE-completeness of Independent Set Games*. 2025. DOI: 10.48550/arXiv.2505.05253. arXiv: 2505.05253 [quant-ph] (cit. on p. 197).

[Man25a] Aareyan Manzoor. *Invariant Random Subgroups, Soficity, and Lück's Determinant Conjecture*. 2025. DOI: 10.48550/arXiv.2508.15154. arXiv: 2508.15154 [math] (cit. on p. 199).

[Man25b] Aareyan Manzoor. *There Is An Equivalence Relation Whose von Neumann Algebra Is Not Connes Embeddable*. 2025. DOI: 10.48550/arXiv.2502.06697. arXiv: 2502.06697 [math] (cit. on p. 199).

[MS24a] Kieran Mastel and William Slofstra. *Two Prover Perfect Zero Knowledge for MIP**. Proceedings of the 56th Annual ACM Symposium on Theory of Computing. STOC 2024. New York, NY, USA: Association for Computing Machinery, 2024, pp. 991–1002. ISBN: 9798400703836. DOI: 10.1145/3618260.3649702 (cit. on p. 193).

[MSZ23] Arthur Mehta, William Slofstra, and Yuming Zhao. *Positivity Is Undecidable in Tensor Products of Free Algebras*. 2023. DOI: 10.48550/arXiv.2312.05617. arXiv: 2312.05617 [math] (cit. on p. 195).

[Mer90] N. David Mermin. *Simple Unified Form for the Major No-Hidden-Variables Theorems*. Physical Review Letters **65**, 27 (1990), pp. 3373–3376. DOI: 10.1103/PhysRevLett.65.3373 (cit. on pp. 3, 92).

[MNY20] Hamoon Mousavi, Seyed Sajjad Nezhadi, and Henry Yuen. *On the Complexity of Zero Gap MIP**. 2020. DOI: 10.48550/arXiv.2002.10490. arXiv: 2002.10490 [quant-ph] (cit. on p. 193).

[MNY22] Hamoon Mousavi, Seyed Sajjad Nezhadi, and Henry Yuen. *Nonlocal Games, Compression Theorems, and the Arithmetical Hierarchy*. Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing. STOC 2022. New York, NY, USA: Association for Computing Machinery, 2022, pp. 1–11. ISBN: 978-1-4503-9264-8. DOI: 10.1145/3519935.3519949 (cit. on pp. 11, 21, 24, 40, 41, 69, 93, 140, 193, 196).

[MS24b] Hamoon Mousavi and Taro Spirig. *A Quantum Unique Games Conjecture*. 2024. DOI: 10.48550/arXiv.2409.20028. arXiv: 2409.20028 [quant-ph] (cit. on pp. 53, 197).

[MP13] Gary L. Mullen and Daniel Panario. *Handbook of Finite Fields*. 0th ed. Chapman and Hall/CRC, 2013. ISBN: 978-0-429-10519-7. DOI: 10.1201/b15006 (cit. on pp. 73, 76).

[NV17] Anand Natarajan and Thomas Vidick. *A Quantum Linearity Test for Robustly Verifying Entanglement*. Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing. STOC 2017. New York, NY, USA: Association for Computing Machinery, 2017, pp. 1003–1015. ISBN: 978-1-4503-4528-6. DOI: 10.1145/3055399.3055468 (cit. on pp. 9, 89, 198).

[NV18] Anand Natarajan and Thomas Vidick. *Two-Player Entangled Games Are NP-hard* (2018), 18 pages. DOI: 10.4230/LIPIcs.CCC.2018.20. arXiv: 1710.03062 [quant-ph] (cit. on pp. 9, 85).

[NW19] Anand Natarajan and John Wright. *NEEXP in MIP**. 2019. arXiv: 1904.05870 [quant-ph] (cit. on pp. 9, 16, 37, 38, 135, 140).

[NZ23a] Anand Natarajan and Tina Zhang. *Bounding the Quantum Value of Compiled Nonlocal Games: From CHSH to BQP Verification*. 2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS). IEEE, 2023, pp. 1342–1348 (cit. on p. 194).

[NZ23b] Anand Natarajan and Tina Zhang. *Quantum Free Games*. Proceedings of the 55th Annual ACM Symposium on Theory of Computing. Orlando FL USA: ACM, 2023, pp. 1603–1616. ISBN: 978-1-4503-9913-5. DOI: 10.1145/3564246.3585208 (cit. on p. 194).

