Strategies for adiabatic state preparation of quantum many-body systems

Dyon van Vreumingen • Strategies for adiabatic state preparation of quantum many-body systems

Strategies for adiabatic state preparation of quantum many-body systems

University of Amsterdam
Institute of Physics

The research for this doctoral thesis was supported by the Dutch Ministry of Economic Affairs and Climate Policy (EZK), through the Quantum Delta NL programme.

Printed and bound by Ipskamp Printing.

ISBN: 978–94–6473–718–9

Strategies for adiabatic state preparation of quantum many-body systems

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op donderdag 20 februari 2025, te 10.00 uur

door Dyon van Vreumingen
geboren te Schoonhoven

Promotiecommissie

Promotores:	prof. dr. C.J.M. Schoutens	Universiteit van Amsterdam
	prof. dr. L. Visscher	Vrije Universiteit Amsterdam
Overige leden:	prof. dr. A. Polkovnikov	Boston University
	dr. J. Tura Brugués	Universiteit Leiden
	prof. dr. J. van Wezel	Universiteit van Amsterdam
	prof. dr. R.M. de Wolf	Universiteit van Amsterdam
	dr. A. Mahomed Abbas	Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

aan hen zonder wier offers dit proefschrift er niet was geweest

List of publications

The following papers form the basis of this dissertation.

[vVS23] **Adiabatic ground state preparation of fermionic many-body systems from a two-body perspective.** Dyon van Vreumingen and Karel-jan Schoutens. *Physical Review A* **108**, 062603. Published 04 December 2023.
Chapter 3 is based on this publication.
The authors contributed equally to this work.

[SvVTP24] **Virtual mitigation of coherent non-adiabatic transitions by echo verification.** Benjamin F. Schiffer, Dyon van Vreumingen, Jordi Tura, and Stefano Polla. *Quantum* **8**, 1346. Published 14 May 2024.
Chapter 4 is based on this publication.
DvV provided the proof of the main theorem of this work (theorem 4.3.1).

[vV24] **Gate-based counterdiabatic driving with complexity guarantees.** Dyon van Vreumingen. *Physical Review A* **110**, 052419. Published 14 November 2024.
Chapter 5 is based on this publication.

Abstract

Quantum computers represent a relatively new and promising development in computing technology. These computers can (at least in theory) use the principles of quantum mechanics to perform certain calculations that would be very time-consuming for classical computers. One of the applications of quantum computers is modelling quantum many-body systems (systems consisting of more than two quantum particles), which describe interactions between particles on atomic and subatomic scales. Such systems are highly complex and challenging to model with classical computers due to the vast number of possible states and entanglement of the particles. In this dissertation, we examine the extent to which quantum computers can accelerate the modelling of quantum many-body systems compared to classical computers. This modelling has two facets: simulation of time evolution on one hand, and preparation of eigenstates (often ground states) on the other. It is known that quantum computers can achieve exponential speedup for simulating time evolution; thus, this dissertation focuses on the aspect of state preparation. Specifically, this dissertation presents research on adiabatic state preparation: a quantum algorithmic technique that uses the adiabatic principle from quantum mechanics to approximate eigenstates. We describe three new techniques that fall within this category and, in certain cases, offer advantages over standard methods.

In **chapter 3**, we consider cases of ground state preparation for fermionic many-body systems, where standard direct interpolation between the initial and final hamiltonian (e.g. from mean-field hamiltonian to full hamiltonian) is hindered by level crossings (also known as gap closures) due to discrete symmetries. As an alternative to direct interpolation, we propose adiabatic paths in a higher-dimensional space. These are introduced through a decomposition of the residual hamiltonian (the difference between final and initial hamiltonians), defined by the spectral decomposition of its two-particle projection. The procedure then involves interpolating a linear combination of terms from this decomposition. We show that this method breaks the appropriate symmetries in selected Fermi-Hubbard

models, thus enabling correct adiabatic passage. At the same time, we demonstrate that a naive approach in which each term is interpolated individually results in a computational complexity that scales unfavourably with system size, through an analysis of the maximum spectral norm of rotated hard-core boson number operators. This suggests that the evolution path should deviate minimally from the direct path, only to the extent necessary to break the specific symmetry.

In **chapter 4**, we present an adiabatic echo verification protocol which mitigates both coherent and incoherent errors, arising from non-adiabatic transitions and hardware noise, respectively. Carrying out a forward and backward adiabatic passage allows for an echo-verified measurement of any observable. In addition to mitigating hardware noise, our method uses positive-time dynamics only. Crucially, the estimator bias of the observable is reduced when compared to standard adiabatic preparation, achieving up to a quadratic improvement.

Chapter 5 discusses a general, fully gate-based quantum algorithm for counterdiabatic driving. The algorithm does not depend on heuristics as in previous variational methods, and exploits regularisation of the adiabatic gauge potential to suppress only the transitions from the eigenstate of interest. This allows for a rigorous quantum gate complexity upper bound in terms of the minimum gap Δ around this eigenstate. We find that, in the worst case, the algorithm can be run with at most $\tilde{O}(\Delta^{-(3+o(1))}\epsilon^{-(1+o(1))})$ quantum gates such that a target state fidelity of at least $1-\epsilon^2$ is achieved. In certain cases, the gap dependence can be improved to quadratic.

Samenvatting

Kwantumcomputers vormen een relatief nieuwe en veelbelovende ontwikkeling in de computertechnologie. Deze computers kunnen (althans vooralsnog in theorie) gebruikmaken van de principes van de kwantummechanica om bepaalde berekeningen uit te voeren die voor klassieke computers zeer tijdrovend zouden zijn. Een van de toepassingen van kwantumcomputers is het modelleren van kwantumveeldeeltjessystemen (systemen bestaande uit meer dan twee kwantumdeeltjes), die de interacties tussen deeltjes op atomaire en subatomaire schaal beschrijven. Dergelijke systemen zijn bijzonder complex en moeilijk te modelleren met klassieke computers, vanwege het enorme aantal mogelijke toestanden en de verstrengeling van de deeltjes. In dit proefschrift bestuderen we de vraag in hoeverre kwantumcomputers de modellering van kwantumveeldeeltjessystemen kunnen versnellen ten opzichte van klassieke computers. Deze modellering kent twee facetten: de simulatie van tijdsevolutie enerzijds, en de preparatie van eigentoestanden (meestal lage-energietoestanden) anderzijds. Van de simulatie van tijdsevolutie is bekend dat kwantumcomputers een exponentiele versnelling kunnen bewerkstelligen; dit proefschrift richt zich daarom op het facet van toestandspreparatie. Specifiek presenteert dit proefschrift onderzoek naar *adiabatische toestandspreparatie*: een kwantumalgoritmische techniek die het adiabatische beginsel uit de kwantummechanica inzet om eigentoestanden te benaderen. We beschrijven drie nieuwe technieken die in deze categorie passen en in zekere gevallen voordelen bieden ten opzichte van standaardmethoden.

In **hoofdstuk 3** beschouwen we gevallen van grondtoestandspreparatie voor fermionische veeldeeltjessystemen waarin standaard directe interpolatie tussen begin- en eindhamiltoniaan (bijvoorbeeld van gemiddeldveldhamiltoniaan naar totale hamiltoniaan) geplaagd wordt door kruisingen in energieniveaus (ook wel sluitingen van de energiekloof) als gevolg van discrete symmetrieën. Als alternatief voor directe interpolatie stellen we adiabatische paden in een hogerdimensionale ruimte voor. Deze worden in het leven geroepen door een ontbinding van de resthamiltoniaan (het verschil tussen eind- en beginhamiltoniaan) die

wordt gedefinieerd door de spectrale ontbinding van diens tweedeeltjesprojectie. De procedure bestaat dan uit het interpoleren van een lineaire combinatie van de termen uit deze ontbinding. We laten zien dat deze methode in geselecteerde Fermi-Hubbard-modellen de juiste symmetrieën breekt en zodoende correcte adiabatische passage mogelijk maakt. Tegelijkertijd tonen we aan dat naïeve toepassing waarbij elke term afzonderlijk wordt geïnterpoleerd leidt tot een rekencomplexiteit die ongunstig schaalt in de systeemgrootte, door middel van een bewijs voor de maximale spectraalnorm van geroteerde hardcore-bosongetaloperatoren. Dit suggereert dat het evolutiepad minimaal van het directe pad dient af te wijken, en slechts in zoverre dat precies de symmetrie wordt gebroken.

In **hoofdstuk 4** presenteren we een adiabatisch echoverificatieprotocol dat coherente afwijkingen door nonadiabatische overgangen en incoherente afwijkingen door ruis in hardware vermindert. Door een voorwaartse en achterwaartse adiabatische passage uit te voeren geeft dit protocol geverifieerde metingen van observabelen, waardoor de zuiverheid van de schatter voor de verwachtingswaarde van de observabele kwadratisch wordt verbeterd ten opzichte van conventionele adiabatische methoden. Hierbij is van belang dat onze methode slechts gebruik maakt van dynamica in positieve tijd, en zodoende geschikt is voor huidige kwantumhardware.

Hoofdstuk 5 bespreekt een algemeen kwantumalgoritme voor contradiabatische aandrijving dat volledig op kwantumpoorten is gebaseerd. Het algoritme is onafhankelijk van heuristieken zoals voorheen gehanteerd in variationele methoden voor het benaderen van de adiabatische ijkpotentiaal; daarbij maakt het gebruik van regulering van de ijkpotentiaal om slechts de overgangen vanuit de van belang zijnde eigentoestand te onderdrukken. Met deze aanpak kunnen we een rigoreuze bovengrens geven voor de rekencomplexiteit uitgedrukt in het aantal kwantumpoorten, als afhankelijkheid van de minimale energiekloof Δ rondom deze eigentoestand. We tonen aan dat het algoritme in het slechtste geval kan worden uitgevoerd met hoogstens $\tilde{O}(\Delta^{-(3+o(1))}\epsilon^{-(1+o(1))})$ kwantumpoorten, zodanig dat een getrouwheid van minstens $1-\epsilon^2$ wordt behaald ten opzichte van de doeltoestand. In bepaalde gevallen kan de afhankelijkheid van de energiekloof naar kwadratisch worden verbeterd.

Contents

Chapter 1

An invitation to the curious layperson

> *There is no general consensus as to what its fundamental principles are, how it should be taught, or what it really "means". Every competent physicist can "do" quantum mechanics, but the stories we tell ourselves about what we are doing are as various as the tales of Scheherazade, and almost as implausible. Niels Bohr said, "if you are not confused by quantum physics, then you haven't really understood it"; Richard Feynman remarked, "I think I can safely say that nobody understands quantum mechanics."*
>
> — David Griffiths

This is a thesis about quantum computing, the science of quantum computers. We are all familiar with computers as devices that automate tasks and perform complicated calculations for us. Quantum computers are also computers, but in a fundamentally different way: unlike classical computers which operate on bits – zeroes and ones –, quantum computers leverage the principles of quantum mechanics, the physics governing the behaviour of the smallest particles in the universe. The world of these particles is fascinating and strange, in ways that challenge our everyday understanding of reality – its peculiar properties allow quantum computers to process information in ways that classical computers cannot, opening up new possibilities for solving complex problems that were previously thought to be intractable. But before I can explain the intricacies of quantum computing to you, dear reader, we first need to go back in time; the story begins over a century ago, in the days when the foundations of quantum mechanics were being laid, and discoveries were made that fundamentally shaped our understanding of physics and computation as it stands today.

