

Hypercontractivity for Quantum Erasure Channels via Variable Multipartite Log-Sobolev Inequality

Zongbo Bao* Yangjing Dong† Fengning Ou‡ Penghui Yao§¶

January 21, 2026

Abstract

We prove an almost optimal hypercontractive inequality for products of quantum erasure channels, generalizing the hypercontractivity for classical binary erasure channels. To our knowledge, this is the first tensorization-type hypercontractivity bound for quantum channels with no fixed states. The traditional inductive arguments for classical hypercontractivity cannot be generalized to the quantum setting due to the nature of the non-commutativity of matrices. To overcome the difficulty, we establish a novel quantum log-Sobolev inequality for Bernoulli entropy, which includes the classical log-Sobolev inequality and the quantum log-Sobolev inequality as one-partite cases. To our knowledge, its classical counterpart is also unknown prior to this work. We establish a connection between our quantum log-Sobolev inequality and the hypercontractivity bound for quantum erasure channels via a refined quantum Gross' lemma, extending the analogous connection between the quantum log-Sobolev inequality and the hypercontractivity for qubit unital channels. As an application, we prove an almost tight bound (up to a constant factor) on the classical communication complexity of two-party common randomness generation assisted with erased-noisy EPR states, generalizing the tight bound on the same task assisted with erased-noisy random strings due to Guruswami and Radhakrishnan.

1 INTRODUCTION

The notion of *hypercontractivity* originates from quantum field theory [Nel66, Gli68, Nel73, SHK72] and functional analysis [Pal32, Bon68, Kie69, Bon70], which has a rich history of research(see [O'D14,

*State Key Laboratory for Novel Software Technology, New Cornerstone Science Laboratory, Nanjing University, Nanjing, China. Email: baozb0407@gmail.com.

†State Key Laboratory for Novel Software Technology, New Cornerstone Science Laboratory, Nanjing University, Nanjing, China. Email: dongmassimo@gmail.com.

‡State Key Laboratory for Novel Software Technology, New Cornerstone Science Laboratory, Nanjing University, Nanjing, China. Email: reverymoon@gmail.com.

§State Key Laboratory for Novel Software Technology, New Cornerstone Science Laboratory, Nanjing University, Nanjing, China. Email: phyao1985@gmail.com.

¶Hefei National Laboratory, Hefei 230088, China.

Notes in Chapter 9] for more background and history). Hypercontractive inequalities nowadays are also a fundamental tool in the analysis of Boolean functions, which found wide applications in various areas of theoretical computer science [KKL88, LMN93, MOO10, KKMO07, MOO10].

A typical hypercontractive inequality is the Bonami-Beckner inequality [Bon68, Bec75]. For a fixed $\rho \in [0, 1]$, consider the noise operator on the space of all functions $f : \{0, 1\}^n \rightarrow \mathbb{R}$, defined by

$$(T_\rho(f))(x) = \mathbb{E}_y[f(y)],$$

where the expectation is taken over y obtained from x by flipping each bit independently with probability $(1-\rho)/2$. Intuitively, T_ρ “smooths” the function f . That is, the peaks of f are smoothed out in $T_\rho(f)$. Consider the p -norm of f : $\|f\|_p = (\mathbb{E}_x[|f(x)|^p])^{1/p}$, where the expectation is over the uniform distribution on $\{0, 1\}^n$. The p -norm is monotone non-decreasing with respect to p . The hypercontractive inequality captures the smoothing effect of T_ρ in terms of p -norms. More precisely, we have

$$\|T_\rho(f)\|_q \leq \|f\|_p, \tag{1}$$

as long as $1 \leq p \leq q$ and $\rho \leq \sqrt{(p-1)/(q-1)}$. It has been further generalized to the non-binary and non-uniform settings [Wol07].

Hypercontractive inequalities in non-commutative settings have received increasing attention since the quantum computing and quantum information technologies emerge [Gro75a, CL93, OZ99, BARdW08, KT13, Kin14, FHL⁺17, Bei21]. To put the classical hypercontractivity in the context of quantum computing, let’s reformulate a Boolean function $f : \{0, 1\}^n \rightarrow \mathbb{R}$ as a $2^n \times 2^n$ diagonal matrix with diagonal entries $f(x)$, i.e., $M_f = \sum_{x \in \{0, 1\}^n} f(x) |x\rangle\langle x|$. Then T_ρ is equivalent to a product of quantum depolarizing channels acting on M_f , and $\|f\|_p$ is the normalized Schatten p -norm of M_f , which will be defined later. Thus it is natural to consider the hypercontractivity bound for the product of depolarizing channels acting on an arbitrary operator. The main challenge in the non-commutative case is the tensorization property, i.e., to generalize the hypercontractive inequality for a product of n quantum channels, from the hypercontractive inequality for a local single copy quantum channel, which is easy in the classical commutative case. Kastoryano and Temme [KT13] proved a hypercontractive inequality for the product of depolarizing channels. Furthermore, King [Kin14] generalized this result to the products of all *unital* qubit channels, which map identity matrices to identity matrices. Recently, Beigi [Bei21] further improved the hypercontractivity bound for the product of depolarizing channels with better parameters. Researchers have also established several variants of non-commutative hypercontractive inequalities, which have found applications in quantum communication complexity [GKK⁺08], quantum coding [BARdW08, AD24], quantum Markov semigroups [BR22] and quantum non-local games [QY21, QY23], etc.

Besides the hypercontractivity for product of unital channels studied in [Kin14, Bei21], there are some results concerning the hypercontractivity on non-unital channels in a different point of view [OZ99, TPK14, MHSFW16]. They focused on the hypercontractive inequalities for an evolution under L_p norm with respect to the fixed state of the evolution, a.k.a. the stationary state of the evolution.

In this paper, we focus on the *quantum erasure channel* (QEC). A QEC acts on the input qubit

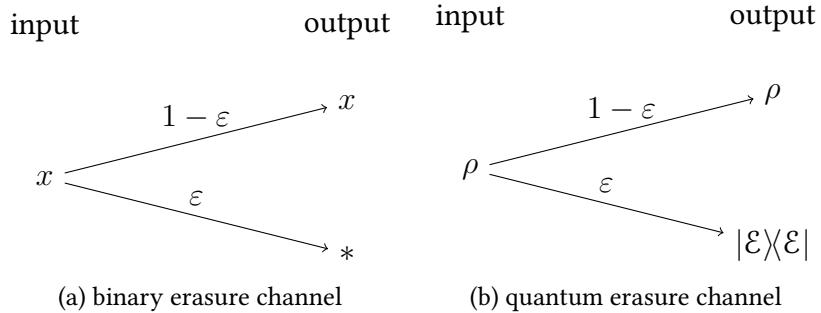


Figure 1: Classical and quantum erasure channel. With probability ε , error occurs and the input bit/state is erased. In the quantum erasure channel, the state $|\mathcal{E}\rangle\langle\mathcal{E}|$ indicating error replaces ρ in case of erasure.

and outputs an error with probability ε , and leaves the input state unchanged with probability $1 - \varepsilon$. Formally, given an input qubit ρ , we define the quantum erasure channel \mathcal{D}_ε as

$$\mathcal{D}_\varepsilon(\rho) = (1 - \varepsilon) \cdot \rho + \varepsilon \cdot |\mathcal{E}\rangle\langle\mathcal{E}|. \quad (2)$$

In case of the error (with probability ε), the input state ρ is replaced by the state $|\mathcal{E}\rangle\langle\mathcal{E}|$ indicating an erasure error. See Fig. 1 for an illustration. The QEC is one of the most fundamental quantum channels, which have received extensive studies. There have been various quantum error correcting codes proposed for the quantum erasure channel [GBP97, DZ13, KCP16, DZ20, ZOJ23]. Besides, they are one of the few quantum channels whose asymptotic capacity for faithful transmission can be computed exactly [BDS97].

In this paper, we establish the first hypercontractivity bound for the tensor product of QECs. It is interesting to note that an almost optimal hypercontractivity bound for quantum depolarizing channels can be obtained from the hypercontractivity for QECs established in this paper¹. The hypercontractivity for classical binary erasure channels has been proved by Nair and Wang [NW16] recently, which have found applications in common randomness generation [GR16] and property testing [BP21]. Recently, Eldan, Wigderson and Wu [EWW23] have independently proved a hypercontractive inequality for the binary erasure channel (in a different language), which was used to give a sharp version of the “it ain’t over till it’s over” theorem [MOO10].

QECs are non-unital and do not have fixed points as the input space and output space have different dimensions. Thus they are not covered by existing conclusions from [Kin14, Bei21, OZ99, TPK14, MHSFW16]. To our knowledge, this is the first quantum channels with no fixed states, whose tensorized hypercontractivity is established.

The most well-known tensorization proofs for classical hypercontractivities are via the inductive argument on the length of the input. For quantum erasure channels, by combining the inductive argument and the norm compression inequality, we could solve the case $1 \leq p \leq 2 \leq q$. However, due to the nature of the norm compression for the case $1 \leq p \leq q \leq 2$, this inductive argument fails to reach any reasonable conclusion. To remedy it, the authors in [OZ99, Kin14, Bei21]

¹See Appendix C for a proof

use a new framework for quantum hypercontractivity, which turns out to be very successful for several quantum channels. A crucial component in this framework is the equivalence between the *log-Sobolev inequality* and hypercontractive inequalities via the *Gross' lemma* [Gro75a, KT13].

Log-Sobolev inequalities are the functional inequalities that upper bound the entropy of a function by their Dirichlet forms. Classical log-Sobolev inequalities are amongst the most studied functional inequalities for semigroups [BH99, FMP13, IM14, CM15, BB21, FOW22], which turn out to be a versatile tool to analyze mixing time of various Markov processes [BGL⁺14, CGM21, CCNW21]. Olkiewicz and Zegarlinski [OZ99] introduced the log-Sobolev inequality in a general non-commutative setting, which is further studied as *quantum log-Sobolev inequalities* on finite dimensional state spaces [KT13, TPK14]. In quantum log-Sobolev inequalities, the entropy is defined on all quantum operators, using a generalized definition of the entropy on functions in the commutative case. This is in contrast with the standard quantum information theory definition of entropy defined primary for quantum states. Quantum log-Sobolev inequalities have played a crucial role in the proof of hypercontractivity bounds for unital qubit channels [Kin14] and have also found applications in the study of quantum systems [CM20, BCG⁺23], noisy quantum devices [SFGP21], quantum Gibbs samplers [KB16], quantum Markov semigroups [BR22], etc.

To our knowledge, all previous classical or quantum log-Sobolev inequalities are concerned with the systems with a fixed size (i.e., number of bits in the classical system or the number of qubits in the quantum system). However, the size of the system varies when considering BECs because of the extra error dimension $|\mathcal{E}\rangle$ in the output system. Hence, we establish a *quantum log-Sobolev inequality for Bernoulli entropy*, where the size of the system varies subject to erasure noise. It includes Kastoryano and Temme's quantum log-Sobolev inequality [KT13] as a single partite case. Our quantum log-Sobolev inequality for Bernoulli entropy is essential to the proof of our QEC hypercontractivity, and is of independent interest as well. To our knowledge, even the classical analog of our log-Sobolev inequality is unknown. Additionally, the quantum Gross' lemma [Gro75a, Kin14], which bridges the hypercontractivity and log-Sobolev inequalities, is replaced by a refined inequality connecting the Schatten p norm of an operator and the Schatten p norm of its reduced operator. This lemma is used to induce a hypercontractivity bound for the QEC from the quantum log-Sobolev inequality for Bernoulli entropy.

As an application of our hypercontractive inequality for the QEC, we prove an almost tight bound (up to a constant factor) on the classical communication complexity of the two-party common randomness generation problem, where maximally entangled states affected by the QEC are shared among the communicating parties. Common randomness generation, a classic topic in information theory and distributed computing, was first raised by Maurer [Mau93] and Ahlswede and Csiszár [AC98], leading to numerous subsequent works. Devetak and Winter [DW04] introduced the problem of entanglement-assisted common randomness generation, and studied the case of classical-quantum correlations. Recently, Lami, Regula, Wang and Wilde [LRWW23] gave efficiently computable upper bounds on the LOCC assisted distillable randomness of a bipartite quantum state. Also, an upper bound in terms of the fidelity-based smooth min-relative entropy was given by Nuradha and Wilde [NW24]. The idea of these two results are strongly linked to those in [DKQ⁺23]. Readers may refer to the excellent survey [STW20] and the references therein.

Most of these works assume that the players share only a limited amount of noisy correlation. In contrast, in this work we assume the noisy correlations are free and unlimited, and try to minimize the classical communication. Guruswami and Radhakrishnan [GR16] investigated the communication complexity of common randomness generation when the parties share unlimited but non-perfect correlation. They considered both the binary symmetric channel and the binary erasure channel and established tight tradeoffs among the communication complexity, min-entropy of the common randomness and errors for both cases. Dong and Yao [DY24] investigated common randomness generation when the players share unlimited depolarized noisy EPR states, and they also established a tight tradeoff among the classical communication complexity, min-entropy of the common randomness and errors. They also establish a tradeoff when quantum communication is allowed, which is not known to be tight. However, the communication complexity of common randomness generation when the players sharing maximally entangled states effected by the quantum erasure channel remains unknown. For the quantum erasure channel, this is due to a lack of a tensorized hypercontractivity bound, which is a key component of all these previous proofs on the communication complexity of common randomness generation. In this work, by using our hypercontractivity bound for the product of quantum erasure channel, we are able to prove a communication lower bound when the players share maximally entangled states affected by the quantum erasure channel. This lower bound matches the classical communication upper bound of [GR16], where the players share classical correlation with the binary erasure channel, and thus is tight up to constant a factor.

1.1 OUR RESULTS

Hypercontractivity Our main result is a hypercontractivity bound for the product of quantum erasure channels. Before illustrating our main result, we first state the hypercontractivity bound for the classical binary erasure channel proved by Nair and Wang [NW16]. A similar inequality has been independently proved by Eldan, Wigderson and Wu [EW23].

Theorem 1.1 ([NW16]). *Let $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ be a function, X be a random variable uniformly distributed over $\{-1, 1\}^n$ and $Y = \text{BEC}_\varepsilon(X)$. Define the function $g : \{-1, 1, *\}^n \rightarrow \mathbb{R}$ as*

$$g(y) = \mathbb{E}[f(X) \mid Y = y]. \quad (3)$$

Then for any $\varepsilon \in [0, 1]$, $p, q > 1$ satisfying $1 - \varepsilon \leq (p - 1)/(q - 1)$, it holds that

$$\|g(Y)\|_q \leq \|f(X)\|_p \quad (4)$$

where the λ -norm is defined as

$$\|Z\|_\lambda = \mathbb{E}\left[|Z|^\lambda\right]^{\frac{1}{\lambda}}. \quad (5)$$

To express our hypercontractivity bounds, we need to decompose a QEC into two operations:

1. Given a 2×2 matrix ρ , we first define an expanding operator \mathcal{D} that creates both the “non-erased” state and the “erased” state, that is

$$\mathcal{D}(\rho) = \rho + \tau(\rho) \cdot |\mathcal{E}\rangle\langle\mathcal{E}| = \begin{pmatrix} \rho & 0 \\ 0 & \tau(\rho) \end{pmatrix}, \quad (6)$$

where $\tau(\cdot)$ stands for the normalized trace operation, and here $\tau(\rho) = \frac{1}{2} \text{Tr } \rho$. This is the analog of the function $g(y) = \mathbb{E}[f(X)|Y = y]$.

2. Then we define the noise matrix Π_ε as

$$\Pi_\varepsilon = (1 - \varepsilon) \cdot \mathbb{1}_2 + 2\varepsilon \cdot |\mathcal{E}\rangle\langle\mathcal{E}| = \begin{pmatrix} 1 - \varepsilon & & \\ & 1 - \varepsilon & \\ & & 2\varepsilon \end{pmatrix}. \quad (7)$$

This matrix reassigned the noise, which simulates taking expectation over Y in the classical case.

Recall that in Eq. (2) we define

$$\mathcal{D}_\varepsilon(\rho) = (1 - \varepsilon) \cdot \rho + \varepsilon \cdot |\mathcal{E}\rangle\langle\mathcal{E}|.$$

We observe that for any 2×2 matrix ρ ,

$$\mathcal{D}_\varepsilon(\rho) = \Pi_\varepsilon \cdot \mathcal{D}(\rho). \quad (8)$$

We now state our hypercontractivity bound for the quantum erasure channel:

Theorem 1.2 (informal version of Theorem 3.2). *Given any $2^n \times 2^n$ matrix X , $1 \geq \varepsilon \geq 0$, $q \geq p \geq 1$ and $c \geq 1$ satisfying $1 - \varepsilon \leq \left(\frac{p-1}{q-1}\right)^c$, let $\|\cdot\|_p$ denote normalized Schatten p -norms. The hypercontractive inequality for the quantum erasure channel*

$$(2^{-n} \text{Tr} [\Pi_\varepsilon^{\otimes n} \cdot |\mathcal{D}^{\otimes n}(X)|^q])^{1/q} \leq \|X\|_p \quad (9)$$

holds if one of the following conditions is satisfied:

- $c = 1$ and $1 \leq p \leq 2 \leq q$.
- $c = 2$ and $1 \leq p \leq q \leq 2$.