[NPA08] Miguel Navascues, Stefano Pironio, and Antonio Acin. *A Convergent Hierarchy of Semidefinite Programs Characterizing the Set of Quantum Correlations*. New Journal of Physics **10**, 7 (2008), p. 073013. DOI: 10.1088/1367-2630/10/7/073013. arXiv: 0803.4290 [quant-ph] (cit. on pp. 9, 12, 68, 198).

[NMY25] Seyed Sajjad Nezhadi, Andrew Marks, and Henry Yuen. *The Recursive Compression Method for Proving Undecidability Results*. In preparation (2025) (cit. on pp. 11, 18, 21, 23, 27, 28).

[Nob22] *Nobel Prize in Physics 2022*. https://www.nobelprize.org/prizes/physics/2022/press-release/ (cit. on p. 2).

[OP04] M. Ohya and Denes Petz. *Quantum Entropy and Its Use*. Springer Science & Business Media, 2004. ISBN: 978-3-540-20806-8 (cit. on pp. 163, 164).

[PS25] Connor Paddock and William Slofstra. *Satisfiability Problems and Algebras of Boolean Constraint System Games*. 2025. DOI: 10.48550/arXiv.2310.07901. arXiv: 2310.07901 [quant-ph] (cit. on p. 199).

[PSZ+24] Connor Paddock, William Slofstra, Yuming Zhao, and Yangchen Zhou. *An Operator-Algebraic Formulation of Self-testing*. Annales Henri Poincaré **25**, 10 (2024), pp. 4283–4319. DOI: 10.1007/s00023-023-01378-y (cit. on p. 198).

[Pau22] Connor Paul-Paddock. *Rounding Near-Optimal Quantum Strategies for Nonlocal Games to Strategies Using Maximally Entangled States*. 2022. arXiv: 2203.02525 [quant-ph] (cit. on pp. 11, 14, 55, 56).

[PSS+16] Vern I. Paulsen, Simone Severini, Daniel Stahlke, Ivan G. Todorov, and Andreas Winter. *Estimating Quantum Chromatic Numbers*. Journal of Functional Analysis **270**, 6 (2016), pp. 2188–2222. DOI: 10.1016/j.jfa.2016.01.010 (cit. on pp. 40, 49).

[Per90] Asher Peres. *Incompatible Results of Quantum Measurements*. Phys. Lett. A **151** (1990), pp. 107–108. DOI: 10.1016/0375-9601(90)90172-K (cit. on pp. 3, 92).

[Raz95] Ran Raz. *A Parallel Repetition Theorem*. Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing. STOC '95. New York, NY, USA: Association for Computing Machinery, 1995, pp. 447–456. ISBN: 978-0-89791-718-6. DOI: 10.1145/225058.225181 (cit. on pp. 167, 169).

[Sch80] J. T. Schwartz. *Fast Probabilistic Algorithms for Verification of Polynomial Identities*. J. ACM **27**, 4 (1980), pp. 701–717. DOI: 10.1145/322217.322225 (cit. on p. 76).

[Sha90] A. Shamir. *IP=PSPACE (Interactive Proof=polynomial Space)*. Proceedings [1990] 31st Annual Symposium on Foundations of Computer Science. 1990, 11–15 vol.1. DOI: 10.1109/FSCS.1990.89519 (cit. on p. 7).

[Sip06] Michael Sipser. *Introduction to the Theory of Computation*. Thomson Course Technology, 2006. ISBN: 978-0-619-21764-8 (cit. on p. 22).

[Slo19a] William Slofstra. *Tsirelson's Problem and an Embedding Theorem for Groups Arising from Non-Local Games*. Journal of the American Mathematical Society **33**, 1 (2019), pp. 1–56. DOI: 10.1090/jams/929. arXiv: 1606.03140 [math-ph, physics:quant-ph] (cit. on p. 193).

[Slo19b] William Slofstra. *The Set of Quantum Correlation Is Not Closed*. Forum of Mathematics, Pi **7** (2019/ed), e1. DOI: 10.1017/fmp.2018.3 (cit. on pp. 3, 41).