For a moment, imagine yourself standing in the shoes of Albert Einstein, at the dawn of the twentieth century. You live in a time when physics stands on

the foundations of newtonian mechanics, electrodynamics and thermodynamics. Newtonian mechanics describes the motion of objects like tennis balls and rockets in terms of quantities such as force, momentum and kinetic energy; electrodynamics deals with electric and magnetic fields, including waves like light and radio signals; and thermodynamics asks questions about the workings of heat, pressure and temperature. These are *deterministic* theories: if the precise state of a system is known at some instant in time, knowledge of the equations and their solutions suffices to predict the system's properties at any later time. It is in this setting, a world described by deterministic theories, that you discover the theory of relativity; with it come the famous equation $E = mc^2$ and the idea that nothing can travel at a speed higher than the speed of light c.[1] While being a revolutionary theory, prescribing completely new laws for how objects move (which, by the way, still "look like" familiar high-school newtonian mechanics in our daily lives, where objects move much slower than light), the theory of relativity is at heart still a completely deterministic theory.

Around the same time, physicists are struggling to understand nature at the smallest scale, and are trying to figure out what exactly atoms are. Experiments have shown that small particles, such as electrons, can also behave like waves: just like light, beams of electrons display diffraction and interference patterns. This phenomenon is known as the *particle-wave duality*. If particles were tiny tennis balls, such patterns would be impossible (see fig. 1.1). Furthermore, it has been observed that atoms can absorb and emit light – but only of particular wave lengths (fig. 1.2). Since a photon of wave length λ carries a definite energy $E = hc/\lambda$ (here, h is Planck's constant, a *very* small number), this has led to the consensus that energy at the smallest scales is *quantised*: unlike tennis balls, which can in principle carry any amount of energy[2], only a few particular energies $E_0, E_1, E_2, \ldots$ are allowed for systems of small particles. Neither the particle-wave duality nor quantisation of energy seem reconcilable with the previously accepted, newtonian understanding of particles; or, in simple words: the description of particles as tiny tennis balls seems to be outright wrong.

A few years later, during the interbellum, Erwin Schrödinger discovers a theory after years of hard work (building upon work by others including Heisenberg,

[1] Relativity predicts that weird things happen if it were possible for objects to travel at speeds larger than the speed of light. For example, if a train were to crash into you at a speed faster than light, you would first observe "being run over" and subsequently see the train moving backwards, away from you. In other words, cause (the train approaching you) would come after effect (you being run over). For this reason, physicists identify causality with the speed of light being a fundamental speed limit.

[2] Remember the formula for kinetic energy, $E = \frac{1}{2}mv^2$? According to Newton's theory, an object can in principle travel at any speed v and therefore carry any (nonnegative) amount of kinetic energy E. Einstein's theory of relativity restricts the possible speed to at most c (and changes the formula for kinetic energy to account for this), but any speed within the *continuous interval* between 0 and c is fine, and in the end, this still allows any (nonnegative) value of kinetic energy.

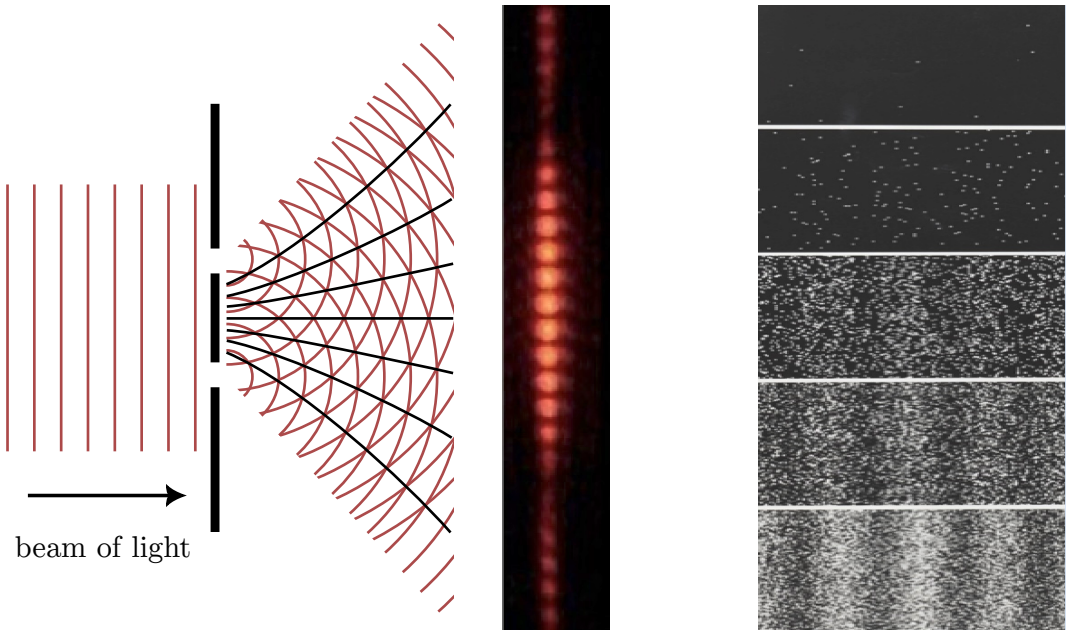


Figure 1.1. *Left*: the double-slit experiment. In 1801, Thomas Young conducted the displayed experiment to show that light is a wave. A planar light wave is directed at a plate with two apertures, and the light passing through the apertures illuminates a screen positioned behind the plate. Because of diffraction, light coming from the apertures behaves as if there were a point source at each aperture. The light from these two sources interferes and creates a pattern of light and dark bands on the screen. The black lines indicate the paths of constructive interference. *Right*: if the same experiment is carried out by shooting electrons at the plate instead of light, a similar interference pattern is observed on the screen. This shows that electrons behave like waves as well: if they didn't, we would observe the electrons landing only in two spots, corresponding to the two apertures. [TEM$^+$89]

Figure 1.2. The emission spectrum of helium [Pog13, KRR$^+$23]. When a helium atom releases energy in the form of light, it can only emit the few specific wave lengths (colours) present in this spectrum. In other words, it can only emit a few specific quantities ("quanta") of energy. The converse is also true: the atom can gain energy by absorbing a photon, but only of these specific wave lengths. The physicists' explanation for this phenomenon is that the energies of the atom are quantised.

De Broglie, Born and Bohr) that manages to explain these phenomena by giving a mathematical description of particle waves. This forms the basis of what is now known as *quantum mechanics*. Quantum mechanics, however, is a theory that is inherently *nondeterministic*, describing the behaviour of small particles as naturally *random*: even if we exactly know the state of a particle at some point, we can never be certain what properties we will measure if we observe the particle at a later time. But it predicts even stranger things: among others, it says that particles can find themselves in multiple states at the same time and, most disturbingly, predicts that the behaviour of a system of particles is dependent on the observations *you*, dear reader, make to it...

Would you accept such a seemingly absurd theory?

Einstein certainly didn't. While he deeply admired the efforts of Schrödinger and the others who had contributed to the development of the theory, he considered its intrinsic nondeterminism nonsensical. In his view, it was an incomplete picture of reality, and there had to be something else we simply hadn't yet discovered – a "hidden variable" of sorts. In a correspondence with his good friend and physicist Max Born, he rejects the idea of "God playing dice with the universe":

> *Die Quantenmechanik ist sehr achtunggebietend. Aber eine innere Stimme sagt mir, daß das noch nicht der wahre Jakob ist. Die Theorie liefert viel, aber dem Geheimnis des Alten bringt sie uns kaum näher. Jedenfalls bin ich überzeugt, daß Der [Gott] nicht würfelt.*
>
> — Albert Einstein

Unfortunately for Einstein, more and more experimental evidence came out in the years after, which supported Schrödinger's theory despite its allegedly preposterous nature. On top of that, in the sixties, John Bell proved that any theory based on the existence of a "hidden variable" is incompatible with quantum mechanics [Bel64], and that Schrödinger's theory had to be the entire story! At this point, Schrödinger's formulation of quantum mechanics became widely accepted, and

established its foothold as a complete theory of the smallest particles.[3]

Did I manage to keep your attention so far? Good. I will now explain the basics of quantum mechanics in more precise, but also more technical, terms. If you are not scared by a little bit of mathematics and abstract reasoning, I encourage you to read on; otherwise, you can skip to the next section.

1.1 A brief introduction to quantum mechanics

In classical mechanics, objects like tennis balls are described in terms of *classical variables* like position, velocity and kinetic energy. For example, an object travelling in a straight line can be modelled by the distance $x(t)$ it has travelled since we started our observation, which is a function of time t. If the object travels at a constant velocity v, we have the simple equation $x(t) = vt$. Like all classical variables, $x(t)$ is a deterministic quantity: if we know $x(t)$ at some time t, we can always predict its value at a later time by using the correct equation. Quantum mechanics, on the other hand, models objects in a completely different way, using *wave functions*. A wave function is simply a function $\Psi(x, t)$ that depends on two variables, position x and time t. In loose terms, if we want to describe particles as waves, like Schrödinger, then $\Psi(x, t)$ tells us what that wave looks like at point x and at moment t. In contrast to classical mechanics, where we know the position of the particle exactly, the wave function models a probability distribution: if a particle is described by $\Psi(x, t)$, and we measure its position, then the probability of finding it between two points x_1 and x_2 at time t is given by

$$P(\text{particle is between } x_1 \text{ and } x_2 \text{ at time } t) = \int_{x_1}^{x_2} |\Psi(x,t)|^2 \, \mathrm{d}x. \tag{1.1}$$

See figure 1.3. This is what I meant with inherent nondeterminism of quantum mechanics: the position at which you will find the particle is random, and you can not be certain where it will end up – all you know is where it is *likely* to appear, namely at those positions where the magnitude of $\Psi(x, t)$ is large. This does not mean, however, that a particle must always be all over the place. Imagine a wave function that is peaked at a specific position x, and (almost) zero elsewhere: since its magnitude is large only in the vicinity of x, a measurement will very likely reveal that the particle is near x. See figure 1.4. Now imagine such a peaked function, centred at the same point x, but with a very, very, *very* narrow peak – one could say, with a width so small it approaches zero –; where would you

[3] By a complete theory, I mean a theory that ignores relativity. Quantum mechanics has been adapted to incorporate *special* relativity, which is a smaller part of Einstein's relativity theory, in the form of quantum field theory. To this day however, quantum mechanics is still at odds with Einstein's *general* relativity theory of gravity and space-time curvature. But these are completely different topics, and beyond the scope of this thesis.