Remark 1.3 (Equivalent form, hypercontractivity for random partial trace). It is not hard to verify that Eq. (9) is equivalent to the following.

$$\left(\sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \|\tau_S X\|_q^q \right)^{1/q} \leq \|X\|_p, \quad (10)$$

where $\tau_S X = \text{Tr}_S X / 2^{|S|}$ is the normalized partial trace. Thus, an alternate interpretation of Theorem 1.2 is a hypercontractive inequality for random partial trace operations.

Remark 1.4. It might seem opaque that Eq. (9) is a direct quantum generalization of Eq. (4). The connection becomes clearer when we notice that the expectation on the left-hand side of Eq. (4) is taken over the random variable $Y = \text{BEC}_\varepsilon(X)$. The random variable Y is certainly not uniformly distributed, and we can verify that for each $i \in [n]$,

$$Y_i = \begin{cases} 0 & \text{w.p. } (1 - \varepsilon)/2 \\ 1 & \text{w.p. } (1 - \varepsilon)/2 \\ * & \text{w.p. } \varepsilon \end{cases} \quad (11)$$

We can expand the left-hand side of Eq. (4) according to Eq. (11) and get

$$\|g(Y)\|_q = \left(\mathbb{E}_Y [|g(y)|^q] \right)^{1/q} = \left(\sum_{y \in Y} \mu(y) |g(y)|^q \right)^{1/q} \quad (12)$$

where $\mu(y) = \prod_{i=1}^n \mu(y_i)$ and each $\mu(y_i)$ takes the value according to Eq. (11). The matrix $\frac{1}{2^n} \Pi_\varepsilon$ in Eq. (9) acts as the role of μ in Eq. (12) as it reassigns the noise of the matrix when tracing out.

Remark 1.5. For the classical BEC hypercontractivity bound, the parameter c is always $c = 1$ and this result is tight when $\varepsilon \leq 1/2$. In our work the tightness of the hypercontractivity results for the QEC remains unknown.

Remark 1.6. The reader may wonder why we consider an inequality of the form in Eq. (9), as opposed to the more natural hypercontractivity bound for the QEC expressed as

$$(2^{-n} \text{Tr} [|\mathcal{D}_\varepsilon^{\otimes n}(X)|^q])^{1/q} \leq (2^{-n} \text{Tr}[|X|^p])^{1/p}. \quad (13)$$

It's easy to check that

$$(2^{-n} \text{Tr} [|\mathcal{D}_\varepsilon^{\otimes n}(X)|^q])^{1/q} \leq (2^{-n} \text{Tr} [\Pi_\varepsilon^{\otimes n} \cdot |\mathcal{D}^{\otimes n}(X)|^q])^{1/q}.$$

Hence Eq. (9) is stronger. Moreover, Eq. (13) fails to prove a tight bound on the communication complexity of common randomness generation considered in Section 5. We posit that the inequality in Eq. (9) will have broader applicability in a majority of scenarios, as it is exactly the quantum generalization of its classical counterpart [NW16] as argued in Remark 1.4. Furthermore, Theorem 1.2 also implies an almost optimal hypercontractivity bound for the quantum depolarizing channel.

Remark 1.7. For the quantum erasure channel, our Theorem 1.2 only proves hypercontractivity inequalities for $p \leq 2$. The $p > 2$ case remains open. For quantum channels generated by primitive and reversible Lindblad generators, i.e., the depolarizing channel, a hypercontractivity inequality for p and q in $[1, 2]$ can be generalized to all parameter ranges. See Proposition 10 and Theorem 11 from [BDR20]. However, for the quantum erasure channel considered in our work, Theorem 1.2 does not seem to be able to imply hypercontractivity inequalities for the quantum erasure channel in the range $q > p > 2$. The barrier seems to be fundamental: For fixed p , we expect (and prove) that the left hand side of Eq. (9) is non-increasing over ε . However, for $q \geq 2$, numerical experiments show that this fails to be true. So our proof framework completely fails.

Quantum Log-Sobolev Inequality for Bernoulli Entropy A key component in the proof of our hypercontractivity bound is a quantum log-Sobolev inequality for Bernoulli entropy. It is an extension of the standard quantum log-Sobolev inequality due to Kastoryano and Temme [KT13].

Theorem 1.8 (informal version of Theorem 4.2). *Let m, n be integers such that $m \leq n$ and X is a positive semi-definite operator. Let τ denote the normalized trace operator. For $q \in [1, 2]$, $\varepsilon \in [0, 1]$, it holds that*

$$\begin{aligned} \text{Ent}_{\varepsilon, [m], q}[X] &\leq 2 \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \sum_{k \in [m] \setminus S} \left(\tau[X_{S^c}^q] - \tau[X_{(S \cup \{k\})^c}^q] \right) \\ &\quad + 2 \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \sum_{k \in [n] \setminus S} \left(\tau[X_{S^c}^q] - \tau \left[\left(\tau_k \left[X_{S^c}^{\frac{q}{2}} \right] \right)^2 \right] \right). \end{aligned} \quad (14)$$

where $X_T = \text{Tr}_{T^c} X / 2^{n-|T|}$ for $T \subseteq [n]$, and $\text{Ent}_{\varepsilon, [m], q}$ is the multipartite Bernoulli entropy defined as

$$\begin{aligned} \text{Ent}_{\varepsilon, [m], q}[X] &= \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \tau[(\tau_S X)^q \ln (\tau_S X)^q] \\ &\quad - \left(\sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \tau[(\tau_S X)^q] \right) \ln \left(\sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \tau[(\tau_S X)^q] \right). \end{aligned} \quad (15)$$

We note that when $m = 0$ and $q = 2$, the first term on the right hand side disappears and thus the inequality boils down to the standard quantum log-Sobolev inequality Theorem 4.1.

Common Randomness Generation We present an application of our hypercontractivity bound for the QEC in the study of common randomness generation, generalizing Guruswami and Radhakrishnan's result [GR16] to the quantum setting.

Theorem 1.9 (informal version of Theorem 5.4). *Let $\varepsilon \in [0, 1]$ and $|\Phi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ be an EPR state. Suppose Alice and Bob share infinitely many copies of the state $(\mathbb{1}_2 \otimes \mathcal{D}_\varepsilon)(\Phi)$, and Alice sends classical messages to Bob. Then to produce a common random string of min-entropy at least k , Alice needs to send Bob at least*

$$t \geq \frac{\varepsilon k}{2} - o(k) \quad (16)$$

bits of classical message.

This lower bound is tight up to a constant factor as it suffices for Alice to send εk bits [GR16].

1.2 PROOF OVERVIEW

The classical hypercontractivity bound for BECs is proved in [NW16] by an inductive argument on the length of the input n . The base case is proved via an information-theoretic argument.

Then, for Boolean functions $g : \{-1, 1, *\}^n \rightarrow \mathbb{R}$ with $n > 1$ and the input $y \in \{-1, 1, *\}^n$ of $g(y)$ is split into $y = y_1 y_2$ where $y_1 \in \{-1, 1, *\}^{n_1}$ and $y_2 \in \{-1, 1, *\}^{n_2}$ and $n_1 + n_2 = n$. Then by fixing y_1 , we may apply the induction on the function $g(y_1, \cdot) : \{-1, 1, *\}^{n_2} \rightarrow \mathbb{R}$. To generalize this inductive argument to quantum hypercontractivity, a natural approach is splitting the matrix into block matrices and applying the induction hypothesis on each block. A key step is to compare

$$\|M\|_p = \left\| \begin{pmatrix} X & Y \\ Y^\dagger & Z \end{pmatrix} \right\|_p \quad \text{and} \quad m := \left\| \begin{pmatrix} \|X\|_p & \|Y\|_p \\ \|Y\|_p & \|Z\|_p \end{pmatrix} \right\|_p$$

for positive semidefinite matrices. It is apparent that the inductive argument holds for the classical BEC hypercontractivity as M is diagonal and thus $\|M\|_p = m$ in the classical case. King [Kin03] proved a *norm-compression inequality* which states that $\|M\|_p \geq m$ if $1 \leq p \leq 2$ and $\|M\|_p \leq m$ if $p \geq 2$. Leveraging this fact, King [Kin14] generalized the inductive argument to the quantum setting and proved a hypercontractivity bound for the quantum depolarizing channel in the case that $1 \leq p \leq 2 \leq q$.

The inductive argument also works for the hypercontractivity for the QEC, when $1 \leq p \leq 2 \leq q$, which, therefore, induces a hypercontractivity bound for the QEC in this region². However, the inductive argument fails for the case $1 \leq p \leq q \leq 2$ as the direction of the norm compression inequality is opposite to what we expect. To resolve this difficulty, we extend the framework in [OZ99, Kin14, BDR20, Bei21].

We give a brief overview of this proof framework of quantum hypercontractivity. A key observation is the connection between quantum log-Sobolev inequality [KT13] and quantum hypercontractivity. Such connections have also been explored in classical settings [PS19, FOW22, KKK⁺21]. A fundamental notion in the quantum log-Sobolev inequality is the *entropy* function,

$$\text{Ent}(X^2) = \tau[X^2 \ln X^2] - \tau[X^2] \ln \tau[X^2], \quad (17)$$

which captures the overall chaos of the system. It can be obtained by taking the derivative of the left-hand side of the hypercontractive inequality. The entropy is upper bounded via a quantum log-Sobolev inequality [KT13] as follows:

$$\text{Ent}(X^2) = \tau[X^2 \ln X^2] - \tau[X^2] \ln \tau[X^2] \leq 2 \sum_{k=1}^n (\tau[X^2] - \tau[(\tau_k X)^2]). \quad (18)$$

Intuitively, the entropy is upper bounded by the correlation between each coordinate and the remaining system. The quantum hypercontractivity can be derived from the quantum log-Sobolev inequality via the quantum Gross' lemma [Gro75a, Kin14]. Nonetheless, it only gives a quantum hypercontractive inequality for positive operators as the quantum log-Sobolev inequality only concerns about positive matrices. To extend it to general matrices, we can further employ Watrous' theorem, which asserts that the $\|\cdot\|_{q \rightarrow p}$ norm of a quantum channel can be achieved by positive matrices [Wat05]. The overall proof is summarized in Fig. 2.

²See Appendix D

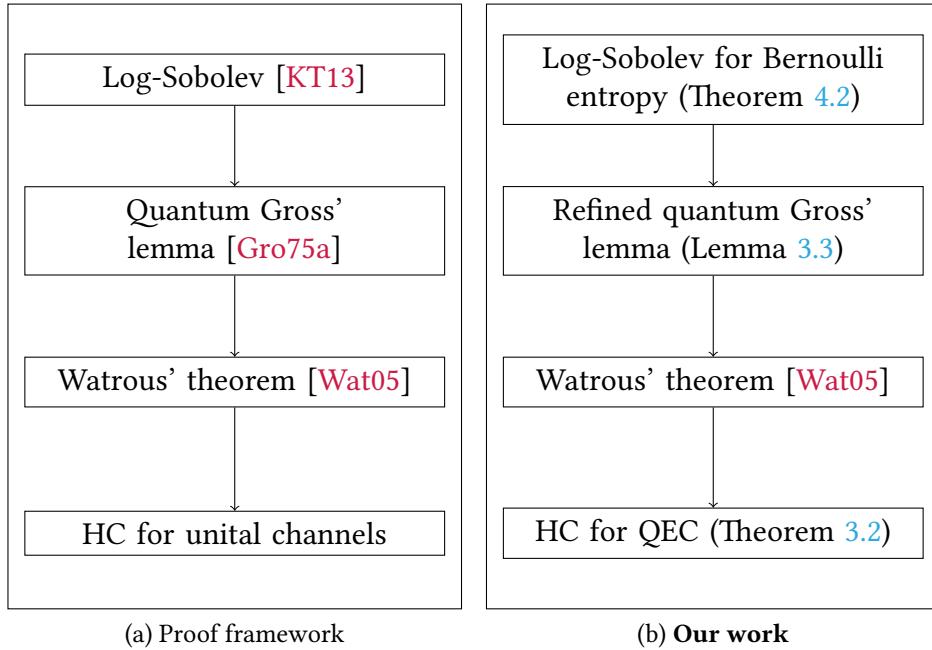


Figure 2: Proof Overview

To prove the hypercontractive inequality for the QEC, we adopt the same framework. However, we need to introduce several new ingredient, which is also depicted in Fig. 2.

The major difficulty is that the quantum log-Sobolev inequality is insufficient for the hypercontractivity of QEC, because the size of the system considered in Eq. (18) is unchanged, while the qubits may be erased and the size of the system decreases when considering the QEC. To capture the “erasing” feature of the QEC, we introduce the notion of *multipartite Bernoulli entropy*. Let X be a $2^n \times 2^n$ positive definite matrix viewed as an operator acting on n qubits, and $m \leq n$. Recall the *multipartite Bernoulli entropy* defined as

$$\text{Ent}_{\varepsilon,[m],q}[X] = \sum_{S \subseteq [m]} (1-\varepsilon)^{m-|S|} \varepsilon^{|S|} \tau[(\tau_S X)^q \ln (\tau_S X)^q] - \left(\sum_{S \subseteq [m]} (1-\varepsilon)^{m-|S|} \varepsilon^{|S|} \tau[(\tau_S X)^q] \right) \ln \left(\sum_{S \subseteq [m]} (1-\varepsilon)^{m-|S|} \varepsilon^{|S|} \tau[(\tau_S X)^q] \right). \quad (19)$$

The multipartite Bernoulli entropy of X captures the *expected* entropy of a random system obtained by removing each qubit system in the first m qubits systems with probability ε , independently. It is a natural extension of the entropy in Eq. (17), which is the case that $m = 0$ and $q = 2$. We establish a quantum log-Sobolev inequality for Bernoulli entropy, which enables us to upper bound the multipartite Bernoulli entropy in terms of the the expected correlation between each coordinate and the remaining system, where the expectation is over the random erasure of the

system.

$$\begin{aligned} \text{Ent}_{\varepsilon, [m], q}[X] &\leq 2 \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \sum_{k \in [m] \setminus S} \left(\tau[X_{S^c}^q] - \tau[X_{(S \cup \{k\})^c}^q] \right) \\ &\quad + 2 \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \sum_{k \in [n] \setminus S} \left(\tau[X_{S^c}^q] - \tau \left[\left(\tau_k \left[X_{S^c}^{\frac{q}{2}} \right] \right)^2 \right] \right). \end{aligned}$$

We also note that when $m = 0$ and $q = 2$ our quantum log-Sobolev inequality degenerates to the quantum log-Sobolev inequality above as Eq. (18). It is worth to note that for $q \neq 2$, the inequality resembles, to some extent, the nonlinear log-Sobolev inequality considered in [PS19, SZ24].

The existing quantum Gross' lemma is also insufficient to derive a hypercontractivity for the QEC from the quantum log-Sobolev inequality. To address it, we prove a refined quantum Gross' lemma connecting the Schatten p norm of an operator and the Schatten p norm of its reduced operator via exploiting the majorization argument between a matrix's eigenvalues and its diagonal entries.

Finally, we conclude our result by employing Watrous' theorem [Wat05], which implies that the above argument for positive semi-definite matrices is sufficient to prove a QEC hypercontractivity for general matrices.

1.3 DISCUSSION AND OPEN PROBLEMS

We prove a hypercontractive inequality for the product of quantum erasure channels. Our hypercontractivity is a natural extension of the hypercontractivity bound for the classical binary erasure channel. To prove the hypercontractivity, we establish a quantum log-Sobolev inequality for Bernoulli entropy, which is interesting in its own right. As an application, we prove an almost tight lower bound on the classical communication complexity of common randomness generation assisted by erased-noisy EPR states. This work raises several open problems for future work.