[Tak70] M. Takesaki. *Tomita's Theory of Modular Hilbert Algebras and Its Applications*. Vol. 128. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer, 1970. ISBN: 978-3-540-04917-3 978-3-540-36267-8. DOI: 10.1007/BFb0065832 (cit. on pp. 17, 45, 159).

[Tak01] M. Takesaki. *Theory of Operator Algebras I*. Springer Science & Business Media, 2001. ISBN: 978-3-540-42248-8 (cit. on pp. 34, 159, 199).

[Tsi93] B. S. Tsirelson. *Some Results and Problems on Quantum Bell-type Inequalities*. Hadronic J. Suppl. **8**, 4 (1993), pp. 329–345 (cit. on p. 3).

[Vid18] Thomas Vidick. *An Expository Note on "a Quantum Linearity Test for Robustly Verifying Entanglement"*. 2018 (cit. on pp. 89, 90).

[Vid20] Thomas Vidick. *A Masters Project*. 2020 (cit. on p. 9).

[Vid22] Thomas Vidick. *Almost Synchronous Quantum Correlations*. Journal of Mathematical Physics **63**, 2 (2022), p. 022201. DOI: 10.1063/5.0056512. arXiv: 2103.02468 (cit. on pp. 11, 14, 49, 50, 55, 56, 60, 100, 132).

[VW16] Thomas Vidick and John Watrous. *Quantum Proofs*. Foundations and Trends® in Theoretical Computer Science **11**, 1-2 (2016), pp. 1–215. DOI: 10.1561/0400000068. arXiv: 1610.01664 [quant-ph] (cit. on p. 66).

[Wat99] John Watrous. *PSPACE Has 2-Round Quantum Interactive Proof Systems*. 1999. DOI: 10.48550/arXiv.cs/9901015. arXiv: cs/9901015 (cit. on p. 7).

[Wat22] John Watrous. *Advanced Topics in Quantum Information Theory Lecture 8: The Hierarchy of Navascués, Pironio, and Acín*. https://jhwatrous.github.io/QIT-notes.08.pdf. 2022 (cit. on p. 68).

[WHK23] Adam Bene Watts, John William Helton, and Igor Klep. *Noncommutative Nullstellensätze and Perfect Games*. Annales Henri Poincaré **24**, 7 (2023), pp. 2183–2239. DOI: 10.1007/s00023-022-01262-1. arXiv: 2111.14928 [quant-ph] (cit. on p. 195).

[Wil13] Mark M. Wilde. *Quantum Information Theory*. Cambridge: Cambridge University Press, 2013. DOI: 10.1017/CBO9781139525343 (cit. on pp. 162, 166).

[Zip79] Richard Zippel. *Probabilistic Algorithms for Sparse Polynomials*. Symbolic and Algebraic Computation. Ed. by Edward W. Ng. Berlin, Heidelberg: Springer, 1979, pp. 216–226. ISBN: 978-3-540-35128-3. DOI: 10.1007/3-540-09519-5_73 (cit. on p. 76).

Nomenclature

Sets, strings and probability distribution.

$|S|$: Cardinality of a set. 33
$|s|_w$: Length of the string s. 21
$\mathbb{E}_{x\sim\mu}$: Expectation over the distribution μ. 33
$\mathcal{M}_n(T)$: n by n matrix of elements of T. 33
$\kappa(\mu)$: The inverse of the spectral gap for the distribution μ over a group. 91
$\mathrm{supp}(\mu)$: The support for the distribution μ. 90
$\mathsf{P}_{Y|X=x}(y)$: Probability of Y conditioning on X. 170
$\|\mathsf{P}_{X_0} - \mathsf{P}_{X_1}\|$: The variation distance between probability distribution P_{X_0} and P_{X_1}, given in (12.14). 170
$\delta_{a,b}$: The delta kronecker product. 33
$s \cdot t$: Dot product between s and t for string s and t. 33
$|s|$: Hamming weight for the string s. 33
$s|_a$: Coordinate of s index by a. 33
$\pi_{>j}(s)$: The map which zeros out the first j entries of the string s. 33
$[n,m]$: The set $\{n, n+1\cdots, m-1\}$. 21
$[n]$: The set $\{0, 1\cdots, n-1\}$. 21
$\mathbf{bin}(n)$: Binary representation for the integer n. 21
$\mathbf{bininv}(s)$: The inverse binary function, i.e. $\mathbf{bininv}(s) = m$ where m is the unique integer such that $\mathbf{bin}(m) = s$ for $s \in \{0,1\}^n$. 21

Turing machines, complexity classes and algorithms.