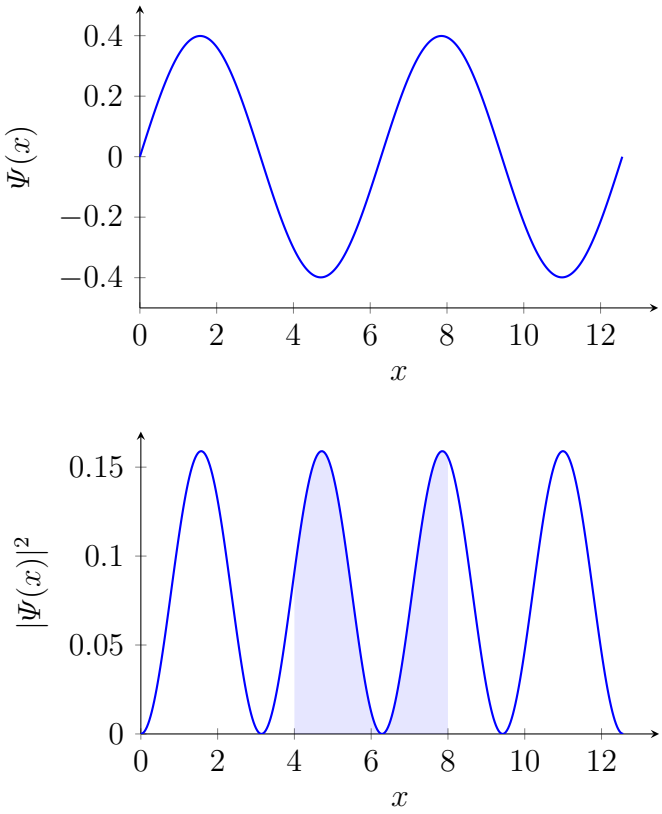


Figure 1.3. *Top*: at some time t, a particle is described by the wave function $\Psi(x)$. *Bottom*: the probability distribution of the particle's position is given by $|\Psi(x)|^2$. The area of the shaded region equals the probability of finding the particle between $x = 4$ and $x = 8$ upon measurement.

measure the particle? Yes, exactly, at point x, because the magnitude of the function is zero anywhere else. Such wave functions, which guarantee you will find the particle at one single position with 100% probability, are special; we call them *position eigenfunctions*.

How about energy? In classical mechanics, energy can be computed from variables like position, velocity and so on, using a collection of well-known formulas. But if the position of a particle is nondeterministic, shouldn't its energy be as well? Yes, that's right! However, computing a particle's energy in the quantum setting is a bit more complicated than just applying the energy formulas of classical mechanics to the measurement outcomes of position and related variables, and understanding how it really works requires a full course on quantum mechanics. What I'll tell you is this: just like position, the energy of a particle (or for that matter, a collection of particles) can be measured; and just like position eigenfunctions, there exist special *energy eigenfunctions* which guarantee you as an observer to find one energy value with 100% probability. These energy eigenfunctions are generally different from position eigenfunctions,

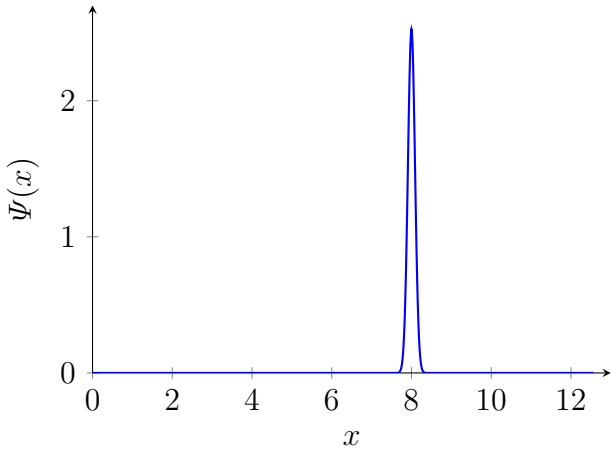


Figure 1.4. If the wave function is sharply peaked, we will only find the particle in the vicinity of the peak ($x = 8$ in this case). The narrower the peak, the smaller the region around $x = 8$ where we can expect to find the particle.

and there exist only an integer number of them, because energy is quantised. (Remember quantisation of energy? It says that quantum systems can only carry particular energies $E_0, E_1, \ldots$ and nothing in between.) We usually label them $\Psi_0(x,t), \Psi_1(x,t), \ldots$, corresponding to their respective energy value; that is, if a particle is in a state described by Ψ_0, you will measure energy E_0, and so on. Most of the time, however, particles are not in a state described by a single energy eigenstate, but in a "mixture" of them, called a *superposition*, and expressed as a weighted sum over multiple energy eigenfunctions:

$$\Psi(x,t) = c_0(t)\Psi_0(x,t) + c_1(t)\Psi_1(x,t) + c_2(t)\Psi_2(x,t) + \cdots \tag{1.2}$$

Here, the coefficients $c_0, c_1, c_2, \ldots$ are simply numbers (which can change with time). If you measure the energy of a particle in this superposition, you will get outcome E_0 with probability $|c_0|^2$, outcome E_1 with probability $|c_1|^2$ and so on.

I haven't said anything about the mathematical form of energy eigenfunctions – but the truth is that calculating them for a given system is actually a very complicated matter. At the same time, knowledge about these energy eigenfunctions (especially the one corresponding to the lowest energy – more on that later) turns out to be crucial for understanding the physical properties of quantum mechanical systems. It is for this reason that thousands of man hours and megawatt hours of computing power are being spent on calculating these functions as we speak. Determining energy eigenfunctions is also the main objective of the work in this thesis: chapters 3, 4 and 5 all present techniques for attacking this problem.

Of course, there are many things to know about a system other than just position and energy; as a general rule in quantum mechanics, anything that can be observed can be measured, and comes with its own eigenfunctions. In the same vein, a system in a state that is a superposition of such eigenfunctions will

display randomness in the measurement outcomes of the property in question.

Alright, this sounds cool and all, wave functions and randomness and stuff – but does it actually make any sense? How should we go about interpreting these things? A physicist's job is to *understand* the world, isn't it? Of course, as a reader, you could simply accept my claim that "experiments have proven quantum mechanics correct" and be done with it. And I wouldn't blame you: there are plenty of physicists who prefer a "shut up and calculate" mentality over questioning the philosophical nature of quantum mechanics, and regard it as nothing more than an accurate model for making predictions of experimental outcomes. But this completely ignores the plethora of conceptual intricacies that comes with quantum mechanics, and any introduction to this beautiful theory would be incomplete without a mention of it.

Say you measured the position of a particle and found the point x. What exactly does that tell you about the reality of the particle? You might be tempted to understand the outcome as an indication that the particle really *was* at x when we measured it, and that quantum mechanics just didn't know about it. But this is the same as saying that quantum mechanics is an incomplete and deterministic theory, which Bell showed cannot be the case. Therefore we are forced to interpret the nature of the smallest particles in a radically different way. As you might have expected, the interpretation of quantum mechanics remains, up to this day, a question without a definite answer. I will present the two most prevalent interpretations here; it is up to you, dear reader, to pick the one that suits you best. And if neither interpretation convinces you, I warmly invite you to do your own research – there are plenty of other philosophies to discover.

The most influential interpretation builds on the nondeterministic ideas discussed above and is known as the *Copenhagen interpretation*. The main idea in this interpretation is that measurement ontologically disturbs the state of the system: before a measurement, the particle did not have a position at all, and its position was created by the measurement. Perhaps the best way to illustrate this point is the following: quantum mechanics postulates that if you measure the position of a particle, you find value x, and you *immediately* measure its position again, you should find the same value. Makes sense, right? After all, the particle had no time to move elsewhere, because the second measurement was immediate. But it has a profound effect: in order to guarantee that the second measurement will yield the same outcome, the wave function that correctly describes the particle after the first measurement must be the position eigenfunction peaked at x, the position where you found it. In the Copenhagen interpretation, we say that the measurement *altered* the wave function. (Physicists refer to such an event as a *wavefunction collapse*.) In the same way, we say that after measuring energy, the wave function collapses onto one of the energy eigenfunctions $\Psi_0, \Psi_1, \ldots$. While this interpretation is the most native to the mathematical formulation of quantum mechanics, there is something very strange about the role of measurement here. After all, what *is* a measurement really? One could define it as the moment your

measurement apparatus interacts with the system under scrutiny. But why is precisely your measurement device capable of making the wave function collapse? In the end, a measurement device is made up of particles, which are quantum mechanical in nature. Does this mean that the interaction between quantum particles causes wave functions to collapse? This seems highly unlikely, as this would mean that the wave function would collapse all the time, and quantum mechanics would be deterministic. Perhaps, then, measurement and collapse do not happen in the device itself, but only at the moment when you become consciously aware of the outcome. Does this mean that consciousness itself has the power to influence nature? Or is the reality which we consciously observe after measurement simply a projection in which the wave function just so happens to "look" collapsed? Clearly, the discussion turns highly speculative at this point, and I'm cutting it short here – if this got you interested, dear reader, there is a big rabbit hole waiting for you to be explored.

The *many-worlds interpretation*, being the second-most well known school of thought, attempts to fix the issues of the Copenhagen interpretation by viewing reality as a multitude of parallel universes. The idea is as follows: say a system is described by a superposition of eigenfunctions $c_A\Psi_A + c_B\Psi_B$, and a measurement is made. In the many-worlds view, a measurement *splits* the state of the universe into multiple parallel universes, one for each possible outcome. In one world, the observer has found outcome A, and in the other, outcome B. But the observers in the different worlds can never know about each other, since what they can observe consists only of what can be measured – namely, the physical properties of the quantum system. The many-worlds interpretation circumvents the problem of the wave function collapse by postulating that in world A, the system was in a sense *predetermined* to be in state Ψ_A, while in world B, it was predetermined to be in Ψ_B.[4] In this way, the global wave function, combined over all worlds, remains the same, and there is no collapse at all – it's just that observer A and observer B find themselves in different worlds, with different "views" of the wave function.

But what about probabilities? Quantum theory says that the probability of finding outcome A equals $|c_A|^2$ and that of outcome B is $|c_B|^2$. But if outcome A was predetermined for world A, can we even talk about probability in that world? Evidently, we need to find a way to describe a "probability of being in world X" in order to make sense of measurement statistics in this interpretation. Unfortunately, this is where efforts by physicists, mathematicians and philosophers have fallen short so far, and satisfactory answers to these questions have remained out of reach. This is the main issue used to argue against the many-worlds interpretation by its opponents.

[4]This may smell like a violation of nondeterminism to you – but it is precisely the possibility of multiple outcomes, even if predetermined for each world, that keeps the mathematical formulation of quantum mechanics intact and ensures it agrees with John Bell's no-hidden-variable theorem.

In any case, there is plenty to philosophise about the true nature of quantum mechanics, and this was my reason for quoting David Griffiths at the beginning of the chapter. If this very brief introduction managed to spark your interest in quantum mechanics, I strongly encourage you to read his book, to which I have included a reference at the end of the chapter.

1.2 Quantum computing

1.2.1 Simulating quantum with quantum

Fast forward a few decades after Schrödinger's discovery. As computers entered the scene and gradually became commonplace, numerical simulations started playing an important role in many scientific fields. For physicists, quantum mechanical systems became an appealing target of simulation, and numerical results yielded new insights in a range of research areas, including chemistry and superconductance.[5] But scientists soon found out that accurately simulating the behaviour of quantum systems takes a *lot* of time on a computer. Put simply, this is because the amount of numbers you have to keep track of to simulate a quantum system of multiple particles generally grows exponentially in the number of particles.[6] In 1981, this issue was addressed by physicist Richard Feynman in a conference talk [Fey82], where he came up with the following solution:[7]

> *Nature isn't classical, dammit, and if you want to make a simulation of nature, you'd better make it quantum mechanical – and by golly it's a wonderful problem, because it doesn't look so easy.*
>
> — Richard Feynman

In other words, if quantum physics is so hard to simulate with a standard computer, we should do it with a computer that is itself quantum mechanical in nature.