1. For parameters $2 \leq p \leq q < \infty$, the hypercontractivity of the quantum erasure channel remains unknown. Our technique fails because the derivative $g'(t)$ in the proof of Theorem 3.2 fails to be non-increasing for $q \geq 2$. More technically, Lemma 3.5 holds only for $q \geq 2$. Thus the hypercontractivity for $2 \leq p \leq q < \infty$ remains open.
2. The tightness of hypercontractivity bound for the QEC is not clear. The hypercontractivity bound for classical BECs [NW16] indicates that the parameter c in our result could be pushed to 1. We leave this to future work.
3. To our knowledge, hypercontractive inequalities for tensored channels, i.e., the exact tensorization property, has been proved only for qubit channels [MO10, KT13, Kin14]. For qudit channels, so far there are only lower bounds from bounds on the log-Sobolev constant [TPK14, MHSFW16] (See the discussion in [BDR20]). Is it possible to establish the tensorization property for qudit channels? The current approach fails even for the easy

case $1 \leq p \leq 2 \leq q$, because the norm-compression inequality does not hold for $t \times t$ block matrices whenever $t \geq 3$ [Aud06]. Hence, we need new techniques to establish quantum hypercontractive inequalities for the product of qudit channels.

4. There have been several works investigating the hypercontractivity bound for quantum channels with respect to the fixed states of the channels over this space [BR22, BDR20]. For hypercontractivity bounds with respect to a non-fixed state, Bardet and Rouzé [BR22] have shown some no-go results: a strong decoherence-free-log-Sobolev inequality does not hold for a quantum Markov semigroups that is neither primitive nor unital. Can we say more about a general quantum hypercontractive inequality in weighted non-commutative L_p spaces with respect to a non-fixed state?
5. Log-Sobolev inequalities are powerful tools to study the mixing time of a random walk on a fixed graph. Recently, there have been several works investigating random walks on a random graph [NP08, CHK12, BKW14]. The entropy in the classical log-Sobolev inequality for Bernoulli entropy captures the average entropy of the random subgraphs of a given graph, where each vertex is deleted with probability ε . Thus, it is interesting to see whether our classical log-Sobolev inequality for Bernoulli entropy can provide better analysis on the mixing time of a random walk on a random graph?

1.4 ORGANIZATION

In Section 2, we introduce some preliminary concepts, basic mathematical foundations, notations, and discuss the erasure channel. Section 3 presents the proof of our hypercontractivity inequality. Section 4 presents the proof of quantum log-Sobolev inequality for Bernoulli entropy. Subsequently, in Section 5 we demonstrate an application of our results to derive a lower bound for common randomness generation.

ACKNOWLEDGMENT

We thank the anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions. We thank Mark M. Wilde for pointing out their work [DKQ⁺23, LRWW23, NW24]. We thank Pei Wu for pointing out their work [EWW23]. We are grateful to Xinyuan Zhang and Xiaoyu Chen for their discussion about log-Sobolev inequalities and entropy functions. We thank Jingcheng Liu and Yitong Yin for the helpful discussion. This research was supported by National Natural Science Foundation of China (Grant No. 62332009, 12347104), Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0302901), NSFC/RGC Joint Research Scheme (Grant no. 12461160276) and Natural Science Foundation of Jiangsu Province (No. BK20243060) and the New Cornerstone Science Foundation.

2 PRELIMINARY

In this paper, we will use $\ln x$ to express the natural logarithm of x , which is based on the mathematical constant $e \approx 2.7182818$. We denote $[n] = \{1, 2, \dots, n\}$ and for the special case $[0] = \emptyset$. Besides, unless explicitly stated, we assume $\varepsilon \in [0, 1]$. For complex number z , we use \bar{z} to denote its complex conjugate.

Linear algebra Let $\mathbb{1}_n$ be the $n \times n$ size identity matrix and we use $\mathbb{1}$ to represent $\mathbb{1}_2$. We let $\mathcal{M}_{n \times m}$ be the set of all $n \times m$ size complex-valued matrices. We use \mathcal{M}_n to denote all $n \times n$ square matrices and \mathcal{H}_n to denote all $n \times n$ Hermitian matrices. For $X \in \mathcal{M}_n$, we define its trace as $\text{Tr}[X] = \sum_{i=1}^n X_{ii}$. Additionally, its *normalized* trace is defined as $\tau[X] = \frac{1}{n} \text{Tr}[X]$.

$X \in \mathcal{M}_n$ is a positive semi-definite (PSD) matrix if X is Hermitian and $\mathbf{x}^\dagger X \mathbf{x} \geq 0$ holds for all vectors $\mathbf{x} \in \mathbb{C}^n$, where \mathbf{x}^\dagger is the complex conjugate transpose of \mathbf{x} . We may write $X \geq 0$ for a PSD matrix X . If $X \geq 0$ satisfies $\text{Tr}[X] = 1$, we say X is a density operator. The set of all PSD matrices and density operators is denoted as Pos_n and \mathcal{D}_n .

For a vector $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{C}^n$ and $1 \leq p < \infty$, we use $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ to denote its p -norm. For $X \in \mathcal{M}_{n \times m}$, the Schatten p -norm is defined to be $\|X\|_p = (\text{Tr}[|X|^p])^{1/p}$. We define the *normalized* Schatten p -norm as $\|X\|_p = (\tau[|X|^p])^{1/p}$, where $|X| = \sqrt{X^\dagger X}$ and X^\dagger is the complex conjugate transpose of X .

Let $N = 2^n$ for positive integer n and $X \in \mathcal{M}_N$. All the rows and columns are indexed by the subsets of $[n]$. The partial trace of X with respect to $S \subseteq [n]$ is defined as a square matrix $\text{Tr}_S[X]$ with indices ranging over the subsets of S^c , typically,

$$\text{Tr}_S[X] = \sum_{s \in \{0,1\}^S} (I_{S^c} \otimes \langle s \rangle) X (I_{S^c} \otimes |s\rangle). \quad (20)$$

Besides, the *normalized* partial trace is defined as $\tau_S[X] = 2^{-|S|} \cdot \text{Tr}_S[X]$. For simplicity, we denote $\tau_S[X]$ by X_{S^c} and $\tau_k[X]$ by $X_{\{k\}^c}$.

Schatten p -norms admit Hölder's inequality ([Wat18, Eq. 1.174]): Let $A, B \in \mathcal{M}_{n \times m}$ and p, p^* be positive real numbers satisfying $1/p + 1/p^* = 1$, we have

$$\|A\|_p = \max \left\{ |\text{Tr}[C^\dagger A]| : C \in \mathcal{M}_{n \times m}, \|C\|_{p^*} \leq 1 \right\} \quad (21)$$

which implies

$$|\text{Tr}[B^\dagger A]| \leq \|A\|_p \cdot \|B\|_{p^*}. \quad (22)$$

p^* above is called the *Hölder conjugate* of p . Hölder's inequality can also be generalized to two sequences of matrices.

Lemma 2.1 (Matrix Hölder's inequality). [SA13, Theorem 2.6] *Let $A_i, B_i (i = 1, 2, \dots, m)$ be two sequences of matrices and p, q be positive real numbers satisfying $1/p + 1/q = 1$. We have*

$$\left| \text{Tr} \left[\sum_{i=1}^m A_i^\dagger B_i \right] \right| \leq \left(\text{Tr} \left[\sum_{i=1}^m |A_i|^p \right] \right)^{1/p} \left(\text{Tr} \left[\sum_{i=1}^m |B_i|^q \right] \right)^{1/q}.$$

Specifically, if A_i and B_i are positive semi-definite for all i , then

$$\mathrm{Tr} \left[\sum_{i=1}^m A_i B_i \right] \leq \left(\mathrm{Tr} \left[\sum_{i=1}^m A_i^p \right] \right)^{1/p} \left(\mathrm{Tr} \left[\sum_{i=1}^m B_i^q \right] \right)^{1/q}.$$

Proof. Let $A = \bigoplus_{i=1}^m A_i$ and $B = \bigoplus_{i=1}^m B_i$. By the Hölder's inequality for Schatten p -norm

$$\left| \mathrm{Tr} \left[\sum_{i=1}^m A_i^\dagger B_i \right] \right| = \left| \mathrm{Tr} [A^\dagger B] \right| \leq \|A\|_p \|B\|_q = \left(\mathrm{Tr} \left[\sum_{i=1}^m |A_i|^p \right] \right)^{1/p} \left(\mathrm{Tr} \left[\sum_{i=1}^m |B_i|^q \right] \right)^{1/q}.$$

□

Majorization A doubly stochastic matrix $X \in \mathbb{R}^{n \times n}$ is a non-negative matrix with all row sums and column sums equal to 1. For $\mathbf{x} \in \mathbb{R}^n$, we let \mathbf{x}^\downarrow be the vector by rearranging the values of \mathbf{x} in decreasing order.

Definition 2.2 (Majorization). We say \mathbf{x} majorizes \mathbf{y} (denoted as $\mathbf{x} \succeq \mathbf{y}$) if

$$\sum_{i=1}^n \mathbf{x}_i = \sum_{i=1}^n \mathbf{y}_i \text{ and } \sum_{i=1}^k \mathbf{x}_i^\downarrow \geq \sum_{i=1}^k \mathbf{y}_i^\downarrow$$

for all $k \in [n]$. Two key properties of majorization are listed below:

Theorem 2.3 (Hardy-Littlewood-Pólya Theorem). [MOA79, page 33] Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $\mathbf{x} \preceq \mathbf{y}$ if and only if there exists a doubly stochastic matrix D such that $\mathbf{x} = D\mathbf{y}$.

Theorem 2.4 (Schur-Horn Theorem). [MOA79, page 300] For $X \in \mathcal{H}_n$, let $\mathrm{diag} = (X_{11}, \dots, X_{nn})^T$ be the diagonal entries, $\boldsymbol{\lambda} = (\lambda_1(X), \dots, \lambda_n(X))^T$ where $\lambda_i(X)$ denotes the i -th biggest eigenvalue of X . Then $\mathrm{diag} \preceq \boldsymbol{\lambda}$.

According to Theorem 2.3, the majorization relation in Theorem 2.4 relates to a doubly stochastic matrix D . The following lemma implicates that D has an intriguing property.

Lemma 2.5. Under the setting of Theorem 2.4, let $X = U\Sigma U^\dagger$ be a spectral decomposition of X with diagonal matrix Σ satisfying $\Sigma_{i,i} = \lambda_i(X)$. Then it holds that $\mathrm{diag} = D\boldsymbol{\lambda}$, where $D_{ij} = |U_{ij}|^2$ is a doubly stochastic matrix.

Proof. For $\forall i \in [n]$, $X_{i,i} = \sum_{j=1}^n U_{ij} \Sigma_{jj} \overline{U_{ij}} = \sum_{j=1}^n |U_{ij}|^2 \Sigma_{jj} = \sum_{j=1}^n D_{ij} \lambda_j$. □

Entropy In order to define our entropy notations used in our proof of hypercontractivity inequality, we start with a weighted relative entropy, which unifies the three definitions of relative entropies below. Note that in this work the entropy is usually defined on quantum operators, as a quantum analog of the entropy for functions in the context of log-Sobolev inequalities. This is in contrast to the standard quantum information theory definition of entropy primary used for quantum states.

Definition 2.6 (Weighted Relative Entropy). Given an index set I , with a sequence of non-negative numbers $(w_i)_{i \in I}$ satisfying $\sum_{i \in I} w_i = 1$ and PSD matrices $(X_i)_{i \in I}$, we define the *weighted relative entropy* as

$$\text{Entropy}[(w_i, X_i)_{i \in I}] = \sum_{i \in I} w_i \cdot \tau[X_i \ln X_i] - \left(\tau \left[\sum_{i \in I} w_i \cdot X_i \right] \right) \ln \left(\tau \left[\sum_{i \in I} w_i \cdot X_i \right] \right).$$

Generally, one can easily check $\text{Entropy}[(w_i, X_i)_{i \in I}] \geq 0$ with Jensen's inequality. In this paper, we consider three special forms of the weighted relative entropy.

Definition 2.7 (Two-point Relative Entropy). For any $a, b > 0, \varepsilon \in [0, 1]$,

$$\begin{aligned} \text{Ent}_\varepsilon[a, b] &= \text{Entropy}[(1 - \varepsilon, a), (\varepsilon, b)] \\ &= (1 - \varepsilon)a \ln a + \varepsilon b \ln b - ((1 - \varepsilon)a + \varepsilon b) \ln ((1 - \varepsilon)a + \varepsilon b). \end{aligned}$$

This definition is used fruitfully when exploring isoperimetry properties over general probability distributions. See [BM00, BL98] for more details.

Definition 2.8 (One-Partite Relative Entropy). For $X \geq 0$ and $q \geq 1$, we define the entropy as:

$$\begin{aligned} \text{Ent}_q[X] &= \text{Entropy}[(1, X^q)] \\ &= \tau[X^q \ln X^q] - \tau[X^q] \ln \tau[X^q]. \end{aligned}$$

This definition is related to the quantum relative entropy. Assume $X = 2^n \phi$ where ϕ is a density operator. Then $\text{Ent}_1[X] = D(\phi \parallel \frac{1_{2^n}}{2^n})$, where $D(\phi \parallel \frac{1_{2^n}}{2^n})$ is the quantum relative entropy between ϕ and the maximally mixed state.

This particular quantity can be traced back to the work presented in [Gro75b]. Besides, this term of relative entropy will inevitably arise when we employ the similar derivative techniques used in [Bei21, BDR20] to prove the quantum hypercontractive inequality.

Here we denote $\text{Ent}_1[X^q]$ by $\text{Ent}_q[X]$ to keep it consistent with the definition below.

Definition 2.9 (Bernoulli Multi-Partite Relative Entropy). For $X \in \text{Pos}_{2^n}$, $S \subseteq [n]$ and $q \geq 1$, we define the entropy as:

$$\begin{aligned} \text{Ent}_{\varepsilon, S, q}[X] &= \text{Entropy} \left[((1 - \varepsilon)^{|S|-|T|} \varepsilon^{|T|}, X_{T^c}^q)_{T \subseteq S} \right] \\ &= \sum_{T \subseteq S} (1 - \varepsilon)^{|S|-|T|} \varepsilon^{|T|} \tau[X_{T^c}^q \ln (X_{T^c}^q)] - \\ &\quad \left(\tau \left[\sum_{T \subseteq S} (1 - \varepsilon)^{|S|-|T|} \varepsilon^{|T|} X_{T^c}^q \right] \right) \ln \left(\tau \left[\sum_{T \subseteq S} (1 - \varepsilon)^{|S|-|T|} \varepsilon^{|T|} X_{T^c}^q \right] \right). \end{aligned} \tag{23}$$

This definition is equivalent to Eq. (40). The term *Bernoulli* is derived from the internal utilization of the Bernoulli distribution. We identify it as *multi-partite* due to the inclusion of all partial traces of X within the expression. This quantity reflects, to some degree, the relative entropy of each constituent party of X . It's important to note that this denotes a natural extension of the one-partite relative entropy.

Quantum Erasure Channel A quantum channel Φ is a linear map in $L(\mathcal{M}_n, \mathcal{M}_{n'})$ with the following properties: (i) completely positive : \forall PSD matrix $X \in \mathcal{M}_{nn'}$, we have $(\Phi_n \otimes \mathbb{1}_{n'})(X) \geq 0$; (ii) trace preserving : $\forall X \in \mathcal{M}_n$, $\text{Tr}[\Phi(X)] = \text{Tr}[X]$. A useful equation is [Wat18, Eq. 2.70]: for $\Phi : \mathcal{M}_n \rightarrow \mathcal{M}_m$, $X \in \mathcal{M}_n$ and $Y \in \mathcal{M}_m$, $\text{Tr}(\Phi(X) \cdot Y) = \text{Tr}(X \cdot \Phi^\dagger(Y))$.

To introduce quantum erasure channels, we first define an expanding operator \mathcal{D} in $L(\mathcal{M}_2, \mathcal{M}_3)$:

$$\mathcal{D}(X) = \begin{pmatrix} X & \\ & \tau(X) \end{pmatrix}.$$

The single-qubit quantum erasure channel (QEC) \mathcal{D}_ε is defined as

$$\mathcal{D}_\varepsilon(X) = \Pi_\varepsilon \cdot \mathcal{D}(X), \text{ where } \Pi_\varepsilon = \begin{pmatrix} 1-\varepsilon & & \\ & 1-\varepsilon & \\ & & 2\varepsilon \end{pmatrix}.$$

Here Π_ε can be viewed as an erasure noise matrix. The definition above can be extended to the n -qubit case as follows. Recall that Pauli matrices

$$\left\{ \sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$

form a base of \mathcal{M}_2 and thus $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}^{\otimes n}$ form a base of \mathcal{M}_{2^n} . For any $X = \sum_{x \in \mathbb{Z}_4^n} (\hat{X}_x \otimes_{i=1}^n \sigma_{x_i})$, we define $\mathcal{D}^{\otimes n}$ in $L(\mathcal{M}_{2^n}, \mathcal{M}_{3^n})$ as

$$\mathcal{D}^{\otimes n}(X) = \sum_{x \in \mathbb{Z}_4^n} \hat{X}_x \left(\bigotimes_{i=1}^n \mathcal{D}(\sigma_{x_i}) \right)$$

and the n -qubit quantum erasure channel is defined as $\mathcal{D}_\varepsilon^{\otimes n}(X) = \Pi_\varepsilon^{\otimes n} \cdot \mathcal{D}^{\otimes n}(X)$. When the context is clear, we omit the symbols $\otimes n$ and use \mathcal{D} , Π_ε and \mathcal{D}_ε to denote the n -qubit case. An n -qubit QEC induces a norm, which we refer to as the (ε, q) -norm in \mathcal{M}_{2^n} .