$\langle \mathtt{A}(x)\rangle$: The description of the Turing machine $\mathtt{A}$ which is hardcoded to run $x \in \{0,1\}^*$ as input (in this case $\langle \mathtt{A}(x)\rangle$ takes the empty tape as input, and will return $\mathtt{A}(x)$) after the computation step.. 21
$\langle \mathtt{A}\rangle$: The minimial description lenght of a Turing machine $|\mathtt{A}|$. 21
$|\mathtt{A}|$: The minimial description lenght of a Turing machine $|\mathtt{A}|$. 21
$\mathsf{TIME}_{\mathtt{A}}(n)$: The maximum of the runtime and decription size for the Turing machine $\mathtt{A}$. 22
coRE : The complement of RE, complete with respect to the non-halting problem.. 22
coD : The complement of the decision problem D.. 22
RE : The set of recursively enumerable languages, complete with respect to the halting problem.. 22
$\mathsf{D}_1 \leq_p \mathsf{D}_2$: D_1 is polynomial-time reducible to D_2 using the standard Karp reduction.. 22
$\mathsf{D}_1 \leq \mathsf{D}_2$: D_1 is reducible to D_2 using the standard Karp reduction.. 22
MIP^* : Multiprover interactive proof system with tensor product model of entanglement (two round, one prover with completness 1 and soundness $\frac{1}{2}$). 66
MIP^{co} : Multiprover interactive proof system with commuting opereator model of entanglement (two round, one prover with completness 1 and soundness $\frac{1}{2}$). 66

$\texttt{searchfrombelow}_\varepsilon$: The search from below algorithm for $\varepsilon \in [0,1]$, Teminates whenever $\omega^*(\mathcal{G}) > \varepsilon$ (Runs forever otherwise). 70

$\texttt{searchfromabove}_\delta$: The search from above algorithm for $\delta \in [0,1]$, Terminates whenever $\omega^{co}(\mathcal{G}) < \delta$ (Runs forever otherwise). 69

🖃 : The input for the Turing machine for the proof of RE/coRE completeness. This is the only non-western character used in this thesis. 25

Finite fields.

$\mathbb{F}_{2^p}$: Finite fields over 2^p, p is always assume to be odd in this thesis. 73

$\hat{e}_i$: Canoical basis for for the field $\mathbb{F}_{2^p}$. 73

$\mathrm{Tr}(a)$: Finite field trace for the element $a \in \mathbb{F}_{2^p}$. 73

$\kappa(a)$: The bijection map between $\mathbb{F}_{2^p}$ to $\{0,1\}^p$ whenever $a \in \mathbb{F}_{2^p}$. The bijection map between $\mathbb{F}_{2^p}^m$ to $\{0,1\}^{pm}$ whenever $a \in \mathbb{F}_{2^p}^m$. 74

$\mathbf{dim}(V)$: Dimension of the subspace $V \subseteq \mathbb{F}_{2^p}^m$. 74

$W^\perp$: The orthogonal subspace of W for $W \subseteq V \subseteq \mathbb{F}_{2^p}^m$. $W^\perp$ is the orthogonal subspace over $\mathbb{F}_{2^p}^m$ if unspecified. 74

W^C : The canonical complement of W for $W \subseteq W \subseteq \mathbb{F}_{2^p}^m$ for a canoical basis subspace V. W^C is the canoical complement over $\mathbb{F}_{2^p}^m$ if unspecified. 75

$V_{<i}$: The union of the first i subspace in a disjoint partition of V. 74

$\pi^m_{>j}$: The map which zeros out the first j entries for elements of $\mathbb{F}_{2^p}^m$. 75

Functions on finite fields.