[5] Superconductance is a phenomenon, discovered by Dutch physicist Heike Kamerlingh Onnes in 1911, where certain materials start conducting electrical current with zero resistance when cooled down to (almost) zero Kelvin ($-273.15\,^{\circ}\mathrm{C}$).

[6] For those readers not familiar with exponential growth, imagine picking a password. If you use only lowercase letters, you have 26 choices for the first character. Then, another 26 choices for the second character, so $26 \times 26 = 676$ possibilities for a two-character password; $26 \times 26 \times 26 = 17576$ for a three-letter password and so on. As you use more characters, the number total number of possible passwords becomes stupidly big. The difficulty of simulating quantum mechanics also has to do with an exploding number of combinations.

[7] Admittedly, this phrase is pretty overused and by now cliché in the scientific quantum computing community (it appears in the introduction of a significant percentage of papers on quantum computation), but it so clearly and concisely conveys the point that I couldn't resist quoting it here as well.

But what exactly is a quantum mechanical computer? I think the best way to think of it is as a *controllable* quantum system. Say you a have single atom, confined to a small region of space. In the lab, electromagnetic fields can be used to isolate atoms and keep them "trapped" in a stable position. Assume the atom is in the state with the lowest energy E_0, and call this the "zero" state. Since it is trapped in a specific position, we can raise its energy by pointing a beam of laser light at it. Say this raises the energy of the atom to E_1, and call this new state of the atom the "one" state. What we've done is build a controllable quantum system that represents a *bit*: something that can be zero or one, and whose state can be controlled by us. In that, it resembles a classical computer, which also works with bits that we can control. But this bit, being represented by an atom, is quantum mechanical in nature, and behaves according to the laws of quantum mechanics; we thus call it a quantum bit, or *qubit* for short.

Of course, classical computers work with more than just one bit; similarly, a useful quantum computer should consist of multiple qubits. Furthermore, the qubits should be able to *interact*, which is to say that it should be possible to perform a computation that is not only dependent on the states of each qubit individually, but also on the combined states of pairs, triples, etc. of qubits. Consider for example multiplying numbers on a classical computer. Say you have two bits, bit A and bit B, and wish to compute their product $\mathrm{A} \times \mathrm{B}$ which should subsequently be stored in bit A. If you have access only to operations that do not allow bits A and B to communicate, then computing this product is impossible. In the same sense, a quantum computer whose qubits cannot communicate is heavily restricted in the kinds of computation it can perform.[8] It turns out that, if you set up a controllable quantum system in the right way, with all the quantum bits happily connected, then such a machine can in principle *efficiently* simulate any other quantum system consisting of a number of particles roughly proportional to the number of qubits. This is called *universal quantum computation*. Here, efficiently means that the computation time does not grow exponentially in the number of particles you're trying to simulate (in many cases, it can be made roughly linear).

The research presented in this disstertation is mainly about the simulation of quantum mechanics, but the field of quantum computing is broader than that. If you are short on time, you can now skip to section 1.3, where I explain the contents of the dissertation; but if you're interested in other things quantum computers can do (section 1.2.2), or where we currently stand with regards to actually realising these quantum computers (section 1.2.3), please read on.

[8] In fact, one can mathematically prove that all these "communication-free" quantum computations can be quickly performed by a classical computer as well, so it would render the quantum computer useless.

1.2.2 Going beyond quantum simulation

About a decade after Feynman's lecture, mathematicians started picking up the topic of quantum computation. They wondered what else a universal quantum computer could be capable of in theory, besides simulating quantum systems. And remarkably, they managed to prove that there exist other, completely unrelated, tasks at which quantum computers are better than classical computers. In other words, they showed the existence of *quantum algorithms* – programs that can be run on a quantum computer – which complete these tasks faster than a classical computer. I will talk about the two most surprising results here.

In the nineties, Lov Grover devised a quantum algorithm for unstructured search. Suppose you have a pile of N paper slips, with a name written on each slip, and you are trying to find one specific name in the pile. Is there a clever way to quickly find the name you're looking for? Unfortunately for you, it turns out the best thing you can do is check all slips until you've found the right name. If you're unlucky, the right slip is the last one you check, and you need to check N slips before finding the right one. On average, you will find the slip with $N/2$ tries. Computers face the same issue when searching for a specific record in a long, unsorted list of data. Grover, on the other hand, showed that a quantum computer can find the right name with a number of checks that is only about $\sqrt{N} \times C$ (where C is a relatively small constant number), provided that the names can somehow be encoded on the quantum computer. Of course, this doesn't make much of a difference when you have only a few names to check, but if you have a billion names to check, then $\sqrt{N} \times C$ becomes a much smaller number than $N/2$.

Around the same time, another quantum algorithm was invented by Peter Shor, for a completely different task: that of prime factorisation. The problem of prime factorisation is the following: given an integer number, how can we write it as a product of prime numbers?[9] If the number is small, this is not so difficult: for example, one can quickly find out that 15 can be factorised as 3×5. But what about a large number, say 321800635057? Can you give me its prime factors? I'm sure you could, if you really cared – but it would take you a *long* time. In fact, the problem is so difficult that the best known classical algorithm needs an amount of calculation time that grows exponentially[10] in the number of digits of the input number. Shor's algorithm however, believe it or not, needs only a number of steps that grows with the third power of the number of digits![11].

You might be wondering why someone without a deep fascination with prime

[9] A prime number is an integer number larger than 1 that is only divisible by 1 and by itself. Every integer can be written as some product of prime numbers.

[10] Strictly speaking I should say *subexponentially*: slightly better than exponentially, but still growing very fast.

[11] If you're not convinced this makes a big difference, compare the functions x^3 and 3^x. The number I gave you has 12 digits; if we take $x = 12$, which of the two is bigger? How about a larger number with say $x = 50$ digits?

numbers would care about the problem of prime factorisation. But as a matter of fact, Shor's discovery has big implications for a lot of people worldwide, even though quantum computers themselves are still very much in development. The issue is that much of our digital communication systems – think websites and financial infrastructure – is currently being kept secure against cyber attacks by something called RSA encryption. In simple terms, these communication systems continuously send out information that only the receiving end should be able to read, and must not be intercepted by a third party, i.e. a hacker. To keep the information out of the hands of the wrong parties, it is *encrypted* into a code, and only the party with the correct key – the receiver – can decode the message. RSA is a method for calculating such codes and keys. However, it was designed based on the assumption that prime factorisation is a difficult problem, since a party who does not own the correct key to decrypt a message can *calculate* it by solving a particular prime factorisation problem. This was fine before quantum computing appeared, as it would take a computer millions of years, if not more, to solve the factorisation problem and calculate correct keys (2048-bit and 4096-bit RSA are currently standard, which correspond to 617 and 1234 decimal digits respectively; the largest factorised number as of today is 250 digits long [BGG$^+$20]). But if anybody had a full-scale quantum computer, they would be able to intercept and decode the messages, and this makes quantum computers a big threat – at least in theory, for now – to our digital security. Fortunately, since Peter Shor's discovery, many proposals have been put forward for *post-quantum cryptography* [BBGP16, KP20], so we will likely be able to change our security protocols to something that can withstand quantum attacks; nevertheless, inplementing this migration in all of our digital infrastructure is expected to come at a large cost [JMM$^+$22].

So why exactly are quantum computers capable of doing things that our current computers cannot? Well, a good answer to this question is rather technical, and probably requires a degree in physics or mathematics, but I'll do my best to give a simple and intuitive explanation. At the beginning of this chapter, I mentioned that, according to Schrödinger, quantum mechanical systems can be in multiple states at the same time. (If you read the last section, you know that this is called a superposition.) A quantum bit, therefore, can be in both a 0 state and a 1 state at the same time. When we look at two qubits, they together can be in the states 00, 01, 10 and 11 simultaneously. Three qubits will form a superposition of the states 000, 001, 010, 011, 100, 101, 110 and 111. Do you see where this is going? As the number of quantum bits grows, the number of states which can be simultaneously occupied grows exponentially – to be precise, with n qubits, there are 2^n possible states. (Of course, the number of possible bit sequences is the same for a classical computer, but the difference is that a classical computer can only be in one of those states, while a quantum computer can be in a combination of all of those states at the same time.) Now the beauty of quantum computing comes from the fact that, when we apply an operation on

the qubits (e.g., "flip the third qubit from 0 to 1 and vice versa"), this operation is applied on *all the states* simultaneously. In this way, a quantum computer can seemingly perform a massive amount of calculations in parallel!

However, there is a catch. Like with any calculation, we will want to read out the result at some point. Since humans (or classical computers, if the result needs to be processed further) cannot deal with superpositions of states, we need to extract a single bit sequence as output. In other words, we need to make a *measurement* to the quantum computer. Schrödinger's theory tells you that such a measurement will indeed give you a single state, but which one you get is *random*. (Also, this measurement changes the configuration of the quantum computer to the single state you just obtained, destroying the superposition. In the language of the Copenhagen interpretation from the last section, the state collapses onto the bit sequence that was found after measurement.[12]) For this reason, quantum computing scientists actually prefer not to talk about parallellism in quantum computation, since there is no way to obtain all possible outputs from the computation. Instead, quantum computing is generally viewed as a game of probabilities. When I say that the output is random, that means that there is a probability associated with each outcome. The art of quantum computing then lies in tweaking these probabilities such that the outcome most likely to be found after measurement is the solution to a problem we want to solve.[13] Just like a classical computer manipulates bits using logical operations, a quantum computer manipulates the probabilities of a quantum system using quantum operations. A sequence of properly chosen quantum operations, followed by an appropriate measurement, then constitutes a quantum algorithm.

When it comes to finding data from an unordered list, or finding the prime factors of a number, it just so turns out to be that the number of quantum operations needed for achieving an acceptable probability of getting the right answer is less than the number of logical operations a classical computer needs in order to find the answer. In other words, for these cases, tweaking probabilities with quantum operations is simply more efficient than calculating the answer deterministically.[14] Of course, unstructured search and prime factorisation are

[12]Some say that the many-worlds interpretation of quantum mechanics (section 1.1) provides a quick quantum algorithm for solving any problem. The algorithm goes as follows: prepare a superposition over all candidate solutions and make a measurement; if the answer is wrong, kill yourself. According to the many-worlds interpretation, the world in which you stay alive is the one in which the problem is solved... (Please don't try this at home.)

[13]Yes, that means that sometimes, a quantum computer will give you the wrong answer! But if we repeat the calculation a number of times and keep the output that we got most often, the probability of obtaining the correct answer gets very close to 100%, provided the correct answer was the most probable one.