Definition 2.10. For any integer $n > 0$, $\varepsilon \in [0, 1]$ and $q \geq 1$, the (ε, q) -norm of $X \in \mathcal{M}_{2^n}$ is

$$\|X\|_{\varepsilon, q} = (2^{-n} \cdot \text{Tr}[\Pi_\varepsilon \cdot |\mathcal{D}(X)|^q])^{\frac{1}{q}}.$$

To see that it is a norm, it suffices to verify that $\|X\|_{\varepsilon, q} = \left\| (\Pi_\varepsilon^{1/2q})^{\otimes n} |\mathcal{D}^{\otimes n}(X)| (\Pi_\varepsilon^{1/2q})^{\otimes n} \right\|_q$.

For a quantum channel with Kraus representation $\Phi(\rho) = \sum_i X_i \rho X_i^\dagger$, we define the conjugate map

$$\Phi^\dagger(\rho) = \sum_i X_i^\dagger \rho X_i.$$

In particular, the conjugate map of the quantum erasure channel is $\mathcal{D}_\varepsilon^\dagger : \mathcal{M}_3 \rightarrow \mathcal{M}_2$:

$$\mathcal{D}_\varepsilon^\dagger(\rho) = (1-\varepsilon) \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \rho \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} + \varepsilon \langle \mathcal{E} | \rho | \mathcal{E} \rangle \cdot \mathbb{1}_2. \quad (24)$$

Note that the set of all possible outputs of a quantum erasure channel forms a subspace of \mathcal{M}_{3^n} , where all elements in the subspaces commute with $\Pi_\varepsilon^{\otimes n}$. Throughout this paper, we will only encounter matrices in this subspace of \mathcal{M}_{3^n} . Such matrices have the form of

$$Y = \sum_{S \subseteq [n]} Y_{S^c} \otimes |\mathcal{E}\rangle\langle \mathcal{E}|^{\otimes S} \quad (25)$$

where each Y_{S^c} lies in the subspace $\mathcal{M}_{2^{n-|S|}}$. The following two facts can be easily verified by the definition.

Fact 2.11. For any operator $X \in \mathcal{M}_{2^n}$, $\mathcal{D}^{\otimes n}(X)$ is of Eq. (25) form.

Fact 2.12. Suppose $Y \in \mathcal{M}_{3^n}$ is of Eq. (25) form. Then Y commutes with $\Pi_\varepsilon^{\otimes n}$ for all $0 \leq \varepsilon \leq 1$.

3 HYPERCONTRACTIVITY

In this section we prove the hypercontractivity for the product of quantum erasure channels. Recall that for $X \in \mathcal{M}_{2^n}$,

$$\|X\|_{\varepsilon,q} = (2^{-n} \cdot \text{Tr}[\Pi_\varepsilon^{\otimes n} \cdot |\mathcal{D}^{\otimes n}(X)|^q])^{\frac{1}{q}} = \left(\sum_{S \subseteq [n]} (1-\varepsilon)^{n-|S|} \varepsilon^{|S|} \|\tau_S X\|_q^q \right)^{\frac{1}{q}}.$$

The following lemma states that it suffices to restrict our argument to positive semi-definite matrices, as the supremum of $\|X\|_{\varepsilon,q}/\|X\|_p$ is achieved by a positive semi-definite matrix. This is implicitly implied by Watrous' paper [Wat05] by noting that $\|X\|_{\varepsilon,q} = \|(\Pi_\varepsilon^{\otimes n})^{1/q} \cdot \mathcal{D}^{\otimes n}(X)\|_q$ and that the map $X \mapsto (\Pi_\varepsilon^{\otimes n})^{1/q} \cdot \mathcal{D}^{\otimes n}(X)$ is completely positive. For the sake of completeness, the proof of the following lemma is provided in Appendix A.

Lemma 3.1 (Watrous' theorem [Wat05]). *For $1 \leq p \leq q, \varepsilon \in [0, 1]$, we have*

$$\sup_{X \in \mathcal{M}_{2^n}, X \neq 0} \frac{\|X\|_{\varepsilon,q}}{\|X\|_p} = \sup_{X \in \text{Pos}_{2^n}, X \neq 0} \frac{\|X\|_{\varepsilon,q}}{\|X\|_p}.$$

The main theorem of this section is the following.

Theorem 3.2 (Hypercontractivity for the product of QECs). *For any $X \in \mathcal{M}_{2^n}$, $q \geq p \geq 1, \varepsilon \in [0, 1], c \geq 1$ satisfying $1 - \varepsilon \leq \left(\frac{p-1}{q-1}\right)^c$, it holds that*

$$\|X\|_{\varepsilon,q} \leq \|X\|_p, \quad (26)$$

whenever either of the following two cases holds.

- **Case 1:** $c = 1$ and $1 \leq p \leq 2 \leq q$.

- **Case 2:** $c = 2$ and $1 \leq p \leq q \leq 2$.

Proof. **Case 1** has a simple inductive proof by using the norm-compression inequality [Kin03]. For the sake of completeness, we provide a proof in Appendix D. **Case 2** is much more challenging. The rest of this section is devoted to this case.

By Lemma 3.1, we can assume X to be positive semi-definite. We note that if $q = 1$, our conclusion holds trivially. Assume $q > 1$ for the following arguments. Fix X and $p \in [1, 2]$. Recall that

$$\|X\|_{\varepsilon, q} = (2^{-n} \cdot \text{Tr}[\Pi_\varepsilon^{\otimes n} \cdot |\mathcal{D}^{\otimes n}(X)|^q])^{\frac{1}{q}} \text{ and } \|X\|_p = (2^{-n} \cdot \text{Tr}[|X|^p])^{\frac{1}{p}}.$$

Introduce a parameter t and let $\frac{p-1}{q-1} = e^{-t/c}$ and $\varepsilon = 1 - e^{-t}$. It is easy to see that Eq. (26) holds with equality at $t = 0$, since in this case $q = p$ and $\varepsilon = 0$.

Let $g(t) = \ln \|X\|_{\varepsilon, q}$, calculating the derivative of $g(t)$ over t yields

$$g'(t) = \frac{q-1}{c \cdot q^2} \cdot \frac{1}{\|X\|_{\varepsilon, q}^q} \cdot \left[\text{Ent}_{\varepsilon, [n], q}[X] - \frac{c \cdot q}{q-1} \sum_{S \subseteq [n]} (1-\varepsilon)^{n-|S|} \varepsilon^{|S|} \sum_{k \in [n] \setminus S} \left(\tau[X_{S^c}^q] - \tau[X_{(S \cup \{k\})^c}^q] \right) \right]$$

and recall that $\text{Ent}_{\varepsilon, [n], q}[X]$ is defined as

$$\sum_{S \subseteq [n]} (1-\varepsilon)^{n-|S|} \varepsilon^{|S|} \tau[X_{S^c}^q \ln X_{S^c}^q] - \left(\sum_{S \subseteq [n]} (1-\varepsilon)^{n-|S|} \varepsilon^{|S|} \tau[X_{S^c}^q] \right) \ln \left(\sum_{S \subseteq [n]} (1-\varepsilon)^{n-|S|} \varepsilon^{|S|} \tau[X_{S^c}^q] \right).$$

To show $g'(t) \leq 0$, we need to prove

$$\text{Ent}_{\varepsilon, [n], q}[X] \leq \frac{c \cdot q}{q-1} \sum_{S \subseteq [n]} (1-\varepsilon)^{n-|S|} \varepsilon^{|S|} \sum_{k \in [n] \setminus S} \left(\tau[X_{S^c}^q] - \tau[X_{(S \cup \{k\})^c}^q] \right). \quad (27)$$

Combining Theorem 4.2 with $m = n$ and Lemma 3.3 for $c' = 4/q^2$, we have

$$\text{Ent}_{\varepsilon, [n], q}[X] \leq \left(2 + \frac{2}{q-1} \right) \sum_{S \subseteq [n]} (1-\varepsilon)^{n-|S|} \varepsilon^{|S|} \sum_{k \in [n] \setminus S} \left(\tau[X_{S^c}^q] - \tau[X_{(S \cup \{k\})^c}^q] \right).$$

Note that

$$2 + \frac{2}{q-1} = \frac{2 \cdot q}{q-1} = \frac{c \cdot q}{q-1}.$$

This completes the proof. □

The following lemma, to some extent, resembles quantum Gross' lemma [Gro75a, Kin14], and the quantum Stroock-Varopoulos inequality [CKMT15, BDR20, Bar17]. Notice that, these inequalities are essential in the proof of the hypercontractivity of unital qubit channels [Gro75a, Kin14, Bei21].

Lemma 3.3 (Refined Gross' lemma). *For $X \in \text{Pos}_{2^n}$ and $q \in (1, 2]$,*

$$\tau[X^q] - \tau\left[\left(\tau_n\left[X^{\frac{q}{2}}\right]\right)^2\right] \leq \frac{1}{q-1}(\tau[X^q] - \tau[(\tau_n X)^q]). \quad (28)$$

Remark 3.4. This lemma differs to the quantum Stroock-Varopoulos inequality from [BDR20] in that the right hand side is not a Dirichlet form, thus is new as far as we can tell.

The proof of Lemma 3.3 is built on the following Lemma.

Lemma 3.5. *Let $q \in (1, 2]$, $\lambda \in \mathbb{R}_{\geq 0}^{2^n}$ and D be a $2^n \times 2^n$ doubly stochastic matrix. It holds that*

$$\|\lambda\|_q^q - \left\|D\lambda^{\frac{q}{2}}\right\|_2^2 \leq \frac{1}{q-1}\left(\|\lambda\|_q^q - \|D\lambda\|_q^q\right). \quad (29)$$

Proof of Lemma 3.3. We will reduce (28) to (29).

Without loss of generality, we assume $\tau_n X$ is diagonal from now on. Otherwise we could replace X by $(V \otimes \mathbb{1})^\dagger X (V \otimes \mathbb{1})$ where V is a unitary to diagonalize $\tau_n X$ without changing any term in the equation above.

Let $\text{diag}(X)$ be the column vector of X 's diagonal entries and $\lambda(X)$ be the column vector of X 's eigenvalues. Both vectors are sorted in non-increasing order. Then there exists a unitary U s.t. $X = U\Sigma U^\dagger$, where $\text{diag}(\Sigma) = \lambda(X)$. According to Theorem 2.4 and Lemma 2.5, we have $\lambda(X) \succeq \text{diag}(X)$ and $\text{diag}(X) = D\lambda(X)$ for some doubly stochastic matrix D , where D depends only on U and is independent of Σ . This property is crucial since we can deduce $\text{diag}(X^{\frac{q}{2}}) = D\lambda(X^{\frac{q}{2}}) = D\lambda(X)^{\frac{q}{2}}$ with the same D , where we use v^q to represent the vector $(v_1^q, \dots, v_n^q)^T$ for column vector $v = (v_1, \dots, v_n)^T$.

Besides, since $\tau_n X$ is diagonal, we know $\lambda(\tau_n X \otimes \mathbb{1}) = \text{diag}(\tau_n X \otimes \mathbb{1}) = M\text{diag}(X) = MD\lambda(X)$ and $\text{diag}(\tau_n\left[X^{\frac{q}{2}}\right] \otimes \mathbb{1}) = M\text{diag}(X^{\frac{q}{2}}) = MD\lambda(X)^{\frac{q}{2}}$, where $M = \frac{1}{2} \cdot \mathbb{1}_{2^{n-1}} \otimes \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. As an intuition, $M\text{diag}(X)$ calculates average values of each two corresponding entries of $\text{diag}(X)$. We use D' to denote $M \cdot D$. Note that D' is also doubly stochastic.

With all the definitions above, we see that

$$2^n \tau[X^q] = \|\lambda(X)\|_q^q \quad (30)$$

$$2^n \tau[(\tau_n X)^q] = 2^n \tau[(\tau_n X)^q \otimes \mathbb{1}_2] = \|D'\lambda(X)\|_q^q \quad (31)$$

$$2^n \tau\left[\left(\tau_n\left[X^{\frac{q}{2}}\right]\right)^2\right] = \left\|\tau_n\left[X^{\frac{q}{2}}\right] \otimes \mathbb{1}\right\|_2^2 \geq \left\|\text{diag}\left(\tau_n\left[X^{\frac{q}{2}}\right] \otimes \mathbb{1}\right)\right\|_2^2 = \left\|D'\lambda(X)^{\frac{q}{2}}\right\|_2^2 \quad (32)$$

Therefore,

$$\begin{aligned} 2^n \left(\tau[X^q] - \tau\left[\left(\tau_n\left[X^{\frac{q}{2}}\right]\right)^2\right] \right) &\leq \|\lambda(X)\|_q^q - \left\|D'\lambda(X)^{\frac{q}{2}}\right\|_2^2 \\ 2^n (\tau[X^q] - \tau[(\tau_n X)^q]) &= \|\lambda(X)\|_q^q - \|D'\lambda(X)\|_q^q. \end{aligned}$$

According to Lemma 3.5, we have

$$\|\boldsymbol{\lambda}(X)\|_q^q - \left\| D' \boldsymbol{\lambda}(X)^{\frac{q}{2}} \right\|_2^2 \leq \frac{1}{q-1} \left(\|\boldsymbol{\lambda}(X)\|_q^q - \|D' \boldsymbol{\lambda}(X)\|_q^q \right)$$

since D' is a doubly stochastic matrix, thus we conclude that

$$\tau[X^q] - \tau \left[\left(\tau_n \left[X^{\frac{q}{2}} \right] \right)^2 \right] \leq \frac{1}{q-1} (\tau[X^q] - \tau[(\tau_n X)^q]).$$

□

Proof of Lemma 3.5. First notice that

$$\sum_{ij} D_{ij} \lambda_j^q = \sum_j \left(\sum_i D_{ij} \right) \lambda_j^q = \sum_j \lambda_j^q = \|\boldsymbol{\lambda}\|_q^q.$$

Thus Eq. (29) is equivalent to

$$\sum_i \sum_j D_{ij} \lambda_j^q + \sum_i \left(\sum_j D_{ij} \lambda_j^{\frac{q}{2}} \right)^2 \leq \frac{1}{q-1} \left(\sum_i \sum_j D_{ij} \lambda_j^q - \sum_i \left(\sum_j D_{ij} \lambda_j \right)^q \right).$$

It suffices to prove for every i ,

$$\sum_j D_{ij} \lambda_j^q + \left(\sum_j D_{ij} \lambda_j^{\frac{q}{2}} \right)^2 \leq \frac{1}{q-1} \left(\sum_j D_{ij} \lambda_j^q - \left(\sum_j D_{ij} \lambda_j \right)^q \right)$$

holds, which again suffices to prove

$$\sum_{i=1}^n \alpha_i \lambda_i^q - \left(\sum_{i=1}^n \alpha_i \lambda_i^{\frac{q}{2}} \right)^2 \leq \frac{1}{q-1} \left(\sum_{i=1}^n \alpha_i \lambda_i^q - \left(\sum_{i=1}^n \alpha_i \lambda_i \right)^q \right)$$

for an arbitrary non-negative vector $\{\alpha_i\}_{i=1}^n$ satisfying $\sum_{i=1}^n \alpha_i = 1$ and an arbitrary non-negative vector $\{\lambda_i\}_{i=1}^n$ and $q \in (1, 2]$.