$\ker(\mathtt{L})$: Kernel subspace for a linear function $\mathtt{L}$. 75

$\mathtt{L}^\perp$: The Linear map which projects onto $\left(\ker(\mathtt{L})^\perp\right)^C$ where $\mathtt{L} : V \to V$ is a linear function over a canoical basis subspace V. 75

$\mathbf{Can}(l)$: Canonical representation of an affine line, define as $\mathrm{Can}(l) := (v, \mathrm{Null}^{\mathrm{LN}}_v(u)) \in \mathbb{F}_{2^p}^{2m}$. 76

$\mathbf{IdPoly}(p,m,d)$: The set of polynomails $\mathbf{g} : \mathbb{F}_{2^p}^m \to \mathbb{F}_{2^p}$ with individual degree of at most d. 76

$\mathbf{RM}_b$: Reed-Muller encoding for the string b. 128

$\eta_{\mathbf{LD}}(p,m,d,\varepsilon)$: The soundness parameter function for the quantum low-individual degree test, given in Theorem 10.3.1. 132

$\eta_{\mathbf{SLD}}(p,m,d,k,\varepsilon)$: The soundness parameter function for the (p,m,d,k)-simultaneous quantum low-individual degree test, given in Lemma 10.4.1. 135

von Neumann algebras.

$\mathscr{H}$: Hilbert space. 34

$\mathscr{A}$: von Neumann algebras. 34

$\mathscr{B}(\mathscr{H})$: Bounded operator acting on $\mathscr{H}$. 34

$||\psi\rangle|$: Vector norm. 34

$\mathscr{A}'$: Commutant of $\mathscr{A}$. 34

$\mathscr{A}^+$: Set of positive elements within $\mathscr{A}$. 34

$\mathrm{Tr}(\cdot)$: Normalized trace for finite dimensional matrix. 35

τ : Trace function (often assocated with a von Neumann algebra $\mathscr{A}$). 35

$||A||_2$: Hilbert schmidt norm for $A \in \mathscr{A}$, where $\mathscr{A}$ is a tracial von Neumann algebra. 35

$||A||_\sigma$: The state dependent norm for a density matrix σ (i.e. $\sigma \in \mathscr{A}^*$ and $\tau(\sigma) = 1$) and $A \in \mathscr{A}$, where $(\mathscr{A}, \tau)$ is a tracial von Neumann algebra. 89

$\|\psi\|$: State norm for the state ψ. 34

$\mathscr{L}^2(\mathscr{A}, \tau)$: Hilbert space for the standard form of $\mathscr{A}$, where $(\mathscr{A}, \tau)$ is a tracial von Neumann algebra. 35

$|\tau\rangle$: Vector state associated to the trace τ for the standard form of $\mathscr{A}$, where $(\mathscr{A}, \tau)$ is a tracial von Neumann algebra. 35

a^{op} : The bijection map between $\mathscr{A}$ in standard form to $\mathscr{A}'$ through the opposite algebra map. 35
$\mathscr{A}^{op}$: The opposite algebra of $\mathscr{A}$. 35
$\mathscr{A} \rtimes_{\boldsymbol{\alpha}} \mathfrak{G}$: The crossed product of $\mathscr{A}$ with the continous automorphism $\boldsymbol{\alpha} : \mathfrak{G} \to \mathrm{Aut}(\mathscr{A})$. 45
$\mathscr{H}^{+}_{|\tau\rangle}$: The canonical positive cone for the von Neumann algebra $\mathscr{A}$ in standard form. 160

Quantum measurements.

$A_{[f|s]}$: Data processing measurement. 36
$\rho_a^{W,p}$: The PVM with outcome a associated with Generalized Pauli meausrement over $\mathbb{F}_{2^q}$ for $W \in \{X, Z\}$. 91
$\rho^{W,p}(a)$: The observable associated with Generalized Pauli meausrement over $\mathbb{F}_{2^p}$ for $W \in \{X, Z\}$ over $a \in \mathbb{F}_{2^p}$, also can be define using $s \in \mathbb{F}_{2^p}^m$, given by (8.6). 91
$\{\rho_s^X\}_{s\in V}$: Generalized Pauli measurement with outcome over a register subspace $V \subseteq \mathbb{F}_{2^p}^m$. 92
$U_{2\to p}$: The Unitary which converts between the generalized (qubit) Pauli measurement over 2^p to the one qubit Pauli measurement given in Lemma 8.2.1. 92
$\approx_\delta$: δ-close when the underlying state and distribution is clear. δ-close is given by (3.6). 37
$\simeq_\delta$: δ-consistant when the underlying state and distribution is clear. δ-consistant is given by (3.5). 37

Quantum correlations and strategies.