[14]A computer scientist would, after reading the last two paragraphs, likely complain that the things I wrote apply just as well to *randomised computation* (a classical computation that takes random steps at one or more points). This is partially true, and in fact, there are several quantum algorithms which were initially claimed to provide a quantum advantage

only two of the many computational problems that can be solved by a computer; since Grover's and Shor's discoveries, quantum algorithms have been designed for a wide range of other problems (see ref. Jor24 for a comprehensive list), and many of them make use of elements from Grover's and Shor's algorithms to achieve similar advantages compared to classical algorithms.

On the other hand, there is also a category of computational problems for which it is not so clear whether a quantum computer can provide an advantage or not. Prime examples are optimisation problems, such as "what is the quickest route for a grocery delivery van?" or "what is the most aerodynamical shape of an airplane wing that can be manufactured cheaply?". Optimisation problems also play a key role in artificial intelligence. Typically, the number of candidate solutions to these problems is unfathomably large, so for most practical applications, approximately optimal solutions suffice. These approximate solutions can be found by algorithms known as *heuristics* which, in very blunt terms, are sophisticated guessing algorithms that just turn out to work well in practice. The theory behind these heuristics and why they work so well is poorly understood, and this makes it difficult to design quantum algorithms which might do better. Numerous attempts have been made to adapt currently known "guessing techniques" to quantum computers, but none of these have yielded a convincing quantum advantage so far.

1.2.3 Quantum computers today

If quantum computers can do numerous things faster than our current computers, does that mean that quantum computers will replace our desktops and laptops in the near future? Probably not. Most importantly, building a controllable quantum system with a few qubits is one thing – quite a few research labs have managed to build such a setup by now – but scaling it up to millions, billions or even more qubits is a completely different story. (Compare it with your own computer: if you have a 1TB hard drive, that's roughly eight trillion bits.) This is because quantum states are incredibly fragile, i.e. highly sensitive to interference from the surrounding environment. That is, tiny disturbances like thermal fluctuation, radiation or sound can be enough to perturb or completely destroy a delicately constructed superposition, which may have been needed for a particular quantum computation. This phenomenon is known as *quantum decoherence*. To combat this, current quantum computing setups cool down their qubits to almost zero degrees Kelvin; still, quantum computations can typically be maintained for only a few hundred nanoseconds before breaking down. On top of that, the quantum operations that current setups can implement are not perfect,

but were later shown to have a fast classical randomised equivalent. Nonetheless, quantum probabilities are subtly different, and according to widely accepted theory, quantum computers are more powerful than classical ones; this is exemplified by Shor's algorithm, for which no fast randomised algorithm is known.

causing errors to accumulate over the course of a computation. And all this gets worse as the number of qubits is increased. For this reason, we are now said to live in the era of "noisy intermediate-scale quantum" (NISQ) computers. These NISQ devices have been succesful at demonstrating that quantum computation is in principle possible, as long as the computation is short and errors are tolerable. However, NISQ devices are not ready to run the famous quantum algorithms invented by Grover and Shor or perform useful quantum simulations, for which long, high-precision calculations are required. On the bright side, mathematicians have figured out a way to deal with accumulating errors in quantum computations. A quantum computer can be made much more robust by representing what is on paper a single qubit (a "logical qubit") with a large number of physical qubits (like the trapped atoms from the beginning of this section), instead of a single physical qubit. Then if one of the physical qubits "crashes", it can be "fixed" using information from the other physical qubits belonging to the same logical qubits. This scheme is known as *error correction*, and is key to achieving what is called *fault-tolerant* quantum computation, which would enable running useful quantum simulations and algorithms such as Grover's and Shor's. However, implementing a stable quantum computer in this way requires many physical qubits. Estimates have been put at 1000–10000 physical qubits per logical qubit [CTV17], depending on the error rate, though experimental results suggest that this figure can be lowered to 24 physical qubits per logical qubit [BCG+24]. In any case, no universal quantum computer exists with more than ten thousand physical qubits as of this writing, meaning no practically useful fault-tolerant quantum algorithm can be run just yet. In practice, developing and operating quantum computers is enormously expensive, given that all technology must work at very high precision and very low temperatures. For all these reasons, quantum computers are expected to remain special-purpose devices and will likely not replace our computers any time soon. After all, our current computers are already

Company	**Year**	**Logical qubits milestone**
QuEra	2025	30
	2026	100
IonQ	2025	64
	2026	256
	2027	384
	2028	1024
IBM	2029	200
	2033+	2000

Table 1.1. Quantum computing roadmaps published by QuEra [QuE24], IonQ [Ion24] and IBM [IBM24]. Each have set milestones the number of logical (i.e. error-corrected) qubits to be achieved in the coming years.

very good at plenty of tasks (such as running your operating system), at a fraction of the cost compared to a quantum computer. Nonetheless, rapid progress is being made as we speak, and leading contenders have put out roadmaps which set milestones in terms of the number of logical qubits in the upcoming years. See table 1.1. Besides the players in this table, other parties to watch are Google, PsiQuantum, Xanadu, Quantinuum, IonQ and Atom Computing.

1.3 This doctoral thesis

This dissertation focusses the topic of simulating quantum systems, as originally imagined by Feynman. In particular, I have conducted research into the simulation of *quantum many-body systems*: collections of many quantum mechanical particles that interact with each other. Part of the dissertation (mainly chapter 3) pays special attention to the simulation of electrons moving around atomic nuclei. This is a rather general setting, since (almost) all matter around us is made up of atomic nuclei surrounded by electrons. Matter made from atoms can take on many different forms; this thesis targets materials such as metals and crystals, which consist of atoms or ions arranged in a lattice, and molecules. Given the complexity of these systems, fully detailed simulations are infeasible, and we use simplified abstractions to model them. Nonetheless, understanding these abstract models gives valuable insights into the behaviour of molecules and materials, as well as physical processes like chemical reactions and electrical conductance.

Before I continue explaining the research in this thesis, I should specify what exactly is meant with quantum simulation. Strictly speaking, quantum simulation deals with the question: "if I know the state of a quantum system now, what will it be in the future?" The process by which a quantum system changes from one state to another state at a later time is called *time evolution*, and the aim of quantum simulation is to calculate this time evolution for any given time. This is exactly what a universal quantum computer is good at, and where it (at least theoretically) has a significant advantage over a classical computer. If we can set up our quantum computer in such a way that the state of its quantum bits is a good representation of the quantum system we wish to simulate, then we can naturally let the quantum computer evolve according to its own time evolution, so that its final state will be a good representation of the studied quantum system at a later time. As mentioned before, the time our quantum computer needs to simulate another system's time evolution is typically not much more than the time the studied quantum system itself would need for its time evolution.

To make simulated time evolution work, we need two ingredients. First of all, we need to make sure the quantum computer simulates the correct quantum physics. Of course, you could set up a handful of qubits to simulate, say, ten electrons. But whether these electrons are part of a molecule or of a crystal makes a big difference, so when we want to simulate a molecule with ten electrons, we

should instruct the quantum computer to simulate the quantum physics of this molecule, and not of some other molecule or a crystal. The mathematical object that encodes the correct quantum physics of a system is called the *hamiltonian*. The good news is that it is relatively straightforward to feed a hamiltonian into a quantum computer, and there exist efficient protocols for doing so. Secondly, any quantum simulation needs a *starting state* from which the time evolution is to be calculated. This starting state should have some physical meaning, in that it should be a reasonably accurate description of a state in which the system could be found in nature. After all, if a quantum computer carries out a realistic simulation of a time evolution, starting from an unrealistic state, then the end result will also be an unrealistic state. In nature, systems are often found in *ground states*: states in which the energy of the system is minimal. Compare, for example, a river: water pouring down from the clouds will (eventually) always flow to the lowest point. This is because it takes energy to lift up anything that has mass, and energy is released when it falls down. Thus, a lake is essentially a body of water sitting at its lowest-energy point. Similarly, molecules at rest, as well as materials cooled down to zero Kelvin (such as in the case of superconductance), are usually in their lowest-energy state. (In the words of section 1.1, this state is an energy eigenfunction corresponding to the lowest possible energy.)

The bad news, however, is that calculating the lowest-energy state of a many-body system is a very difficult problem. As far as we know, no classical computer can do it in a time that scales better than exponentially in the number of particles. But the consensus among scientists is that it is so difficult that even a quantum computer will generally need an exponential amount of time, despite its advantages over a classical computer. Nevertheless, it may turn out that the base of this exponential[15] is smaller for practical purposes when we use a quantum computer. This is the main motivation for the currently big research field of *quantum ground state preparation*, which is concerned with developing quantum computational methods for calculating ground states. This dissertation is a contribution to that field.

1.3.1 Adiabatic quantum computing

The research conducted for this dissertation is aimed at a niche within quantum ground state preparation known as *adiabatic state preparation*.[16] These methods hinge on a concept in physics that is known as the *adiabatic principle*. It is quite likely that you have encountered the adiabatic principle in your daily life, without knowing it had that name. Imagine holding a string with a metal ball attached

[15] In the functions 2^x and 4^x, the numbers 2 and 4 are the bases of the exponential, respectively. For large x, 4^x is much larger than 2^x, although both scale exponentially in x.

[16] For those who will continue reading after this chapter: some of the methods presented in this dissertation, especially those in chapters 4 and 5, are also applicable to higher-energy states.

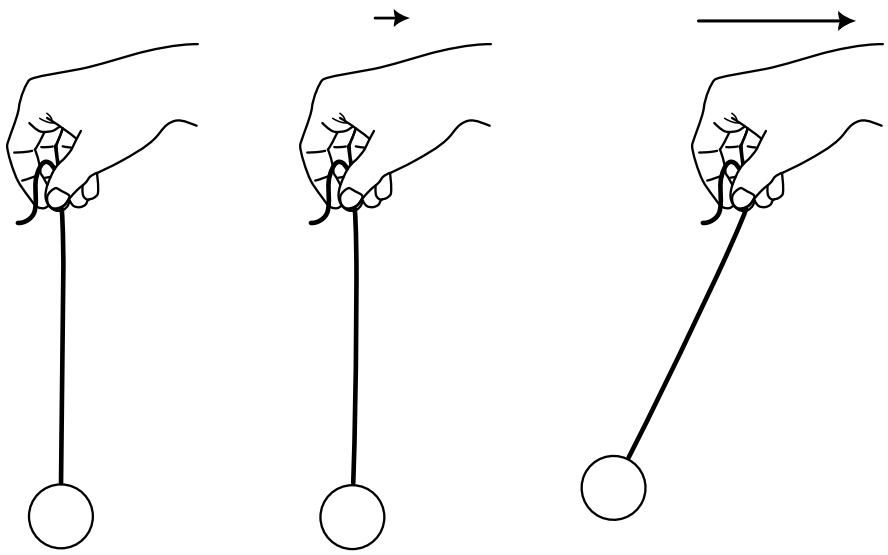


Figure 1.5. *Left*: holding a ball on a string. The system is at rest. *Middle*: adiabatic motion. The system is slowly moved out of position, without disturbing the ball and making it swing. *Right*: diabatic motion. When the hand is moved fast, the ball will start swinging.

at the bottom, hanging there at rest without swinging to the side. Say you would want to rotate your body ninety degrees, still holding the string, without making the ball swing. How would you do it? Right, you would turn very slowly, in order to keep the ball at rest; if you turned very abruptly, you would knock the ball out of equilibrium, making it swing. The slow turn is an example of what we call *adiabatic motion*; the abrupt turn is *diabatic motion*. See fig. 1.5. Another example should be familiar to those who have worked in hospitality: picture yourself as a waiter (or waitress), holding a tray with a drink. If you were to walk to your customer quickly, the drink would likely start waving and splashing inside the glass, and you might spill some of it (diabatic motion); but if you walked very slowly, the drink would stay at rest in the glass (adiabatic motion). Of course, you cannot let your customers wait for too long, so you need to find a balance between walking quickly and not spilling your drink.