Treating $\{\alpha_i\}_{i=1}^n$ as a distribution over $[n]$ and considering the normalized p -norms

$$\|\lambda\|_p = \left(\sum_i \alpha_i |\lambda_i|^p \right)^{1/p} = \left(\mathbb{E}_i [|\lambda_i|^p] \right)^{1/p}. \quad (33)$$

Our goal can be tidied into

$$\|\lambda\|_1^q \leq (2-q) \|\lambda\|_q^q + (q-1) \|\lambda\|_{q/2}^q. \quad (34)$$

We first use the Hölder's inequality for probability spaces. Let $p = q/(2-q) \geq 1$ and $p^* = q/(2(q-1))$ be p 's Hölder conjugate, we obtain

$$\begin{aligned}
\|\lambda\|_1 &= \mathbb{E}[\lambda_i] = \mathbb{E}_i[\lambda_i^{2-q} \cdot \lambda_i^{q-1}] \\
&\leq \|\lambda^{2-q}\|_p \cdot \|\lambda^{q-1}\|_{p^*} = \left(\mathbb{E}_i[\lambda_i^q]\right)^{(2-q)/q} \cdot \left(\mathbb{E}_i[\lambda_i^{q/2}]\right)^{2(q-1)/q} = \|\lambda\|_q^{2-q} \cdot \|\lambda\|_{q/2}^{q-1}. \quad (35)
\end{aligned}$$

We now use the following inequality: Let $0 \leq s \leq 1$ and $a, b \geq 0$, then

$$a^s b^{1-s} \leq sa + (1-s)b. \quad (36)$$

This follows simply by the concavity of the log function:

$$\log a^s b^{1-s} = s \log a + (1-s) \log b \leq \log(sa + (1-s)b).$$

Then we apply Eq. (36) and get

$$\|\lambda\|_q^{2-q} \cdot \|\lambda\|_{q/2}^{q-1} \leq (2-q)\|\lambda\|_q + (q-1)\|\lambda\|_{q/2}. \quad (37)$$

Finally, by convexity of the function $x \mapsto x^q$,

$$\|\lambda\|_1^q \leq \left((2-q)\|\lambda\|_q + (q-1)\|\lambda\|_{q/2}\right)^q \leq (2-q)\|\lambda\|_q^q + (q-1)\|\lambda\|_{q/2}^q. \quad (38)$$

□

4 VARIABLE MULTIPARTITE LOG-SOBOLEV INEQUALITY

In this section, we will prove a quantum log-Sobolev inequality for multipartite Bernoulli entropy, which plays a crucial role in the proof of the hypercontractivity for quantum erasure channels. The standard quantum log-Sobolev inequality is due to Kastoryano and Temme [KT13].

Theorem 4.1 (Quantum log-Sobolev inequality). *[KT13] For $A \in \text{Pos}_{2^n}$, it holds that*

$$\text{Ent}_2[A] = \tau[A^2 \ln A^2] - \tau[A^2] \ln \tau[A^2] \leq 2 \sum_{k=1}^n (\tau[A^2] - \tau[(\tau_k[A])^2]).$$

In Theorem 4.1, the dimension of the system is fixed. We now state our main result of the variable multipartite quantum log-Sobolev inequality, where the dimension of the system varies subject to a random partial trace. Recall the definition

$$\begin{aligned}
\text{Ent}_{\epsilon,[m],q}[X] &= \sum_{S \subseteq [m]} (1-\epsilon)^{m-|S|} \epsilon^{|S|} \tau[(\tau_S X)^q \ln (\tau_S X)^q] \\
&\quad - \left(\sum_{S \subseteq [m]} (1-\epsilon)^{m-|S|} \epsilon^{|S|} \tau[(\tau_S X)^q] \right) \ln \left(\sum_{S \subseteq [m]} (1-\epsilon)^{m-|S|} \epsilon^{|S|} \tau[(\tau_S X)^q] \right).
\end{aligned}$$

Theorem 4.2 (Variable Quantum log-Sobolev inequality). *Let m, n be integers such that $m \leq n$ and $X \in \text{Pos}_{2^n}$. For $q \in [1, 2]$, $\varepsilon \in [0, 1]$, it holds that*

$$\begin{aligned} \text{Ent}_{\varepsilon, [m], q}[X] &\leq 2 \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \sum_{k \in [m] \setminus S} \left(\tau[X_{S^c}^q] - \tau[X_{(S \cup \{k\})^c}^q] \right) \\ &\quad + 2 \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \sum_{k \in [n] \setminus S} \left(\tau[X_{S^c}^q] - \tau \left[\left(\tau_k \left[X_{S^c}^{\frac{q}{2}} \right] \right)^2 \right] \right). \end{aligned} \quad (39)$$

where $X_T = \text{Tr}_{T^c} X / 2^{n-|T|}$ for $T \subseteq [n]$, and $\text{Ent}_{\varepsilon, [m], q}$ is the multipartite Bernoulli entropy defined as

$$\begin{aligned} \text{Ent}_{\varepsilon, [m], q}[X] &= \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \tau[(\tau_S X)^q \ln (\tau_S X)^q] \\ &\quad - \left(\sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \tau[(\tau_S X)^q] \right) \ln \left(\sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \tau[(\tau_S X)^q] \right). \end{aligned} \quad (40)$$

To the best of our knowledge, even the classical counterpart of Theorem 4.2 is also unknown. We write it down for readers who are interested in classical log-Sobolev inequalities and their applications.

Corollary 4.3 (Variable Classical log-Sobolev inequality). *For $f : \{0, 1\}^n \rightarrow \mathbb{R}_{\geq 0}$, integers $m \leq n$, $T \subseteq [n]$, we define $f_T : \{0, 1\}^{|T|} \rightarrow \mathbb{R}_{\geq 0}$ to be $f_T(x) = \mathbb{E}_{y \sim \{0, 1\}^{|T^c|}} [f(x, y)]$, where x and y are placed on T and T^c , respectively. For $\varepsilon \in [0, 1]$, $q \in [1, 2]$, define*

$$\begin{aligned} \text{Ent}_{\varepsilon, [m], q}[f] &= \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \mathbb{E}[f_{S^c}^q \ln (f_{S^c}^q)] \\ &\quad - \left(\sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \mathbb{E}[f_{S^c}^q] \ln \left(\sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \mathbb{E}[f_{S^c}^q] \right) \right). \end{aligned}$$

It holds that

$$\begin{aligned} \text{Ent}_{\varepsilon, [m], q}[f] &\leq 2 \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \sum_{k \in [m] \setminus S} \left(\mathbb{E}[f_{S^c}^q] - \mathbb{E}[f_{(S \cup \{k\})^c}^q] \right) \\ &\quad + 2 \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \sum_{k \in [n] \setminus S} \left(\mathbb{E}[f_{S^c}^q] - \mathbb{E} \left[\left(f_{S^c}^{q/2} \right)_{(S \cup \{k\})^c}^2 \right] \right). \end{aligned}$$

We are now ready to prove Theorem 4.2.

Proof of Theorem 4.2. Define

$$J(X, m) = \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \sum_{k \in [m] \setminus S} \left(\tau[X_{S^c}^q] - \tau[X_{(S \cup \{k\})^c}^q] \right),$$

$$K(X, m) = \sum_{S \subseteq [m]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \sum_{k \in [n] \setminus S} \left(\tau[X_{S^c}^q] - \tau \left[\left(\tau_k \left[X_{S^c}^{\frac{q}{2}} \right] \right)^2 \right] \right).$$

It suffices to prove

$$\text{Ent}_{\varepsilon, [m], q}[X] \leq 2 \cdot J(X, m) + 2 \cdot K(X, m). \quad (41)$$

We will prove by induction on m . Recall $[0] = \emptyset$ which implies $J(X, 0) = 0$. Thus our inequality in the base case of $m = 0$ is equivalent to the standard quantum log-Sobolev inequality Theorem 4.1 where $A = X^{\frac{q}{2}}$.

Now suppose $m \geq 1$ and Eq. (41) holds for all integers in $[m-1]$. Recall the definition of two-point relative entropy $\text{Ent}_\varepsilon[\cdot, \cdot]$ in Definition 2.7. We expand the left-hand side

$$\begin{aligned} \text{Ent}_{\varepsilon, [m], q}[X] &= (1 - \varepsilon) \cdot \text{Ent}_{\varepsilon, [m-1], q}[X] + \varepsilon \cdot \text{Ent}_{\varepsilon, [m-1], q}[X_{\{m\}^c}] \\ &\quad + \text{Ent}_\varepsilon \left[\sum_{S \subseteq [m-1]} (1 - \varepsilon)^{m-1-|S|} \varepsilon^{|S|} \tau[X_{S^c}^q], \sum_{S \subseteq [m-1]} (1 - \varepsilon)^{m-1-|S|} \varepsilon^{|S|} \tau[X_{(S \cup \{m\})^c}^q] \right]. \end{aligned}$$

By the induction hypothesis, we have

$$\begin{aligned} &(1 - \varepsilon) \cdot \text{Ent}_{\varepsilon, [m-1], q}[X] + \varepsilon \cdot \text{Ent}_{\varepsilon, [m-1], q}[X_{\{m\}^c}] \\ &= 2(1 - \varepsilon) \cdot J(X, m-1) + 2(1 - \varepsilon) \cdot K(X, m-1) + 2\varepsilon \cdot J(X_{\{m\}^c}, m-1) + 2\varepsilon \cdot K(X_{\{m\}^c}, m-1) \\ &= 2J(X, m) + 2K(X, m) - 2 \sum_{S \subseteq [m-1]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \left(\tau[X_{S^c}^q] - \tau[X_{(S \cup \{m\})^c}^q] \right) \end{aligned}$$

since

$$(1 - \varepsilon) \cdot K(X, m-1) + \varepsilon \cdot K(X_{\{m\}^c}, m-1) = K(X, m),$$

$$(1 - \varepsilon) \cdot J(X, m-1) + \varepsilon \cdot J(X_{\{m\}^c}, m-1) = J(X, m) - \sum_{S \subseteq [m-1]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \left(\tau[X_{S^c}^q] - \tau[X_{(S \cup \{m\})^c}^q] \right).$$

Finally with Claim 4.4, we obtain

$$\text{Ent}_{\varepsilon, [m], q}[X] \leq 2 \cdot J(X, m) + 2 \cdot K(X, m),$$

which completes the proof of the induction step. □

Claim 4.4.

$$\begin{aligned} &\text{Ent}_\varepsilon \left[\sum_{S \subseteq [m-1]} (1 - \varepsilon)^{m-1-|S|} \varepsilon^{|S|} \tau[X_{S^c}^q], \sum_{S \subseteq [m-1]} (1 - \varepsilon)^{m-1-|S|} \varepsilon^{|S|} \tau[X_{(S \cup \{m\})^c}^q] \right] \\ &\leq 2 \sum_{S \subseteq [m-1]} (1 - \varepsilon)^{m-|S|} \varepsilon^{|S|} \left(\tau[X_{S^c}^q] - \tau[X_{(S \cup \{m\})^c}^q] \right). \end{aligned} \quad (42)$$

Proof. By Lemma B.2 and the fact that $1 \leq q \leq 2$, we have for each S ,

$$\tau\left[X_{(S \cup \{m\})^c}^q\right] \leq \tau[X_{S^c}^q] \leq 16\tau\left[X_{(S \cup \{m\})^c}^q\right].$$

Thus we are able to apply Lemma B.1, to get

$$\begin{aligned} & \text{Ent}_\varepsilon\left[\sum_{S \subseteq [m-1]} (1-\varepsilon)^{m-1-|S|} \varepsilon^{|S|} \tau[X_{S^c}^q], \sum_{S \subseteq [m-1]} (1-\varepsilon)^{m-1-|S|} \varepsilon^{|S|} \tau\left[X_{(S \cup \{m\})^c}^q\right]\right] \\ & \leq 2(1-\varepsilon) \sum_{S \subseteq [m-1]} (1-\varepsilon)^{m-1-|S|} \varepsilon^{|S|} \left(\tau[X_{S^c}^q] - \tau\left[X_{(S \cup \{m\})^c}^q\right]\right) \\ & = 2 \sum_{S \subseteq [m-1]} (1-\varepsilon)^{m-1-|S|} \varepsilon^{|S|} \left(\tau[X_{S^c}^q] - \tau\left[X_{(S \cup \{m\})^c}^q\right]\right). \end{aligned}$$

□

5 COMMON RANDOMNESS GENERATION

In this section we give a lower bound on the communication cost of *common randomness* (CR) generation, in the case where Alice and Bob share maximally entangled states affected by the quantum erasure channel as free resource.

CR generation refers to the task of producing a common random string X accessible for both Alice and Bob. In our work we restrict the actions of Alice and Bob to local quantum operations and one-way classical communication from Alice to Bob. We measure the amount of randomness of a random variable X by its min-entropy.

Definition 5.1 (Min-Entropy). Let R be a random variable with distribution μ . We define the min-entropy of R to be

$$H_{\min}(R) = \min_{r \in R} \log(1/\mu(r)). \quad (43)$$

We assume $\log(1/0) = +\infty$ in this definition.

Let $|\Phi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ be an EPR state and write $\Phi = |\Phi\rangle\langle\Phi|$. If Alice and Bob share maximally entangled states, they can apply the same measurements $\{M_0 = |0\rangle\langle 0|, M_1 = |1\rangle\langle 1|\}$ on the shared entanglement and get exactly identical results. Thus they can generate CR without any classical communication. However, EPRs may suffer various noise in reality. Thus, it is natural to investigate CR generation with non-perfect quantum correlation.

Recall that the quantum erasure channel $\mathcal{D}_\varepsilon : \mathcal{M}_2 \rightarrow \mathcal{M}_3$ for $0 \leq \varepsilon \leq 1$ is

$$\mathcal{D}_\varepsilon(X) = \Pi_\varepsilon \mathcal{D}(X) = (1-\varepsilon) \cdot X + \varepsilon \text{Tr}(X) \cdot |\mathcal{E}\rangle\langle\mathcal{E}|.$$

We study the communication cost of CR generation, when Alice and Bob share EPR states while Bob's quantum states are affected by a quantum erasure channel. Note that the quantum erasure channel is not symmetric regarding to maximally entangled states:

$$(\mathcal{D}_\varepsilon \otimes \mathbb{1})(\Phi) \neq (\mathbb{1} \otimes \mathcal{D}_\varepsilon)(\Phi), \quad (44)$$

so we make explicit here that the noise is applied to Bob's part, i.e., they share the state $(\mathbb{1} \otimes \mathcal{D}_\varepsilon)(\Phi)$ where Alice has access to the left side and Bob has access to the right side. This is more interesting because it is Alice who sends one-way information to Bob. Otherwise if the noise is applied to Alice's part, Alice can simply tell Bob which qubits are faulty.

Suppose Alice and Bob share n copies of the state $(\mathbb{1} \otimes \mathcal{D}_\varepsilon)(\Phi)$, and their goal is to produce a CR of min-entropy at least k , with one-way classical communication from Alice to Bob. Alice's strategy can be represented by a POVM $\{X_{a,\pi}\}_{a,\pi}$, where $X_{a,\pi} \in \text{Pos}_{2^n}$ and $\sum_{a,\pi} X_{a,\pi} = \mathbb{1}_{2^n}$. After receiving the message π from Alice, Bob performs his measurement described by $\{Y_a^\pi\}_a$, where $Y_a^\pi \in \text{Pos}_{3^n}$ and $\sum_a Y_a^\pi = \mathbb{1}_{3^n}$ for each π . Then their success probability is

$$\Pr[\text{Success}] = \sum_{a,\pi} \text{Tr} [(X_{a,\pi} \otimes Y_a^\pi) \cdot (\mathbb{1} \otimes \mathcal{D}_\varepsilon^{\otimes n})(\Phi^{\otimes n})].$$

Our first lemma will be that Bob's strategy, i.e. the measurement operators Y_a^π , will be of the form of Eq. (25). That is,

$$Y_a^\pi = \sum_{S \subseteq [n]} Y_{a,S^c}^\pi \otimes |\mathcal{E}\rangle\langle\mathcal{E}|^{\otimes S}$$

where for all $S \subseteq [n]$, Y_{a,S^c}^π lies in the subspace $\mathcal{M}_{2^{[n]-S}}$. As for an intuitive perspective, this means Bob is the one who knows which qubits are polluted and which are not, and can make his decision only based on the clean qubits. Define

$$\bar{\Pi}_0 = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix} \text{ and } \bar{\Pi}_1 = \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix},$$

and $\bar{\Pi}_b = \bar{\Pi}_{b_1} \otimes \dots \otimes \bar{\Pi}_{b_1}$ for $b \in \{0,1\}^n$.

Lemma 5.2. *Let $X \in \text{Pos}_{3^n}$ be of Eq. (25) form. Then*

$$\max_{Y \in \mathcal{M}_{3^n}} |\text{Tr}[XY]|$$

is achieved by a Y of Eq. (25) form.