$C_{x,y,a,b}$: Correlations, the subscipt x, y, a, b are often omitted. 38
C_q : The set of all quantum tensor product correlations. $C_q(\mathscr{X}, \mathscr{A})$ if the questionset and the asnwer set are specifed, sometime written as C_*. 39
C_q^n : The set of all quantum tensor correlations realizable in $\mathscr{M}_n(\mathscr{C}) \otimes \mathscr{M}_n(\mathscr{C})$. 39
C_{qc} : The set of all quantum commuting operator correlations. $C_{qc}(\mathscr{X}, \mathscr{A})$ if the questionset and the asnwer set are specifed. 39
C_t^s : The set of all quantum syncronous correlations under model $t \in \{*, co\}$. 39
$\delta_{\mathbf{sync}}(\mu, C)$: The synchronicity for the correlation C under the discribution μ, sometimes written as $\delta_{\mathrm{sync}}(\mu, \mathscr{S})$ for quantum strategy instead. 55

Non-local games.

$\mathscr{G}$: Non-local games, often denoted with $\mathscr{G} = (\mathscr{X}, \mathscr{A}, \mu, D)$ where $\mathscr{X}$ is the question set, answer set $\mathscr{A}$, question distribution μ and validation function D. 40
$\mathscr{G}^{\mathbf{accept}}$: The accepting syncronous game, with $\omega^*(\mathscr{G}^{\mathrm{accept}}) = \omega^{co}(\mathscr{G}^{\mathrm{accept}}) = 1$. 68
$\mathscr{G}^{\mathbf{reject}}$: The rejecting syncronous game, with $\omega^*(\mathscr{G}^{\mathrm{reject}}) = \omega^{co}(\mathscr{G}^{\mathrm{reject}}) = \frac{1}{3}$. 68
$\mathscr{G}_\perp$: Oracularization transformation for the game $\mathscr{G}$. 141
$\mathscr{G}_\perp$: Anchoring transformation for the game $\mathscr{G}$. 168
$\mathscr{G}^{\otimes r}$: r-fold parallel repetition of a game $\mathscr{G}$. 167
$(\vec{x}, \vec{y})$: Question pair for r-fold parallel repetition. 168
$(\vec{a}, \vec{b})$: Answer pair for r-fold parallel repetition. 168
$\omega(\mathscr{G}, C)$: The value of the game under correlation C or strategy $\mathscr{S}$. 40
$\omega^*(\mathscr{G})$: Tensor product value of $\mathscr{G}$. 41
$\omega^{co}(\mathscr{G})$: Commuting operator value of $\mathscr{G}$. 41
$\omega_s^t(\mathscr{G})$: Syncronous value. 41

Conditional Linear functions and Conditional linear verifier.

$\mathtt{L}$: Conditionally linear function. 76
$\mathtt{L}_{j,s}$: The jth level CL function used to define $\mathtt{L}$ when the previous function returns s.. 76

$\mathtt{L}^P$: The CL function which is used to define a CL distribution (Definition 7.2.5) for $P \in \{A, B\}$. Same notation ($\mathtt{L}^V$) is used to define a typed CL distribution (Definition 7.2.7) for $v \in \mathtt{T}$. 78

$\mathbf{neigh}_{\mathtt{E}}(v)$: Indicator vector for $v \in \mathtt{T}$, where $(\mathtt{T}, \mathtt{E})$ is a graph. 79

$\mathscr{V}$: A CL Verifier sequence for a $\mathsf{MIP}^*/\mathsf{MIP}^{\mathsf{co}}$ protocol, as specified in Definition 7.3.1. 83

$\mathtt{Q}_{\mathscr{V}}$: The sampler for $\mathscr{V}$, as specified in Definition 7.3.1. 83

$\mathtt{D}_{\mathscr{V}}$: The decider for $\mathscr{V}$, as specified in Definition 7.3.1. 83

$\mathbf{k}(n)$: The level function for a CL verifier. 83

$\mathbf{m}(n)$: The cardinality function for a CL verfier. 83

$\mathbf{p}(n)$: The field size function for a CL verfier. 83

Notations for the paralllel repetition theorem.