The adiabatic principle extends to a wide range of physical phenomena, and can be formulated in a very general way. It states that if the *external parameters* of some system (e.g. the position where you hold the string, or the tray with the glass) are changed very slowly, then the *internal physics* of that system (e.g. whether the ball starts swinging, or how the drink behaves inside the glass) remains (approximately) unchanged.

Quantum mechanics also behaves this way. Say we have a system of quantum mechanical particles described by some external parameters, and the system is in its ground state. Then the adiabatic principle says that if we slowly change the external parameters, the system will (approximately) remain in its ground state.

The new ground state may have a different mathematical description as compared to the old ground state, and its energy may be different, but we are guaranteed that it will be the state with the lowest possible energy. To make this a bit more concrete, picture two molecules separated so far away from each other that they don't interact. The molecules are completely uncorrelated, and we can describe the quantum mechanical state of the two molecules by considering each molecule separately. Say both molecules are in their ground states. Now we want to bring them closely together; the external parameter here is the distance between the two molecules. When we close the distance, the molecules begin interacting, and their quantum mechanical states become correlated. But if we do it very slowly (adiabatically), the correlated state will be (approximately) the collective ground state of the two molecules together. If we do it too fast (diabatically), the molecules will end up in a state with higher energy.

Adiabatic state preparation exploits the adiabatic principle to calculate ground states of complicated quantum mechanical systems. It works in two steps, and I will explain it with help of the analogy of the waiter I just introduced. First, pick a simple quantum system whose ground state can be easily computed and given as input to the quantum computer. This step can be likened to the waiter's task of walking to the kitchen and filling up a glass, after having taken the customer's order. This part is easy: the waiter can just walk to the kitchen quickly as he is not holding any drink that could be spilled, and filling up the glass can be done in a matter of seconds. In terms of quantum mechanics, a typical example of a simple system would be one in which the electrons of the material or molecule we're studying don't interact: we pretend they are completely blind to each other's presence. Although this is not a very realistic description, calculating hypothetical ground states of such systems is easy enough that their mathematical description can be written out on paper using a relatively short formula.[17] After the waiter has filled up his glass – and the easy noninteracting ground state has been loaded onto the quantum computer – the less easy journey towards the table begins. To deliver the drink to the table successfully, the waiter must carefully walk towards it; in doing so, he slowly changes the external parameters of the glass – namely, its position in 3D space – towards the final position on the table. Meanwhile, the quantum computer simulates a time evolution in which the noninteracting quantum system slowly changes to a fully interacting system. Here, the external parameter is a somewhat abstract variable which determines the strength with which the electrons feel each other, from noninteracting to fully

[17]Noninteracting systems are typically much easier to describe than interacting systems because in noninteracting systems, we can study every constituent (every electron in this case) separately. Think of it like driving: if cars were purely noninteracting (meaning they could drive through one another), calculating the driving time from point A to point B would be as simple as dividing the distance by the speed limit, since we could completely ignore the other cars on the road. But the reality is that cars are interacting objects, and as a result, we all have to deal with traffic and congestion, which makes travel times much less predictable.

interacting.[18] If the change of the external parameter and the corresponding time evolution is carried out slowly enough – i.e. adiabatically –, then we should end up with a reasonably accurate representation of the ground state of the complex, interacting system on the quantum computer. And the waiter, careful enough not to spill a drop, ensures that the customer enjoys his drink with a happy smile.

1.3.2 Contributions of this doctoral thesis

The contributions made in this thesis are extensions to the concept of adiabatic state preparation. Chapters 3, 4 and 5 each present an improvement, or different viewpoint, to the basic method. Since these contributions are rather technical, I will stick to the analogy of the waiter in an attempt to explain them in more accessible language.

Chapter 3 deals with situations in which the waiter must avoid obstacles that are on his way to the customer. Imagine a banana peel having been left on the floor by an unorderly customer; if the waiter weren't careful to avoid it, the drink might end up somewhere completely unintended... Sometimes, such "banana peels" can occur in adiabatic state preparation as well, in that the quantum calculation can run through a point where the adiabatic principle cannot be maintained, and the preparation procedure is derailed. (The technical term for such a point is a *gap closure*.) In such cases, it is impossible to reach the ground state and the output will be a state completely different from what we wanted. But just like a waiter can walk around a banana peel, so too can a quantum computer "walk around" a gap closure, at least in certain cases. Chapter 3 studies these cases for specific systems of interacting electrons and presents methods for finding paths around the obstacles.

Chapter 4 presents a strategy for speeding up conventional adiabatic state preparation. Admittedly, the analogy with the waiter may seem a bit strange in this case, but here goes. In this protocol, the waiter walks to the table with a drink, pours a few drops into the customer's glass on the table, waits for a random time and then walks back to the kitchen. This process is repeated until the glass on the table is full. The quantum computer does something similar: it carries out an adiabatic time evolution, does nothing for a random amount of time, and then does the evolution in reverse; while it is idling, we extract a little bit of information from simulated system. Repeating this procedure and

[18] To understand this variable, it may help to think about gravity on earth. In high school, you learned that earth pulls you down with a force equal to your mass times Newton's constant g, which equals 9.81m/s^2. If g were zero, there would be no gravity, and you would fly off into space; in physics language, you and the earth would not interact through gravity. If g were a small number, you would feel gravity only slightly, and you would be said to be "weakly interacting" with the earth through gravity. If you were to simulate gravity on a computer, you could pick g, which determines the interaction strength, as you like, and run the simulation to see what happens. When modelling electrons on a quantum computer, the interaction strength between them can be modified in a similar way.

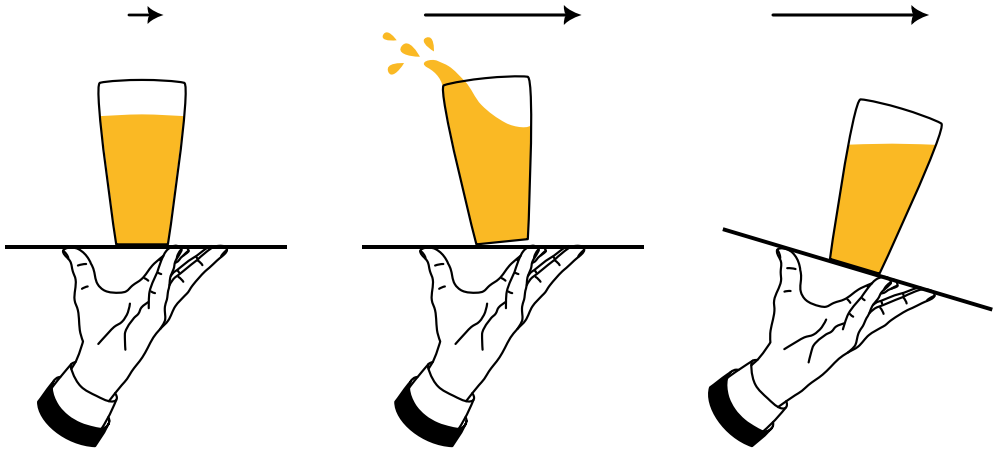


Figure 1.6. *Left*: an adiabatic waiter, carrying the drink very slowly so as to keep the drink at rest inside the glass. *Middle*: if the waiter walks too fast (diabatic motion), some of the drink will be spilled. *Right*: tilting the tray counteracts the force that causes the liquid to splash (counterdiabatic motion), allowing the waiter to walk faster without spilling anything.

combining the extracted pieces of information lets you calculate, for example, the energy of the state that was prepared by the time evolution. Why is this faster than standard adiabatic preparation, you ask? If you read footnote 13, you know that a quantum computation anyway needs to be repeated multiple times to give a reliable answer; with this method, one simply needs fewer repetitions than in the standard case.

Chapter 5 tells the story of a clever waiter wishing to deliver his drink more quickly. Instead of walking slowly and carefully, the clever waiter realises that the forces which make the liquid wave and splash in the glass can be counteracted by tilting the tray while walking (see fig. 1.6). Adjusting the tilt depending on his direction of movement, he thus manages to make his way through the restaurant more quickly. A quantum computer can do something similar, by time evolving a modified quantum system. However, calculating precisely how the quantum computer should "tilt its tray" is a very difficult problem. This chapter presents a state preparation method in which the quantum computer itself figures out how to tilt the tray, and compares this method to standard adiabatic preparation.

* * *

This is where my attempt at explaining the context and content of this thesis in understandable language comes to an end; the remainder of the thesis is technical, and only accessible with a good understanding of quantum mechanics and the associated mathematical tools. For those who wish to become more familiar with these topics, I would recommend the following introductory texts:

- *Introduction to quantum mechanics* by David Griffiths and Darrell Schroeter [GS18]. This book introduces the concepts of quantum physics in a pedagogical manner and provides a beautiful discussion on its philosophical implications.

- *Quantum computation and quantum information* by Michael Nielsen and Isaac Chuang [NC10] is a self-contained introduction to quantum computing, and extensively explains the basic principles as well as key quantum algorithms and quantum information.

If your reading ends here, I hope you enjoyed doing so and learned something interesting. In any case, I thank you, dear reader, for your attention and your valuable time.

Chapter 2

Preliminaries

This chapter serves to set notational conventions to be used throughout this thesis as well as introduce the fundamental technical concepts that form the cornerstone of the work presented in the following chapters. We set off with a description of quantum computation expressed in its most common formulation, the quantum circuit model, also known as gate-based quantum computing. We discuss how this formalism emerges from traditional quantum mechanics, and consider a few relevant quantum algorithmic routines as well as quantum computational complexity. The chapter continues with a brief review of adiabatic state preparation, which is a different paradigm of quantum computing as well as the main theme of this thesis. We quote the adiabatic theorems to be used in the following chapters. We then turn our attention to the application and the main motivation for these algorithms: the preparation of eigenstates, and specifically ground states, of (fermionic) many-body quantum systems. We consider the formulation of many-body systems in second quantisation, as well as electronic structure hamiltonians which serve as our models of condensed matter and molecular systems. The difficulty of obtaining eigenenergies and eigenstates is discussed, and followed by an outline of prevalent classical and quantum methods.

2.1 Conventions and norms

2.1.1 Conventions

Throughout this thesis, we will write matrices or operators in upright roman style, and real-valued vectors in upright bold style, instead of italics. For example, $\mathrm{A}\mathbf{v} = \lambda\mathbf{v}$. Matrix elements A_{ij} and vector elements v_i are written in italics. Complex-valued vectors of unit norm are written in *bra-ket* notation, e.g. $|\psi\rangle$ for such a vector and $\langle\psi|$ for its hermitian conjugate. The inner product between $\langle\phi|$ and $|\psi\rangle$ is written $\langle\phi|\psi\rangle$. Sometimes other styles, e.g. calligraphic ($\mathcal{C}$) are used, but these do not follow any specific convention.