Proof. Notice that $X = \sum_{b \in \{0,1\}^n} \bar{\Pi}_b X \bar{\Pi}_b$. Suppose we have an optimal $Y \in \text{Pos}_{3^n}$. Let

$$\tilde{Y} = \sum_{b \in \{0,1\}^n} \bar{\Pi}_b Y \bar{\Pi}_b.$$

Then

$$\begin{aligned}
\mathrm{Tr}[X\tilde{Y}] &= \sum_b \mathrm{Tr}[X\bar{\Pi}_b Y \bar{\Pi}_b] \\
&= \mathrm{Tr}\left[\left(\sum_b \bar{\Pi}_b X \bar{\Pi}_b\right) Y\right] \\
&= \mathrm{Tr}[XY].
\end{aligned}$$

□

Corollary 5.3. *Bob's strategy can be replaced by measurement operators \tilde{Y}_a^π of Eq. (25) form.*

Proof. Suppose Bob has an optimal strategy $\{Y_a^\pi\}$. Then

$$\begin{aligned}
\mathrm{Pr}[\text{Success}] &= \sum_{a,\pi} \mathrm{Tr}[(X_{a,\pi} \otimes Y_a^\pi) \cdot (\mathbb{1} \otimes \mathcal{D}_\varepsilon^{\otimes n})(\Phi^{\otimes n})] \\
&= \sum_{a,\pi} \mathrm{Tr}\left[\left(X_{a,\pi} \otimes \mathcal{D}_\varepsilon^{\dagger \otimes n}(Y_a^\pi)\right) \cdot \Phi^{\otimes n}\right] \\
&= 2^{-n} \sum_{a,\pi} \mathrm{Tr}\left[X_{a,\pi} \cdot \mathcal{D}_\varepsilon^{\dagger \otimes n}(Y_a^\pi)^T\right] \\
&= 2^{-n} \sum_{a,\pi} \mathrm{Tr}\left[\mathcal{D}_\varepsilon^{\otimes n}(X_{a,\pi}) \cdot (Y_a^\pi)^T\right].
\end{aligned}$$

Notice that for each a, π , the operator $\mathcal{D}_\varepsilon^{\otimes n}(X_{a,\pi})$ is of Eq. (25) form. Let

$$\tilde{Y}_a^\pi = \sum_{b \in \{0,1\}^n} \bar{\Pi}_b Y_a^\pi \bar{\Pi}_b.$$

By Lemma 5.2, we only need to check they form a POVM measurement. For each π ,

$$\sum_a \tilde{Y}_a^\pi = \sum_{a,b} \bar{\Pi}_b Y_a^\pi \bar{\Pi}_b = \sum_b \bar{\Pi}_b \left(\sum_a Y_a^\pi \right) \bar{\Pi}_b = \sum_b \bar{\Pi}_b \mathbb{1} \bar{\Pi}_b = \sum_b \bar{\Pi}_b = (\bar{\Pi}_0 + \bar{\Pi}_1)^{\otimes n} = \mathbb{1}_{3^n}.$$

□

We are now ready to prove our main theorem for this section.

Theorem 5.4. *Let $0 \leq \varepsilon \leq 1/2$. Suppose Alice and Bob share infinitely many pairs of $(\mathbb{1} \otimes \mathcal{D}_\varepsilon)(\Phi)$, and Alice is allowed to send classical messages to Bob. Then for any $k \geq 1, \gamma \in (0, 1)$, to produce a common random string $R \in \{0, 1\}^*$ with $H_{\min}(R) \geq k$ and a success probability at least $2^{-\gamma k}$, Alice needs to send at least*

$$\left(\varepsilon'(1 - \gamma) - 2\sqrt{\varepsilon'(1 - \varepsilon')\gamma}\right) \cdot k$$

bits, where $\varepsilon' = \varepsilon/2$.

Remark 5.5. Grusuwami and Radhakrishnan [GR16] showed that it suffices for Alice to send $(\varepsilon(1 - \gamma) - 2\sqrt{\varepsilon(1 - \varepsilon)\gamma})k$ bits to Bob to achieve the same task, even if they share erased-noisy random strings. Hence, the lower bound given in Theorem 5.4 is tight up to a constant.

Proof of Theorem 5.4. Our proof follows closely the proof for the classical counterpart [GR16].

$$\begin{aligned}
\Pr[\text{Success}] &= \sum_{a,\pi} \text{Tr} \left[(X_{a,\pi} \otimes Y_a^\pi) \cdot (\mathbb{1} \otimes \mathcal{D}_\varepsilon^{\otimes n}) (\Phi^{\otimes n}) \right] \\
&= \sum_{a,\pi} \text{Tr} \left[\left(X_{a,\pi} \otimes \mathcal{D}_\varepsilon^{\dagger \otimes n} (Y_a^\pi) \right) \cdot \Phi^{\otimes n} \right] \\
&= 2^{-n} \sum_{a,\pi} \text{Tr} \left[X_{a,\pi} \cdot \mathcal{D}_\varepsilon^{\dagger \otimes n} (Y_a^\pi)^T \right] \\
&= 2^{-n} \sum_{a,\pi} \text{Tr} \left[\mathcal{D}_\varepsilon^{\otimes n} (X_{a,\pi}) \cdot (Y_a^\pi)^T \right] \\
&= 2^{-n} \sum_{a,\pi} \text{Tr} \left[\Pi_\varepsilon^{\otimes n} \mathcal{D}^{\otimes n} (X_{a,\pi}) \cdot (Y_a^\pi)^T \right].
\end{aligned}$$

Then, by Lemma 2.1 with parameters q and q^* , since $(Y_a^\pi)^T$ commutes with Π_ε , we have

$$\Pr[\text{Success}] \leq 2^{-n} \sum_{\pi} \left(\sum_a \text{Tr} \left[\Pi_\varepsilon^{\otimes n} \mathcal{D}^{\otimes n} (X_{a,\pi})^q \right] \right)^{1/q} \cdot \left(\sum_a \text{Tr} \left[\Pi_\varepsilon^{\otimes n} (Y_a^\pi)^{Tq^*} \right] \right)^{1/q^*}.$$

We first upper bound the second term. Since $(Y_a^\pi)^T \leq \mathbb{1}_3$, we have $(Y_a^\pi)^{Tq^*} \leq (Y_a^\pi)^T$, and thus

$$\sum_a \text{Tr} \left[\Pi_\varepsilon (Y_a^\pi)^{Tq^*} \right] \leq \sum_a \text{Tr} \left[\Pi_\varepsilon (Y_a^\pi)^T \right] = \text{Tr} \left[\Pi_\varepsilon \sum_a (Y_a^\pi)^T \right] = \text{Tr} \left[\Pi_\varepsilon^{\otimes n} \mathbb{1}_{3^n} \right] = 2^n. \quad (45)$$

Applying Eq. (45), we get

$$\Pr[\text{Success}] \leq 2^{-n+n/q^*} \sum_{\pi} \left(\sum_a \text{Tr} \left[\Pi_\varepsilon^{\otimes n} \mathcal{D}^{\otimes n} (X_{a,\pi})^q \right] \right)^{1/q}.$$

Now let t be the number of bits sent from Alice to Bob. Then the number of different choices of the message π is at most 2^t . So by the concavity of the function $x \mapsto x^{1/q}$, we can put the summation over π into the parenthesis and get

$$\Pr[\text{Success}] \leq 2^{-n+n/q^*+t/q^*} \left(\sum_{a,\pi} \text{Tr} \left[\Pi_\varepsilon^{\otimes n} \mathcal{D}^{\otimes n} (X_{a,\pi})^q \right] \right)^{1/q}.$$

We now apply hypercontractivity Theorem 3.2 with $p = 1 + (1 - \epsilon)^{1/c}(q - 1)$ and $c = 2$ to get

$$\Pr[\text{Success}] \leq 2^{t/q^*} \left(\sum_{a,\pi} (2^{-n} \text{Tr}[X_{a,\pi}^p])^{q/p} \right)^{1/q}.$$

From the assumption that $H_{\min}(R) \geq k$, we have $2^{-n} \text{Tr}[X_{a,\pi}] \leq 2^{-k}$. Notice $\|X_{a,\pi}\| \leq 1$. Hence,

$$\begin{aligned} \Pr[\text{Success}] &\leq 2^{t/q^*} \left(\sum_{a,\pi} (2^{-n} \text{Tr}[X_{a,\pi}])^{q/p} \right)^{1/q} \\ &= 2^{t/q^*} \left(\sum_{a,\pi} (2^{-n} \text{Tr}[X_{a,\pi}]) \cdot (2^{-n} \text{Tr}[X_{a,\pi}])^{(q-p)/p} \right)^{1/q} \\ &\leq 2^{t/q^*} \left(\sum_{a,\pi} (2^{-n} \text{Tr}[X_{a,\pi}]) \cdot 2^{-k(q-p)/p} \right)^{1/q} \\ &= 2^{t/q^* - k \frac{q-p}{pq}}. \end{aligned}$$

Let $\delta = q-1$, then $q = 1+\delta$ and $p = 1+(1-\epsilon)^{1/c}\delta$. From the assumption that $\Pr[\text{Success}] \geq 2^{-\gamma k}$, we get for every $\delta > 0$,

$$\begin{aligned} t &\geq -\frac{(1+\delta)\gamma k}{\delta} + \frac{1 - (1 - \epsilon)^{1/c}}{1 + (1 - \epsilon)^{1/c}\delta} k \\ &\geq -\frac{(1+\delta)\gamma k}{\delta} + \frac{\epsilon/c}{1 + (1 - \epsilon/c)\delta} k \\ &= \left(\frac{\epsilon/c}{1 + (1 - \epsilon/c)\delta} - \frac{\gamma}{\delta} - \gamma \right) \cdot k. \end{aligned}$$

Maximizing over δ , with $c = 2$, we get

$$t \geq \left((\epsilon/c)(1 - \gamma) - 2\sqrt{(\epsilon/c)(1 - \epsilon/c)\gamma} \right) k.$$

□

REFERENCES

[AC98] Rudolf Ahlswede and Imre Csiszár. Common randomness in information theory and cryptography. II. CR capacity. *IEEE Transactions on Information Theory*, 44(1):225–240, 1998. 4

[AD24] Srinivasan Arunachalam and Joao F. Doriguello. Matrix hypercontractivity, streaming algorithms and LDCs: the large alphabet case. *ACM Trans. Comput. Theory*, 16(4), November 2024. [2](#)

[Aud06] Koenraad M.R. Audenaert. A norm compression inequality for block partitioned positive semidefinite matrices. *Linear Algebra and its Applications*, 413(1):155–176, 2006. [12](#)

[Bar17] Ivan Bardet. Estimating the decoherence time using non-commutative functional inequalities. *arXiv preprint arXiv:1710.01039*, 2017. [18](#)

[BARdW08] A. Ben-Aroya, O. Regev, and R. de Wolf. A hypercontractive inequality for matrix-valued functions with applications to quantum computing and LDCs. In *2008 IEEE 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 477–486, Los Alamitos, CA, USA, oct 2008. IEEE Computer Society. [2](#)

[BB21] Roland Bauerschmidt and Thierry Bodineau. Log-Sobolev inequality for the continuum sine-Gordon model. *Communications on Pure and Applied Mathematics*, 74(10):2064–2113, 2021. [4](#)

[BCG⁺23] Ivan Bardet, Ángela Capel, Li Gao, Angelo Lucia, David Pérez-García, and Cambyse Rouzé. Rapid thermalization of spin chain commuting Hamiltonians. *Phys. Rev. Lett.*, 130:060401, Feb 2023. [4](#)

[BDR20] Salman Beigi, Nilanjana Datta, and Cambyse Rouzé. Quantum reverse hypercontractivity: its tensorization and application to strong converses. *Communications in Mathematical Physics*, 376:753–794, 2020. [7](#), [9](#), [11](#), [12](#), [15](#), [18](#), [19](#)

[BDS97] Charles H Bennett, David P DiVincenzo, and John A Smolin. Capacities of quantum erasure channels. *Physical Review Letters*, 78(16):3217, 1997. [3](#)

[Bec75] William Beckner. Inequalities in Fourier analysis. *Annals of Mathematics*, 102(1):159–182, 1975. [2](#)

[Bei21] Salman Beigi. Improved Quantum Hypercontractivity Inequality for the Qubit Depolarizing Channel. *Journal of Mathematical Physics*, 62(12):122201, 12 2021. [2](#), [3](#), [9](#), [15](#), [18](#)

[BGL⁺14] Dominique Bakry, Ivan Gentil, Michel Ledoux, et al. *Analysis and Geometry of Markov Diffusion Operators*, volume 103. Springer, 2014. [4](#)

[BH99] Th Bodineau and B Helffer. The log-Sobolev inequality for unbounded spin systems. *Journal of functional analysis*, 166(1):168–178, 1999. [4](#)

[BKW14] Itai Benjamini, Gady Kozma, and Nicholas Wormald. The mixing time of the giant component of a random graph. *Random Structures & Algorithms*, 45(3):383–407, 2014. [12](#)

[BL98] S.G Bobkov and M Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *Journal of Functional Analysis*, 156(2):347–365, 1998. [15](#)

[BM00] F. Barthe and B. Maurey. Some remarks on isoperimetry of Gaussian type. *Annales de l’Institut Henri Poincaré (B) Probability and Statistics*, 36(4):419–434, 2000. [15](#)

[Bon68] Aline Bonami. Ensembles $\lambda(p)$ dans le dual de D^∞ . *Annales de l’institut Fourier*, 18(2):193–204, 1968. [1](#), [2](#)

[Bon70] Aline Bonami. Étude des coefficients de Fourier des fonctions de $L^p(G)$. *Annales de l’Institut Fourier*, 20:335–402, 1970. [1](#)

[BP21] Andrej Bogdanov and Gautam Prakriya. Direct Sum and Partitionability Testing over General Groups. In Nikhil Bansal, Emanuela Merelli, and James Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021)*, volume 198 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 33:1–33:19, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. [3](#)

[BR22] Ivan Bardet and Cambyse Rouzé. Hypercontractivity and logarithmic Sobolev inequality for non-primitive quantum Markov semigroups and estimation of decoherence rates. *Annales Henri Poincaré*, 23(11):3839–3903, November 2022. [2](#), [4](#), [12](#)

[CCNW21] Hong-Bin Chen, Sinho Chewi, and Jonathan Niles-Weed. Dimension-free log-Sobolev inequalities for mixture distributions. *Journal of Functional Analysis*, 281(11):109236, 2021. [4](#)

[CGM21] Mary Cryan, Heng Guo, and Giorgos Mousa. Modified log-Sobolev inequalities for strongly log-concave distributions. *The Annals of Probability*, 49(1):506 – 525, 2021. [4](#)

[CHK12] David Croydon, Ben Hambly, and Takashi Kumagai. Convergence of mixing times for sequences of random walks on finite graphs. *Electronic Journal of Probability*, 17(none):1 – 32, 2012. [12](#)

[CKMT15] Toby Cubitt, Michael Kastoryano, Ashley Montanaro, and Kristan Temme. Quantum Reverse Hypercontractivity. *Journal of Mathematical Physics*, 56(10):102204, 10 2015. [18](#)

[CL93] Eric A. Carlen and Elliott H. Lieb. Optimal Hypercontractivity for Fermi Fields and Related Noncommutative Integration Inequalities. *Communications in Mathematical Physics*, 155(1):27 – 46, 1993. [2](#)

[CM15] Raffaella Carbone and Andrea Martinelli. Logarithmic Sobolev inequalities in non-commutative algebras. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 18(02):1550011, 2015. [4](#)

[CM20] Eric A. Carlen and Jan Maas. Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems. *Journal of Statistical Physics*, 178(2):319–378, Jan 2020. [4](#)

[DKQ⁺23] Dawei Ding, Sumeet Khatri, Yihui Quek, Peter W. Shor, Xin Wang, and Mark M. Wilde. Bounding the forward classical capacity of bipartite quantum channels. *IEEE Transactions on Information Theory*, 69(5):3034–3061, 2023. [4](#), [12](#)

[DW04] Igor Devetak and Andreas Winter. Distilling common randomness from bipartite quantum states. *IEEE Transactions on Information Theory*, 50(12):3183–3196, 2004. [4](#)

[DY24] Yangjing Dong and Penghui Yao. Communication complexity of common randomness generation with isotropic states. *IEEE Trans. Inf. Theor.*, 70(8):5681–5691, August 2024. [5](#)

[DZ13] Nicolas Delfosse and Gilles Zémor. Upper bounds on the rate of low density stabilizer codes for the quantum erasure channel. *Quantum Info. Comput.*, 13(9–10):793–826, sep 2013. [3](#)

[DZ20] Nicolas Delfosse and Gilles Zémor. Linear-time maximum likelihood decoding of surface codes over the quantum erasure channel. *Phys. Rev. Res.*, 2:033042, Jul 2020. [3](#)

[EWW23] Ronen Eldan, Avi Wigderson, and Pei Wu. An optimal “it ain’t over till it’s over” theorem. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, page 853–866, New York, NY, USA, 2023. Association for Computing Machinery. [3](#), [5](#), [12](#)

[FHL⁺17] Uwe Franz, Guixiang Hong, François Lemeux, Michaël Ulrich, and Haonan Zhang. Hypercontractivity of heat semigroups on free quantum groups. *Journal of Operator Theory*, 77(1):pp. 61–76, 2017. [2](#)