$\psi^{\mathscr{A}_1\mathscr{A}_2}$: A state define within $\mathscr{A}_1 \otimes \mathscr{A}_2$, also notation for the restriction of $\mathscr{A}_1 \otimes \mathscr{A}_2$ if the state is defined in a bigger Hilbert space. 159

$\phi^{\mathscr{X}\mathscr{A}}$: Classical quantum state with the classical component being $\mathscr{X}$. 165

$\mathrm{D}(\psi_1 \| \psi_2)$: Relative entropy betweenm the state ψ_1 and ψ_2. 162

$\mathrm{I}(\mathscr{A}_1 : \mathscr{A}_2)_\psi$: Mutual information between algebra $\mathscr{A}_1$ and $\mathscr{A}_2$ for the state ψ. 164

$\eta_{\mathbf{Anchor}}$: The noise parameter. 171

C : The critical set given by Proposition 12.3.3. 171

$\eta_{\mathbf{PR}}$: The constant given in Lemma 12.3.9. 172

$\mathbf{R}_C$: The question/answer correlation for coordinates in the critical set C, consist of question distribution $\mathbf{Q}_C = (\mathbf{X}_C, \mathbf{Y}_C)$ and answer distribution $\mathbf{S}_C = (\mathbf{A}_C, \mathbf{B}_C)$. 172

$\mathbf{\Omega}$: The dependence breaking probability distribution define for coordinates outside of the critical set, given in Definition 12.3.7. We further use $\mathbf{\Omega}_{-i}$ to denote the distribution without the coordiate i. 172

$|\Phi_{(\omega_{-i},\vec{r}_C),s,y}\rangle$: The (unnormalized) post measurement state $|\psi\rangle$ with measurements condition on $\mathbf{\Omega}_{-i} = \omega_{-i}$, $\mathbf{R}_C = \vec{r}_C$, s and y are either the normal measurement, or the slanted measuremnt given in (12.21) if s or y are $\perp/x$. 173

$\gamma_{(\omega_{-i},\vec{r}_C),s,y}$: The normalized factor for the state $|\Phi_{(\omega_{-i},\vec{r}_C),s,y}\rangle$. 173

$|\widetilde{\Phi}_{(\omega_{-i},\vec{r}_C),s,y}\rangle$: Normalized version of $|\Phi_{(\omega_{-i},\vec{r}_C),s,y}\rangle$. 174

$\Xi^{\mathbf{\Omega X}_C\mathbf{Y}\mathbf{Q}_C\mathscr{A}}$: The classical quantum state which packages all of the post-measurement state from the strategy $\mathscr{S}^{\otimes r}$, with outcome being restricted to $(\mathbf{a}_C, \mathbf{b}_C) \in \mathbf{S}_C$. For each index, Alice's measurement is being condition on $\Omega = \omega$ and $\mathbf{Q}_C = (\vec{x}_C, \vec{y}_C)$, Bob's mesurement only depends on $\vec{y}$. 179

$\xi^{\mathbf{\Omega X}_C\mathbf{Y}\mathbf{Q}_C\mathscr{A}}$: $\Xi^{\mathbf{\Omega X}_C\mathbf{Y}\mathbf{Q}_C\mathscr{A}}$ being additionally conditioning on $(\vec{x}_C, \vec{y}_C, \vec{a}_C, \vec{b}_C)$ giving a winning question/answer on all coordinates in the critical set C. 179

$\Lambda^{\mathbf{\Omega}\mathbf{X}\mathbf{Y}_C\mathbf{Q}_C\mathscr{A}}$: The classical quantum state which packages all of the post-measurement state from the strategy $\mathscr{S}^{\otimes r}$, with outcome being restricted to $(\mathbf{a}_C, \mathbf{b}_C) \in \mathbf{S}_C$. For each index, Bob's measurement is being condition on $\Omega = \omega$ and $\mathbf{Q}_C = (\vec{x}_C, \vec{y}_C)$, Alice's mesurement only depends on $\vec{x}$. 179

$\lambda^{\mathbf{\Omega}\mathbf{X}\mathbf{Y}_C\mathbf{Q}_C\mathscr{A}}$: $\lambda^{\mathbf{\Omega X}_C\mathbf{Y}\mathbf{Q}_C\mathscr{A}}$ being additionally conditioning on $(\vec{x}_C, \vec{y}_C, \vec{a}_C, \vec{b}_C)$ giving a winning question/answer on all coordinates in the critical set C. 179