We write $\|\mathbf{v}\|$ for the euclidean norm of a real-valued vector $\mathbf{v}$ and $\||\psi\rangle\|$ or $\|\psi\|$ for that of a complex-valued unit vector $|\psi\rangle$. For a vector $\mathbf{v}$, denote by $\|\mathbf{v}\|_1$ its 1-norm $\sum_i |v_i|$.

We will make use of two matrix norms: the spectral norm $\|\cdot\|$ (also known as operator norm) and the trace norm $\|\cdot\|_1$. The spectral norm is defined as $\|\mathrm{A}\| = \sup_{|\psi\rangle:\||\psi\rangle\|=1} \|\mathrm{A}|\psi\rangle\|$, or equivalently as the largest singular value of A. For hermitian matrices, this corresponds to the largest absolute eigenvalue. The trace norm is defined by $\|\mathrm{A}\|_1 = \mathrm{tr}\sqrt{\mathrm{A}^\dagger\mathrm{A}}$, which equates to the sum over all singular values of A, or over the absolute eigenvalues if A is hermitian.

The quantity $|\langle\phi|\psi\rangle|^2$ is called the *fidelity* or *overlap* between $|\psi\rangle$ and $|\phi\rangle$; conversely, $1 - |\langle\phi|\psi\rangle|^2$ is termed *infidelity* and $\sqrt{1 - |\langle\phi|\psi\rangle|^2}$ *square-root infidelity*.

Lastly, we will frequently express quantities in asymptotic or "big-O" notation. Formally, given a real- or complex-valued function g, $O(g)$ is the set of functions f such that $|f(x)| \leq C|g(x)|$ for some constant C and all x larger than some x_0. In other words, $\lim_{x\to\infty} |f(x)|/|g(x)| < \infty$. In addition, we have $f \in \Omega(g)$ if $\lim_{x\to\infty} |f(x)|/|g(x)| > 0$ (or equivalently, if $g \in O(f)$); $f \in o(g)$ if $\lim_{x\to\infty} |f(x)|/|g(x)| = 0$; and $\Theta(g) = O(g) \cap \Omega(g)$. Often, we will use the limit to zero instead of to infinity, especially if g is a function of ϵ or Δ. Sometimes we abuse notation and use these sets as if they were functions themselves, for example in expressions such as $f \leq O(g)$ which formally indicates $f(x) \leq h(x)\ \forall x$ where $h \in O(g)$.

2.1.2 Properties of vector and matrix norms

The euclidean norm, as well as the spectral and the trace norm, satisfy a triangle inequality (using subscript $_{(1)}$ to indicate either spectral or trace norm),

$$\begin{aligned} \|\mathbf{v} + \mathbf{w}\| &\leq \|\mathbf{v}\| + \|\mathbf{w}\|; \\ \||\psi\rangle + |\phi\rangle\| &\leq \|\psi\| + \|\phi\|; \\ \|\mathrm{A} + \mathrm{B}\|_{(1)} &\leq \|\mathrm{A}\|_{(1)} + \|\mathrm{B}\|_{(1)}. \end{aligned} \tag{2.1}$$

For complex-valued unit vectors $|\psi\rangle$ and $|\phi\rangle$, it may be shown by direct calculation that the square-root infidelity between them is upper bounded by the euclidean distance:

$$\begin{aligned} &\||\psi\rangle - |\phi\rangle\|^2 = 2(1 - \mathrm{Re}|\langle\phi|\psi\rangle|) \geq 2(1 - |\langle\phi|\psi\rangle|) \\ &\Leftrightarrow |\langle\phi|\psi\rangle| \geq 1 - \frac{1}{2}\||\psi\rangle - |\phi\rangle\|^2 \\ &\Leftrightarrow |\langle\phi|\psi\rangle|^2 \geq 1 - \||\psi\rangle - |\phi\rangle\|^2 + \frac{1}{4}\||\psi\rangle - |\phi\rangle\|^4 \geq 1 - \||\psi\rangle - |\phi\rangle\|^2 \\ &\Leftrightarrow \sqrt{1 - |\langle\phi|\psi\rangle|^2} \leq \||\psi\rangle - |\phi\rangle\|. \end{aligned} \tag{2.2}$$

The spectral and trace norm are submultiplicative,

$$\|\mathrm{AB}\|_{(1)} \leq \|\mathrm{A}\|_{(1)}\,\|\mathrm{B}\|_{(1)}, \tag{2.3}$$

as well as unitarily invariant, meaning that

$$\|\mathrm{UAV}\|_{(1)} = \|\mathrm{A}\|_{(1)} \tag{2.4}$$

for unitary matrices U, V.

The spectral and trace norms are related by the inequality $\|\mathrm{AB}\|_1 \leq \|\mathrm{A}\|\|\mathrm{B}\|_1$ for any matrices A and B, which may be derived, for example, from Hölder's inequality or Von Neumann's inequality [Bha13].

Both norms satisfy the so-called telescoping property: let $\mathrm{U}_1, \mathrm{U}_2, \mathrm{V}_1, \mathrm{V}_2$ be unitary matrices; then

$$\|\mathrm{U}_1\mathrm{U}_2 - \mathrm{V}_1\mathrm{V}_2\|_{(1)} \leq \|\mathrm{U}_1 - \mathrm{V}_1\|_{(1)} + \|\mathrm{U}_2 - \mathrm{V}_2\|_{(1)}, \tag{2.5}$$

This follows from the unitary invariance of the norms and a triangle equality:

$$\begin{aligned}\|\mathrm{U}_1\mathrm{U}_2 - \mathrm{V}_1\mathrm{V}_2\|_{(1)} &= \|\mathrm{U}_1\mathrm{U}_2 - \mathrm{V}_1\mathrm{U}_2 + \mathrm{V}_1\mathrm{U}_2 - \mathrm{V}_1\mathrm{V}_2\|_{(1)} \\ &\leq \|\mathrm{U}_1\mathrm{U}_2 - \mathrm{V}_1\mathrm{U}_2\|_{(1)} + \|\mathrm{V}_1\mathrm{U}_2 - \mathrm{V}_1\mathrm{V}_2\|_{(1)} \\ &= \|(\mathrm{U}_1 - \mathrm{V}_1)\mathrm{U}_2\|_{(1)} + \|\mathrm{V}_1(\mathrm{U}_2 - \mathrm{V}_2)\|_{(1)} \\ &= \|\mathrm{U}_1 - \mathrm{V}_1\|_{(1)} + \|\mathrm{U}_2 - \mathrm{V}_2\|_{(1)}. \end{aligned} \tag{2.6}$$

By induction, this property can be extended to products of any size:

$$\left\| \prod_{i=1}^{m} \mathrm{U}_i - \prod_{i=1}^{m} \mathrm{V}_i \right\| \leq \sum_{i=1}^{m} \|\mathrm{U}_i - \mathrm{V}_i\|_{(1)}. \tag{2.7}$$

The spectral and trace norms give rise to the spectral distance $\|\mathrm{A} - \mathrm{B}\|$ and trace distance $\frac{1}{2}\|\mathrm{A} - \mathrm{B}\|_1$ respectively. The trace distance is especially useful when dealing with matrices constructed as the outer product of a complex-valued unit vector with its hermitian conjugate. In particular, it holds that

$$\frac{1}{2}\||\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|\|_1 = \sqrt{1 - |\langle\phi|\psi\rangle|^2}. \tag{2.8}$$

The derivation follows from explicit calculation of the eigenvalues of $|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|$ after Gram-Schmidt orthogonalisation: if we write $|\phi\rangle = \alpha|\psi\rangle + \beta|\psi^\perp\rangle$ for some $|\psi^\perp\rangle$ orthogonal to $|\psi\rangle$ and $\alpha^2 + \beta^2 = 1$, then in the basis $\{|\psi\rangle, |\psi^\perp\rangle\}$ we have

$$\frac{1}{2}\big(|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|\big) = \frac{1}{2}\begin{bmatrix} 1 - |\alpha|^2 & -\alpha\beta^* \\ -\alpha^*\beta & -(1 - |\alpha|^2) \end{bmatrix}. \tag{2.9}$$

This matrix has eigenvalues $\pm\frac{1}{2}\sqrt{1 - |\alpha|^2}$, and since $\alpha = \langle\phi|\psi\rangle$, its trace norm equals $\sqrt{1 - |\langle\phi|\psi\rangle|^2}$.

2.2 Fundamentals of quantum computation

Physicists, mathematicians and computer scientists like to describe the reality around us in terms of states: collections of informational units that represent the configuration of an object, and which may change with time. In computer science, the basic unit of information is a bit: a variable that can take either the value 0 or 1. A state can then be formed by strings of concatenated bits, or bitstrings. In classical physics, such as electrodynamics and classical mechanics, states may be represented by real-valued vector fields like position, momentum, electromagnetic fields and the like. *Quantum states*, on the other hand, are represented by a vector $|\psi\rangle \in \mathbb{C}^N$ with unit norm, i.e. $\||\psi\rangle\| = 1$. The space of quantum states (plus all mathematical necessities such as a properly defined inner product) is called *Hilbert space*. Quantum computation is the art of processing information through the manipulation of quantum states. The smallest piece of information in quantum computing is a quantum bit, also known as *qubit*, which is defined as a two-dimensional quantum state $|\psi\rangle \in \mathbb{C}^2$. In analogy to classical bits, it is customary to choose a basis of orthonormal states $\{|0\rangle, |1\rangle\}$. With respect to this basis, a qubit is expressed as

$$|\psi\rangle = \psi_0|0\rangle + \psi_1|1\rangle \tag{2.10}$$

with $\psi_0, \psi_1 \in \mathbb{C}$. Unless either ψ_0 or ψ_1 equals zero, this state is said to be in superposition of the states $|0\rangle$ and $|1\rangle$. It is neither $|0\rangle$, nor $|1\rangle$, but something in between – "both 0 and 1", in a sense. As such, a qubit is a fundamentally different information carrier as compared to a classical bit, since it carries two complex numbers as information. Multi-qubit states can be constructed by concatenating single qubits through a tensor product, and one obtains composite states of the form

$$\begin{aligned} |\psi\rangle &= \psi_{0\cdots00}|0\rangle \otimes \cdots \otimes |0\rangle \otimes |0\rangle + \psi_{0\cdots01}|0\rangle \otimes \cdots \otimes |0\rangle \otimes |1\rangle + \cdots \\ &\quad + \psi_{1\cdots11}|1\rangle \otimes \cdots \otimes |1\rangle \otimes |1\rangle \\ &= \psi_{0\cdots00}|0\cdots00\rangle + \psi_{0\cdots01}|0\cdots01\rangle + \cdots + \psi_{1\cdots11}|1\cdots11\rangle. \end{aligned} \tag{2.11}$$

The basis consisting of these bitstring states is often referred to as the computational basis.