[FMP13] Alessio Figalli, Francesco Maggi, and Aldo Pratelli. Sharp stability theorems for the anisotropic Sobolev and log-Sobolev inequalities on functions of bounded variation. *Advances in Mathematics*, 242:80–101, 2013. [4](#)

[FOW22] Yuval Filmus, Ryan O’Donnell, and Xinyu Wu. Log-Sobolev inequality for the multislice, with applications. *Electronic Journal of Probability*, 27:1 – 30, 2022. [4](#), [9](#)

[GBP97] Markus Grassl, Th Beth, and Thomas Pellizzari. Codes for the quantum erasure channel. *Physical Review A*, 56(1):33, 1997. [3](#)

[GKK⁺08] Dmitry Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald de Wolf. Exponential separation for one-way quantum communication complexity, with applications to cryptography. *SIAM J. Comput.*, 38(5):1695–1708, December 2008. [2](#)

[Gli68] James Glimm. Boson Fields with Nonlinear Selfinteraction in Two Dimensions. *Communications in Mathematical Physics*, 8(1):12 – 25, 1968. [1](#)

[GR16] Venkatesan Guruswami and Jaikumar Radhakrishnan. Tight Bounds for Communication-Assisted Agreement Distillation. In Ran Raz, editor, *31st Conference on Computational Complexity (CCC 2016)*, volume 50 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 6:1–6:17, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. [3](#), [5](#), [8](#), [27](#)

[Gro75a] Leonard Gross. Hypercontractivity and Logarithmic Sobolev Inequalities for the Clifford-Dirichlet Form. *Duke Mathematical Journal*, 42(3):383 – 396, 1975. [2](#), [4](#), [9](#), [10](#), [18](#)

[Gro75b] Leonard Gross. Logarithmic Sobolev inequalities. *American Journal of Mathematics*, 97(4):1061–1083, 1975. [15](#)

[IM14] Emanuel Indrei and Diego Marcon. A quantitative log-Sobolev inequality for a two parameter family of functions. *International Mathematics Research Notices*, 2014(20):5563–5580, 2014. [4](#)

[KB16] Michael J. Kastoryano and Fernando G. S. L. Brandão. Quantum Gibbs samplers: The commuting case. *Communications in Mathematical Physics*, 344(3):915–957, Jun 2016. [4](#)

[KCP16] Santhosh Kumar, Robert Calderbank, and Henry D. Pfister. Reed-Muller codes achieve capacity on the quantum erasure channel. In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 1750–1754, 2016. [3](#)

[Kie69] Konrad Kiener. *Über Produkte von quadratisch integrierbaren Funktionen endlicher Vielfalt*. PhD thesis, PhD thesis, Universität Innsbruck, 1969. 1.1. 1, 1969. [1](#)

[Kin03] Christopher King. Inequalities for Trace Norms of 2×2 Block Matrices. *Communications in Mathematical Physics*, 242(3):531–545, January 2003. [9](#), [18](#), [40](#)

[Kin14] Christopher King. Hypercontractivity for semigroups of unital qubit channels. *Communications in Mathematical Physics*, 328(1):285–301, 2014. [2](#), [3](#), [4](#), [9](#), [11](#), [18](#)

[KKK⁺21] Esty Kelman, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. Theorems of KKL, Friedgut, and Talagrand via Random Restrictions and Log-Sobolev Inequality. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 26:1–26:17, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. [9](#)

[KKL88] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, SFCS ’88, page 68–80, USA, 1988. IEEE Computer Society. [2](#)

[KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007. [2](#)

[KT13] Michael J Kastoryano and Kristan Temme. Quantum logarithmic Sobolev inequalities and rapid mixing. *Journal of Mathematical Physics*, 54(5), 2013. [2](#), [4](#), [8](#), [9](#), [10](#), [11](#), [21](#)

[LMN93] Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, fourier transform, and learnability. *J. ACM*, 40(3):607–620, jul 1993. [2](#)

[LRWW23] Ludovico Lami, Bartosz Regula, Xin Wang, and Mark M. Wilde. Upper bounds on the distillable randomness of bipartite quantum states. In *2023 IEEE Information Theory Workshop (ITW)*, pages 203–208, 2023. [4](#), [12](#)

[Mau93] Ueli M Maurer. Secret key agreement by public discussion from common information. *IEEE transactions on information theory*, 39(3):733–742, 1993. [4](#)

[MHSFW16] Alexander Müller-Hermes, Daniel Stilck França, and Michael M. Wolf. Entropy production of doubly stochastic quantum channels. *Journal of Mathematical Physics*, 57(2):022203, 02 2016. [2](#), [3](#), [11](#)

[MO10] Ashley Montanaro and Tobias J Osborne. Quantum Boolean functions. *Chicago Journal OF Theoretical Computer Science*, 1:1–45, 2010. [11](#)

[MOA79] Albert W Marshall, Ingram Olkin, and Barry C Arnold. *Inequalities: Theory of Majorization and its Applications*. Springer, 1979. [14](#)

[MOO10] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Annals of Mathematics*, pages 295–341, 2010. [2](#), [3](#)

[Nel66] Edward Nelson. A quartic interaction in two dimensions. In *Mathematical Theory of Elementary Particles, Proc. Conf., Dedham, Mass., 1965*, pages 69–73. MIT press, 1966. [1](#)

[Nel73] Edward Nelson. The free Markoff field. *Journal of Functional Analysis*, 12(2):211–227, 1973. [1](#)

[NP08] Asaf Nachmias and Yuval Peres. Critical random graphs: Diameter and mixing time. *The Annals of Probability*, 36(4):1267 – 1286, 2008. [12](#)

[NW16] Chandra Nair and Yan Nan Wang. Evaluating hypercontractivity parameters using information measures. In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 570–574, 2016. [3](#), [5](#), [7](#), [8](#), [11](#), [40](#)

[NW24] Theshani Nuradha and Mark M. Wilde. Fidelity-based smooth min-relative entropy: Properties and applications. *IEEE Trans. Inf. Theor.*, 70(6):4170–4196, June 2024. [4](#), [12](#)

[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, New York, 2014. [2](#)

[OZ99] Robert Olkiewicz and Bogusław Zegarlinski. Hypercontractivity in noncommutative L_p spaces. *Journal of Functional Analysis*, 161(1):246–285, 1999. [2](#), [3](#), [4](#), [9](#)

[Pal32] REAC Paley. A remarkable series of orthogonal functions (I). *Proceedings of the London Mathematical Society*, 2(1):241–264, 1932. [1](#)

[PS19] Yury Polyanskiy and Alex Samorodnitsky. Improved log-sobolev inequalities, hypercontractivity and uncertainty principle on the hypercube. *Journal of Functional Analysis*, 277(11):108280, 2019. [9](#), [11](#)

[QY21] Minglong Qin and Penghui Yao. Nonlocal games with noisy maximally entangled states are decidable. *SIAM Journal on Computing*, 50(6):1800–1891, 2021. [2](#)

[QY23] Minglong Qin and Penghui Yao. Decidability of Fully Quantum Nonlocal Games with Noisy Maximally Entangled States. In Kousha Etessami, Uriel Feige, and Gabriele Puppis, editors, *50th International Colloquium on Automata, Languages, and Programming (ICALP 2023)*, volume 261 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 97:1–97:20, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. [2](#)

[Ras12] Alexey E Rastegin. Relations for certain symmetric norms and anti-norms before and after partial trace. *Journal of Statistical Physics*, 148:1040–1053, 2012. [37](#)

[SA13] Khalid Shebrawi and Hussien Albadawi. Trace inequalities for matrices. *Bulletin of the Australian Mathematical Society*, 87(1):139–148, 2013. [13](#)

[SFGP21] Daniel Stilck França and Raul García-Patrón. Limitations of optimization algorithms on noisy quantum devices. *Nature Physics*, 17(11):1221–1227, Nov 2021. [4](#)

[SHK72] Barry Simon and Raphael Høegh-Krohn. Hypercontractive semigroups and two dimensional self-coupled Bose fields. *Journal of Functional Analysis*, 9(2):121–180, 1972. [1](#)

[STW20] Madhu Sudan, Himanshu Tyagi, and Shun Watanabe. Communication for generating correlation: A unifying survey. *IEEE Transactions on Information Theory*, 66(1):5–37, 2020. [4](#)

[SZ24] Ludovic Stephan and Yizhe Zhu. Sparse random hypergraphs: non-backtracking spectra and community detection†. *Information and Inference: A Journal of the IMA*, 13(1):iaae004, 02 2024. [11](#)

[TPK14] Kristan Temme, Fernando Pastawski, and Michael J Kastoryano. Hypercontractivity of quasi-free quantum semigroups. *Journal of Physics A: Mathematical and Theoretical*, 47(40):405303, sep 2014. [2](#), [3](#), [4](#), [11](#)

[Wat05] John Watrous. Notes on super-operator norms induced by Schatten norms. *Quantum Info. Comput.*, 5(1):58–68, jan 2005. [9](#), [10](#), [11](#), [17](#), [35](#)

[Wat18] John Watrous. *The Theory of Quantum Information*. Cambridge university press, 2018. [13](#), [16](#)

[Wol07] Paweł Wolff. Hypercontractivity of simple random variables. *Studia Mathematica*, 180(3):219–236, 2007. [2](#)

[Zha19] Pingping Zhang. On some inequalities related to positive block matrices. *Linear Algebra and its Applications*, 576:258–267, 2019. Proceedings of the ILAS 2017 Conference in Ames, Iowa. [37](#)

[ZOJ23] Changchun Zhong, Changhun Oh, and Liang Jiang. Information transmission with continuous variable quantum erasure channels. *Quantum*, 7:939, March 2023. [3](#)

A PROOF OF LEMMA 3.1

Lemma 3.1 (Watrous’ theorem [Wat05]). *For $1 \leq p \leq q, \varepsilon \in [0, 1]$, we have*

$$\sup_{X \in \mathcal{M}_{2^n}, X \neq 0} \frac{\|X\|_{\varepsilon, q}}{\|X\|_p} = \sup_{X \in \text{Pos}_{2^n}, X \neq 0} \frac{\|X\|_{\varepsilon, q}}{\|X\|_p}.$$

Proof. Note that for any $X \in \mathcal{M}_{2^n}$,

$$(2^{-n} \text{Tr}[\Pi_\varepsilon \cdot |D(X)|^q])^{\frac{1}{q}} = \left(\sum_{S \subseteq [n]} (1 - \varepsilon)^{n - |S|} \varepsilon^{|S|} \|\tau_S[X]\|_q^q \right)^{\frac{1}{q}}.$$

Let $X \in \mathcal{M}_{2^n}$ satisfy $\|X\|_p = 1$. For any $S \subseteq [n]$, let $Y_{S^c} \in \mathcal{M}_{2^{n-|S|}}$ satisfy $\|Y_{S^c}\|_{q^*} = 1$, where q^* is the Hölder conjugate of q . Consider the singular value decompositions of X and Y_{S^c} :

$$X = \sum_{i=1}^{2^n} l_i |u_i\rangle\langle v_i| \text{ and } Y_{S^c} = \sum_{j=1}^{m_S} r_{S,j} |w_{S,j}\rangle\langle x_{S,j}|$$

where m_S denotes the number of singular values of Y_{S^c} .

Define $X_L, X_R, Y_{S^c,L}, Y_{S^c,R}$ as follows:

$$X_L = \sum_{i=1}^{2^n} l_i |u_i\rangle\langle u_i| \text{ and } \forall S \subseteq [n], Y_{S^c,L} = \sum_{j=1}^{m_S} r_{S,j} |w_{S,j}\rangle\langle w_{S,j}|,$$

$$X_R = \sum_{i=1}^{2^n} l_i |v_i\rangle\langle v_i| \text{ and } \forall S \subseteq [n], Y_{S^c, R} = \sum_{j=1}^{m_S} r_{S,j} |x_{S,j}\rangle\langle x_{S,j}|.$$

We point out that $\|X_L\|_p = \|X_R\|_p = \|Y_{S^c, L}\|_{q^*} = \|Y_{S^c, R}\|_{q^*} = 1$. Also, $X_L, X_R, Y_{S^c, L}, Y_{S^c, R}$ are positive semi-definite.

Therefore,

$$\begin{aligned} \sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \left| \tau \left[Y_{S^c}^\dagger \cdot (\tau_S X) \right] \right| &= \sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \frac{1}{2^{n-|S|}} \left| \sum_{j=1}^{m_S} r_{S,j} \langle w_{S,j} | \tau_S X | x_{S,j} \rangle \right| \\ &= \sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \frac{1}{2^n} \left| \sum_{j=1}^{m_S} \sum_{k=1}^{2^{|S|}} r_{S,j} \langle w_{S,j}, k | X | x_{S,j}, k \rangle \right| \\ &= \sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \frac{1}{2^n} \left| \sum_{i=1}^{2^n} \sum_{j=1}^{m_S} \sum_{k=1}^{2^{|S|}} l_i r_{S^c, j} \langle w_{S,j}, k | u_i \rangle \langle v_i | x_{S,j}, k \rangle \right| \\ &\leq \frac{1}{2^n} \sqrt{\sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \sum_{i,j,k} l_i r_{S,j} |\langle w_{S,j}, k | u_i \rangle|^2} \cdot \sqrt{\sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \sum_{i,j,k} l_i r_{S,j} |\langle v_i | x_{S,j}, k \rangle|^2} \\ &= \sqrt{\sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \left| \tau \left[Y_{S^c, L}^\dagger \cdot \tau_S X_L \right] \right|} \cdot \sqrt{\sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \left| \tau \left[Y_{S^c, R}^\dagger \cdot \tau_S X_R \right] \right|} \\ &\leq \sqrt{\sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \|Y_{S^c, L}\|_{q^*} \cdot \|\tau_S X_L\|_q} \cdot \sqrt{\sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \|Y_{S^c, R}\|_{q^*} \cdot \|\tau_S X_R\|_q} \\ &\leq \sqrt{\left(\sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \|Y_{S^c, L}\|_{q^*}^{q^*} \right)^{\frac{1}{q^*}} \left(\sum_{S \subseteq [n]} (1 - \varepsilon)^{n-|S|} \varepsilon^{|S|} \|Y_{S^c, R}\|_{q^*}^{q^*} \right)^{\frac{1}{q^*}}} \|X_L\|_{\varepsilon, q} \|X_R\|_{\varepsilon, q} \\ &= \sqrt{\|X_L\|_{\varepsilon, q} \cdot \|X_R\|_{\varepsilon, q}}. \end{aligned}$$

Taking the supremum over all Y_{S^c} , we conclude that for $\forall X \in \mathcal{M}_{2^n}$, satisfying $\|X\|_p = 1$,

$$\|X\|_{\varepsilon, q} \leq \sqrt{\|X_L\|_{\varepsilon, q} \cdot \|X_R\|_{\varepsilon, q}}.$$

Without loss of generality, suppose $\|X_L\|_{\varepsilon, q} \leq \|X_R\|_{\varepsilon, q}$, then

$$\|X\|_{\varepsilon, q} \leq \|X_R\|_{\varepsilon, q}.$$

Note that X_R is positive semi-definite and $\|X_R\|_p = 1$, so the lemma is proved. \square

B LEMMAS FOR PROOF OF CLAIM 4.4

We need two technical lemmas. Recall the two-point relative entropy function

$$\text{Ent}_\varepsilon[a, b] = (1 - \varepsilon) \cdot a \ln a + \varepsilon \cdot b \ln b - [(1 - \varepsilon)a + \varepsilon b] \ln[(1 - \varepsilon)a + \varepsilon b].$$

Lemma B.1. For $16b \geq a \geq b > 0$ and $0 \leq \varepsilon \leq 1$,

$$\text{Ent}_\varepsilon[a, b] \leq 2(1 - \varepsilon)(a - b). \quad (46)$$

Proof. Fix a and b . Let $f(\varepsilon) = \text{Ent}_\varepsilon[a, b] - 2(1 - \varepsilon)(a - b)$. We see that $f(1) = 0$, so we only need to prove that f is non-decreasing for ε in the interval $[0, 1]$. Calculating the derivatives of f over ε we get

$$\frac{d}{d\varepsilon}f(\varepsilon) = -a \ln a + b \ln b + 3(a - b) + (a - b) \ln((1 - \varepsilon)a + \varepsilon b),$$

and

$$\frac{d^2}{d\varepsilon^2}f(\varepsilon) = -\frac{(a - b)^2}{(1 - \varepsilon)a + \varepsilon b} \leq 0.$$

Thus f is concave. To prove $\frac{d}{d\varepsilon}f(\varepsilon) \geq 0$ for $\varepsilon \in [0, 1]$, we only need $\frac{d}{d\varepsilon}f(1) \geq 0$, which is

$$-a \ln a + 3(a - b) + a \ln b \geq 0$$

for any $16b \geq a \geq b$.

Now we only fix on a and define the function $g : [a/16, a] \rightarrow \mathbb{R}$ as

$$g(b) = -a \ln a + 3(a - b) + a \ln b.$$

Again we see that $g(b)$ is concave because $\frac{d^2}{db^2}g(b) = -\frac{1}{b^2} \leq 0$. To prove $g(b) \geq 0$ in the range $[a/16, a]$, it suffices to check

- $g(a) = -a \ln a + 3(a - a) + a \ln a = 0$.
- $g(a/16) = \frac{45}{16}a - a \ln 16 = a \cdot \left(\frac{45}{16} - \ln 16\right) \geq 0$ since $\ln 16 \approx 2.77 < \frac{45}{16} = 2.8125$.

□

Lemma B.2. Let $A \in \mathcal{M}_2 \times \mathcal{M}_n$ be positive semi-definite. Then for $q \geq 1$,

$$\tau(\tau_1[A])^q \leq \tau A^q \leq 4^q \tau(\tau_1[A])^q. \quad (47)$$

Proof. The inequality $\tau(\tau_1[A])^q \leq \tau A^q$ is Proposition 1 in [Ras12] in disguise. The inequality $\tau A^q \leq 4^q \tau(\tau_1[A])^q$ is a simple corollary of the inequality by Zhang [Zha19], who as Proposition 3.4 proved that

$$4\tau_1[A] \otimes \mathbb{1}_2 = 2\text{Tr}_1[A] \otimes \mathbb{1}_2 \succeq A.$$

Then

$$4^q(\tau_1[A])^q \otimes \mathbb{1}_2 \succeq A^q$$

and our lemma follows. □

C QEC HYPERCONTRACTIVITY IMPLIES QUANTUM DEPOLARIZING CHANNEL HYPERCONTRACTIVITY

Definition C.1 (quantum depolarizing channel). Let $0 \leq \rho \leq 1$, the quantum depolarizing channel Δ_ρ is defined as

$$\Delta_\rho(X) = \rho X + (1 - \rho)\tau X \cdot \mathbb{1}_2. \quad (48)$$

Lemma C.2. For any p and q , suppose Eq. (9) holds for any matrix X . Then for any Y of Eq. (25) form, it holds that

$$|2^{-n} \text{Tr}[Y^\dagger \mathcal{D}_\varepsilon^{\otimes n}(X)]| \leq \|X\|_p \left(2^{-n} \text{Tr}[\Pi_\varepsilon^{\otimes n} |Y|^{q^*}] \right)^{1/q^*}, \quad (49)$$

$$\left(2^{-n} \text{Tr}[\mathcal{D}_\varepsilon^{\otimes n\dagger}(Y)]^{p^*}\right)^{1/p^*} \leq \left(2^{-n} \text{Tr}[\Pi_\varepsilon^{\otimes n} |Y|^{q^*}] \right)^{1/q^*}, \quad (50)$$

where p^* and q^* are the Hölder conjugates of p and q , respectively. Moreover, Eq. (9), Eq. (49), Eq. (50) are all equivalent to each other.

Proof.

$$\bullet \text{ Eq. (9)} \implies \text{Eq. (49)}$$

$$\begin{aligned} |2^{-n} \text{Tr}[Y^\dagger \mathcal{D}_\varepsilon^{\otimes n}(X)]| &= |2^{-n} \text{Tr}[\Pi_\varepsilon^{\otimes n} Y^\dagger \mathcal{D}_\varepsilon^{\otimes n}(X)]| \\ &\leq (2^{-n} \text{Tr}[\Pi_\varepsilon^{\otimes n} |\mathcal{D}_\varepsilon^{\otimes n}(X)|^q])^{1/q} \left(2^{-n} \text{Tr}[\Pi_\varepsilon^{\otimes n} |Y|^{q^*}] \right)^{1/q^*} \\ &\leq \|X\|_p \left(2^{-n} \text{Tr}[\Pi_\varepsilon^{\otimes n} |Y|^{q^*}] \right)^{1/q^*}. \end{aligned}$$

$$\bullet \text{ Eq. (49)} \implies \text{Eq. (9)}$$

$$\begin{aligned} &(2^{-n} \text{Tr}[\Pi_\varepsilon^{\otimes n} |\mathcal{D}_\varepsilon^{\otimes n}(X)|^q])^{1/q} \\ &= \max_Y \left\{ \left| 2^{-n} \text{Tr}[Y^\dagger (\Pi_\varepsilon^{\otimes n})^{1/q} \mathcal{D}_\varepsilon^{\otimes n}(X)] \right| : Y \in \mathcal{M}_{3^n}, \|Y\|_{q^*} \leq 1 \right\} \\ &= \max_Y \left\{ \left| 2^{-n} \text{Tr}[Y^\dagger \Pi_\varepsilon^{\otimes n} \mathcal{D}_\varepsilon^{\otimes n}(X)] \right| : Y \in \mathcal{M}_{3^n}, \left(2^{-n} \text{Tr}[\Pi_\varepsilon^{\otimes n} |Y|^{q^*}] \right)^{1/q^*} \leq 1 \right\} \\ &\leq \|X\|_p. \end{aligned}$$

where the maximizations are over Y of Eq. (25) form and the first equality follows from Eq. (21) and Lemma 5.2.

$$\bullet \text{ Eq. (50)} \implies \text{Eq. (49)}$$

$$\begin{aligned} |2^{-n} \text{Tr}[Y^\dagger \mathcal{D}_\varepsilon^{\otimes n}(X)]| &= \left| 2^{-n} \text{Tr}[\mathcal{D}_\varepsilon^{\otimes n\dagger}(Y^\dagger) X] \right| \\ &\leq \|X\|_p \left(2^{-n} \text{Tr}[\mathcal{D}_\varepsilon^{\otimes n\dagger}(Y)]^{p^*}\right)^{1/p^*} \\ &\leq \|X\|_p \left(2^{-n} \text{Tr}[\Pi_\varepsilon^{\otimes n} |Y|^{q^*}] \right)^{1/q^*}. \end{aligned}$$

- Eq. (49) \implies Eq. (50)

$$\begin{aligned}
\left(2^{-n} \operatorname{Tr} \left| \mathcal{D}_\varepsilon^{\otimes n\dagger}(Y) \right|^{p^*} \right)^{1/p^*} &= \max_X \left\{ \left| 2^{-n} \operatorname{Tr} \left[X^\dagger \mathcal{D}_\varepsilon^{\otimes n\dagger}(Y) \right] \right| : \|X\|_p \leq 1 \right\} \\
&= \max_X \left\{ \left| 2^{-n} \operatorname{Tr} \left[\mathcal{D}_\varepsilon^{\otimes n}(X^\dagger) Y \right] \right| : \|X\|_p \leq 1 \right\} \\
&\leq \|X\|_p \left(2^{-n} \operatorname{Tr} \left[\Pi_\varepsilon^{\otimes n} |Y|^{q^*} \right] \right)^{1/q^*} \\
&\leq \left(2^{-n} \operatorname{Tr} \left[\Pi_\varepsilon^{\otimes n} |Y|^{q^*} \right] \right)^{1/q^*}.
\end{aligned}$$

where the first equality is Eq. (21). \square

Lemma C.3. Suppose there exist parameters ε, p, q, r such that Eq. (9) holds for

- $q \leftarrow r^*$ and $p \leftarrow q^*$ and
- $q \leftarrow r$ and $p \leftarrow p$,

where r^* and q^* are the Hölder conjugates of r and q , respectively. Then for any matrix X , for $\rho \leq 1 - \varepsilon$, we have the hypercontractivity inequality

$$\|\Delta_\rho^{\otimes n}(X)\|_q \leq \|X\|_p. \quad (51)$$

Proof. Note that

$$\Delta_\rho^{\otimes n}(X) = \mathcal{D}_\varepsilon^{\otimes n\dagger}(\mathcal{D}^{\otimes n}(X)).$$

So we have

$$\begin{aligned}
\|\Delta_\rho^{\otimes n}(X)\|_q &= \left(2^{-n} \operatorname{Tr} \left| \mathcal{D}_\varepsilon^{\otimes n\dagger}(\mathcal{D}^{\otimes n}(X)) \right|_q^q \right)^{1/q} \\
&\leq \left(2^{-n} \operatorname{Tr} \left[\Pi_\varepsilon^{\otimes n} |\mathcal{D}^{\otimes n}(X)|^r \right] \right)^{1/r} \\
&\leq \|X\|_p.
\end{aligned}$$

\square

Corollary C.4. Suppose Eq. (9) holds for all $1 - \varepsilon \leq \left(\frac{p-1}{q-1} \right)^c$ and any matrix X , then Eq. (51) holds for all $\rho = 1 - \varepsilon \leq \left(\frac{p-1}{q-1} \right)^{c/2}$ and any matrix X .

Proof. Let $r = 1 + (p-1)/(1-\varepsilon)^{1/c}$, then we can check that

$$\left(\frac{r-1}{q-1} \right)^c = 1 - \varepsilon,$$

and

$$\left(\frac{p-1}{r-1} \right)^c = 1 - \varepsilon.$$

Then we can apply Lemma C.3 with the parameters ε, p, q, r to prove Eq. (51). \square

D PROOF OF HYPERCONTRACTIVITY FOR $1 \leq p \leq 2 \leq q$

Lemma D.1 ([Kin03]). *Let M be a $2n \times 2n$ positive semidefinite matrix. It can be written in the block form*

$$M = \begin{pmatrix} X & Y \\ Y^\dagger & Z \end{pmatrix} \quad (52)$$

Define the 2×2 matrix

$$m_p = \begin{pmatrix} \|X\|_p & \|Y\|_p \\ \|Y\|_p & \|Z\|_p \end{pmatrix} \quad (53)$$

Then m_p is positive semidefinite and

- for $1 \leq p \leq 2$

$$\|M\|_p \geq \|m_p\|_p \quad (54)$$

- for $2 \leq p \leq \infty$

$$\|M\|_p \leq \|m_p\|_p \quad (55)$$

Lemma D.2. *Let X, Y be two 2×2 real valued positive semidefinite matrices such that for all $i, j \in \{1, 2\}$*

$$X_{ij} \leq Y_{ij} \quad (56)$$

then for any Schatten p -norm $\|\cdot\|$ we have

$$\|X\| \leq \|Y\| \quad (57)$$

Proof. Define the intermediate matrix

$$Z = \begin{pmatrix} Y_{11} & X_{12} \\ X_{12} & Y_{22} \end{pmatrix} = X + \begin{pmatrix} Y_{11} - X_{11} & \\ & Y_{22} - X_{22} \end{pmatrix} \quad (58)$$

it is clear that $X \preceq Z$ and thus $\|X\| \leq \|Z\|$. We only need to prove $\|Z\| \leq \|Y\|$.

Let $\lambda_1 \geq \lambda_2 \geq 0$ be the eigenvalues of Y and $\gamma_1 \geq \gamma_2 \geq 0$ be the eigenvalues of Z . Since $\text{Tr } Y = \text{Tr } Z$, we have $\lambda_1 + \lambda_2 = \gamma_1 + \gamma_2$. Since $\det(Y) \leq \det(Z)$, we have $\lambda_1 \cdot \lambda_2 \leq \gamma_1 \cdot \gamma_2$. Thus $\{\lambda_1, \lambda_2\}$ majorises $\{\gamma_1, \gamma_2\}$ and thus $\|Y\| \geq \|Z\|$. \square

Lemma D.3. *Let $A \in \mathcal{M}_{2^n}$, then for all $1 \leq p \leq 2 \leq q$ and $0 \leq \varepsilon \leq 1$ satisfying $1 - \varepsilon \leq \frac{p-1}{q-1}$,*

$$(\text{Tr} [\Pi_\varepsilon^{\otimes n} |\mathcal{D}^{\otimes n}(X)|^q])^{1/q} \leq \|X\|_p \cdot 2^{n \cdot (\frac{1}{q} - \frac{1}{p})}. \quad (59)$$

Proof. By Lemma 3.1, we can assume X to be positive semi-definite. For $n = 1$, the inequality collapses to the classical case and has been proven by Nair and Wang [NW16]. For $n > 1$, we can write X as a block matrix

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$$

then

$$(\Pi_\varepsilon^{\otimes n})^{1/q} \mathcal{D}^{\otimes n}(X) = \begin{pmatrix} Z_{11} & Z_{12} & \\ Z_{21} & Z_{22} & \\ & & Z_{33} \end{pmatrix}$$

where we let $Y_{ij} = (\Pi_\varepsilon^{\otimes n-1})^{1/q} \mathcal{D}^{\otimes n-1}(X_{ij})$ for $i, j \in \{1, 2\}$ and $Z_{ij} = (1 - \varepsilon)^{1/q} Y_{ij}$ for $i, j \in \{1, 2\}$ and $Z_{33} = (2\varepsilon)^{1/q} \frac{Y_{11} + Y_{22}}{2}$. Since $q \geq 2$, Lemma D.1 implies

$$\begin{aligned} (\mathrm{Tr}[\Pi_\varepsilon^{\otimes n} \mathcal{D}^{\otimes n}(X)^q])^{1/q} &= \left(\mathrm{Tr} \left[\begin{pmatrix} Z_{11} & Z_{12} & \\ Z_{21} & Z_{22} & \\ & & Z_{33} \end{pmatrix}^q \right] \right)^{1/q} \\ &= \left\| \begin{pmatrix} \left\| \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} \right\|_q & \\ & \left\| Z_{33} \right\|_q \end{pmatrix} \right\|_q \\ &\leq \left\| \begin{pmatrix} \left\| \begin{pmatrix} \left\| Z_{11} \right\|_q & \left\| Z_{12} \right\|_q \\ \left\| Z_{12} \right\|_q & \left\| Z_{22} \right\|_q \end{pmatrix} \right\|_q & \\ & \quad \left\| Z_{33} \right\|_q \end{pmatrix} \right\|_q \end{aligned}$$

By triangle inequality we have

$$\left\| Z_{33} \right\|_q = (2\varepsilon)^{1/q} \left\| \frac{Y_{11} + Y_{22}}{2} \right\|_q \leq \frac{(2\varepsilon)^{1/q}}{2} \left\| Y_{11} \right\|_q + \frac{(2\varepsilon)^{1/q}}{2} \left\| Y_{22} \right\|_q$$

Thus

$$(\mathrm{Tr}[\Pi_\varepsilon^{\otimes n} \mathcal{D}^{\otimes n}(X)^q])^{1/q} \leq \left\| \Pi_\varepsilon^{1/q} \mathcal{D} \left(\begin{pmatrix} \left\| Y_{11} \right\|_q & \left\| Y_{12} \right\|_q \\ \left\| Y_{12} \right\|_q & \left\| Y_{22} \right\|_q \end{pmatrix} \right) \right\|_q \quad (60)$$

By induction hypothesis we also have

$$\left\| Y_{ij} \right\|_q = (\mathrm{Tr} \Pi_\varepsilon^{\otimes n-1} \mathcal{D}^{\otimes n-1}(X_{ij})^q)^{1/q} \leq \|X_{ij}\|_p \cdot 2^{(n-1) \cdot (\frac{1}{q} - \frac{1}{p})} \quad (61)$$

thus by Lemma D.2 and again by induction hypothesis

$$\begin{aligned} (\mathrm{Tr}[\Pi_\varepsilon^{\otimes n} \mathcal{D}^{\otimes n}(X)^q])^{1/q} &\leq \left\| \Pi_\varepsilon^{1/q} \mathcal{D} \left(\begin{pmatrix} \|X_{11}\|_p & \|X_{12}\|_p \\ \|X_{12}\|_p & \|X_{22}\|_p \end{pmatrix} \right) \right\|_q \cdot 2^{(n-1) \cdot (\frac{1}{q} - \frac{1}{p})} \\ &\leq \left\| \begin{pmatrix} \|X_{11}\|_p & \|X_{12}\|_p \\ \|X_{12}\|_p & \|X_{22}\|_p \end{pmatrix} \right\|_p \cdot 2^{n \cdot (\frac{1}{q} - \frac{1}{p})} \quad (62) \end{aligned}$$

Finally we use Lemma D.1 again for $1 \leq p \leq 2$ which implies

$$\left\| \begin{pmatrix} \|X_{11}\|_p & \|X_{12}\|_p \\ \|X_{12}\|_p & \|X_{22}\|_p \end{pmatrix} \right\|_p \leq \left\| \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^\dagger & X_{22} \end{pmatrix} \right\|_p = \|X\|_p$$

and the lemma is proved. \square