As we can see, a state built from n qubits carries 2^n complex numbers as information. This is a vastly bigger amount of information than contained in a string of n classical bits. The catch, however, is that most of the information in a quantum state is not directly accessible. In order to extract information from a quantum state, we must apply a *measurement* to it. According to the postulates of quantum mechanics, the outcome ω of any measurement to a quantum state is associated with a subspace S_ω of state space onto which the state is *collapsed*

upon measurement:

$$|\psi\rangle \xmapsto{\text{measurement with outcome } \omega} \frac{\Pi_{S_\omega}|\psi\rangle}{\|\Pi_{S_\omega}|\psi\rangle\|}, \tag{2.12}$$

and the probability of obtaining outcome ω is given by the normalisation factor $\|\Pi_{S_\omega}|\psi\rangle\|$. For example, if a state as in eq. 2.11 is measured in the computational basis (where the S_ω are the one-dimensional spaces spanned by $|0\cdots00\rangle$, $|0\cdots01\rangle$ and so on), we will obtain outcome $\mathbf{x}$ with probability $|\psi_{\mathbf{x}}|^2$. For the remainder of this work, we will associate measurement outcomes with real numbers. In this case, the measurement outcomes and projectors can be conveniently combined into an hermitian operator called an *observable*:

$$\mathrm{O} = \sum_\omega \omega \Pi_{S_\omega}. \tag{2.13}$$

Note that the eigenvalues of an observable are exactly the measurement outcomes ω, and the associated eigenvectors, also known as eigenstates, lie in the respective subspaces S_ω. The expected outcome of a measurement of observable with respect to a state can then be expressed as

$$\langle \mathrm{O} \rangle = \sum_\omega \omega \|\Pi_{S_\omega}|\psi\rangle\|^2 = \langle\psi|\mathrm{O}|\psi\rangle. \tag{2.14}$$

When an observable encodes the physics of a quantum mechanical system, the observable is called the *hamiltonian* of that system, and we use the symbol H for it. By "encoding the physics", we mean that it determines the time dynamics of a system in the nonrelativistic hamiltonian formulation of quantum mechanics – in other words, the hamiltonian satisfies the *Schrödinger equation*[1]:

$$\mathrm{i}\frac{\partial}{\partial t}|\psi(t)\rangle = \mathrm{H}(t)|\psi(t)\rangle. \tag{2.15}$$

The eigenvalues of the hamiltonian are called *energies*, denoted E_n. The corresponding *energy eigenstates* shall be denoted $|E_n\rangle$ or $|n\rangle$, depending on context and readability. Since H is hermitian, solving the Schödinger equation tells us that quantum states at different times are connected by the following linear map:

$$\begin{aligned} |\psi(t)\rangle &= \mathrm{U}(t,0)|\psi(0)\rangle; \\ \mathrm{U}(t,0) &= \mathcal{T}\exp\left[-\mathrm{i}\int_0^t \mathrm{d}t'\,\mathrm{H}(t')\right] \end{aligned} \tag{2.16}$$

where $\mathcal{T}$ is the time ordering superoperator. Because the hamiltonian is involved in the Schrödinger equation and therefore dictates the form of the connection of two states separated in time, hamiltonians are said to *generate motion (in time)*.

[1] We work in units such that $\hbar = 1$.

The connecting map U can alternatively be expressed as

$$\mathrm{U}(t,0) = \lim_{r\to\infty} \prod_{m=r}^{0} \mathrm{e}^{-\mathrm{i}\mathrm{H}(t_m)\delta t} \tag{2.17}$$

where $\delta t := t/r$ and $t_m := m\delta t$. If H is constant in time, or *time-independent*, eqs. 2.16 and 2.17 reduce to $\mathrm{U}(t,0) = \exp[-\mathrm{i}\mathrm{H}t]$. Since H is hermitian, U is a unitary map. That the unitarity of U is a necessity, we expected beforehand, since a quantum state must remain normalised at any point in time for the measurement probabilities to be validly defined.

The unitarity condition now allows us to specify the processing of quantum information. In analogy to logical gates in a classical processor, a quantum gate can in principle be defined as any unitary operation on a quantum state. However, in the same way that a classical processor cannot execute every boolean operation without decomposing it into a sequence of smaller gates (e.g. NAND gates), large, multi-qubit operations must be composed from smaller units as well. The Solovay-Kitaev theorem [Kit97, DN05] tells us that this can be done by using gates from only a small set, called a *universal gate set*. We will refer to such gates as *elementary gates*. Typical examples of universal gate sets are the set $\{\mathrm{R_X}(\theta), \mathrm{R_Y}(\theta), \mathrm{R_Z}(\theta), {}^{C}\mathrm{X}\}$, where

$$\mathrm{R_X}(\theta) = \mathrm{e}^{-\mathrm{i}\mathrm{X}\theta/2} = \begin{bmatrix} \cos\theta & -\mathrm{i}\sin\theta \\ -\mathrm{i}\sin\theta & \cos\theta \end{bmatrix}, \tag{2.18}$$

$$\mathrm{R_Y}(\theta) = \mathrm{e}^{-\mathrm{i}\mathrm{Y}\theta/2} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \tag{2.19}$$

$$\mathrm{R_Z}(\theta) = \mathrm{e}^{-\mathrm{i}\mathrm{Z}\theta/2} = \begin{bmatrix} \mathrm{e}^{-\mathrm{i}\theta/2} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\theta/2} \end{bmatrix}, \tag{2.20}$$

$${}^{C}\mathrm{X} = \mathrm{I}_2 \oplus \mathrm{X} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{2.21}$$

with X, Y and Z the familiar Pauli matrices. The operation ${}^{C}\mathrm{X}$ is often denoted as CNOT (our notation here is nonstandard). Another freqently used set is the so-called Clifford+T set $\{\mathrm{D}, \sqrt{\mathrm{Z}}, \sqrt[4]{\mathrm{Z}}, {}^{C}\mathrm{X}\}$, where

$$\mathrm{D} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{2.22}$$

is the Hadamard gate (this operator is usually written H, but we use nonstandard notation to avoid confusion with hamiltonians). ${}^{C}\mathrm{X}$ is said to be an entangling gate, because it cannot be written as a tensor product of two single-qubit gates. Entangling gates, when applied to a product states, produce *entangled states*.

Entangling gates are indispensable for universal quantum computation, since otherwise the space of computed quantum states would be restricted to product states.

All of what we have discussed so far applies to closed quantum systems, which are described by *pure states*, i.e. normalised vectors in $\mathbb{C}^N$. In this thesis, we will also encounter *mixed states* in open systems. A mixed state can be viewed as a probability distribution over pure states. The corresponding mathematical object is a *density operator*, which is a matrix of the form

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{2.23}$$

with mutually orthogonal $|\psi_i\rangle$, each $p_i \geq 1$ and $\sum_i p_i = 1$. (If any of the p_1 equals 1, the density operator represents a pure state.) Alternatively, a density operator may be defined as a positive-semidefinite, hermitian matrix with unit trace. In density operator notation, unitary operations carry out the map

$$\rho \overset{\mathrm{U}}{\mapsto} \mathrm{U}\rho\mathrm{U}^\dagger \tag{2.24}$$

and the expected outcome of measuring an observable O with respect to a state ρ can be written

$$\langle \mathrm{O} \rangle = \mathrm{tr}[\mathrm{O}\rho] = \sum_i p_i \langle\psi_i|\mathrm{O}|\psi_i\rangle. \tag{2.25}$$

As an extension to unitary operations, we can also define probabilistic operations on quantum states, which are known as channels:

$$\rho \overset{\text{channel}}{\longmapsto} \sum_j q_j \mathrm{U}_j \rho \mathrm{U}_j^\dagger \tag{2.26}$$

where each U_j is unitary, each $q_j \geq 0$ and $\sum_j q_j = 1$. We will make use of these channels in chapter 4.

In this thesis, we presume access to a device that can apply unitary transformations and channels to quantum states. Such a device is called a *quantum computer*. We also make the assumption of *fault tolerance*, where a quantum computer can implement any elementary gate with zero error, and where a quantum state can be maintained for any amount of time without decoherence occurring.

2.3 Adiabatic state preparation

2.3.1 The adiabatic approximation

In this thesis, we will mostly be concerned with a specific kind of quantum dynamics known as *adiabatic dynamics*. Adiabatic dynamics occurs when a

hamiltonian $\mathrm{H}(t)$ changes slowly in time. In a similar way that the internal physics of a classical system is approximately unchanged when its external parameters are varied slowly, so too will an eigenstate of a slowly varied hamiltonian mostly remain in its instantaneous ground space. Specifically, when at an initial time 0 a system starts out in an eigenstate $|n(0)\rangle$ of a hamiltonian $\mathrm{H}(0)$, and the parameters of $\mathrm{H}(t)$ are varied slowly in time until a final time T, then the adiabatic approximation says that the final state $|\psi(T)\rangle$, according to the Schrödinger equation, will be close to the corresponding eigenstate $|n(T)\rangle$. The most basic argument for this approximation goes as follows: let $\mathrm{H}(\boldsymbol{\lambda}(t))$ be parametrised by a time-dependent parameter vector $\boldsymbol{\lambda}(t)$ and let $|\psi(t)\rangle = \sum_n \psi_n(t)|n(t)\rangle$ be a generic quantum state. Then by applying the Schrödinger equation and multiplying by $\langle m(t)|$ on both sides we get

$$\begin{aligned} \langle m(t)|\mathrm{i}\partial_t|\psi(t)\rangle = \langle m(t)|\mathrm{H}(\boldsymbol{\lambda}(t))|\psi(t)\rangle \Leftrightarrow \\ \partial_t\psi_m(t) + \sum_n \psi_n(t)\,\langle m(t)|\partial_{\boldsymbol{\lambda}}|n(t)\rangle\,\partial_t\boldsymbol{\lambda}(t) = -\mathrm{i}E_m(t)\psi_m(t), \end{aligned} \tag{2.27}$$

and we see that in the limit $\partial_t\boldsymbol{\lambda}(t) \to 0$, this becomes a decoupled system of equations in the coefficients $\psi_m(t)$. Hence, if $|\psi(0)\rangle$ is an eigenstate of $\mathrm{H}(0)$, it will remain an eigenstate at later times.

The notion of adiabaticity gives us a nice way to construct quantum algorithms. The idea is to encode the solution to a computational problem into an eigenspace of some hamiltonian $\mathrm{H_f}$; if we set up a system in the corresponding eigenstate of some other (usually simple) hamiltonian $\mathrm{H_i}$, and slowly evolve the system from $\mathrm{H}(t = 0) = \mathrm{H_i}$ to $\mathrm{H}(t = T) = \mathrm{H_f}$, then by the adiabatic approximation, we obtain a state that is close to a solution to our problem. This strategy for quantum computing is known as *adiabatic quantum computation* (AQC). From a physicist's point of view, one is often more interested in preparing an eigenstate of a given hamiltonian in order to obtain insight into its physical characteristics rather than solving some computational problem; in this context, the same procedure will be referred to as *adiabatic state preparation* (ASP). Since the motivation for the work in this thesis is of the latter kind, we will use the name ASP. Oftentimes, the target eigenstate is the ground state of the final hamiltonian. The intermediate eigenstates $|n(\boldsymbol{\lambda}(t))\rangle$ are called the *instantaneous* eigenstates $\mathrm{H}(\boldsymbol{\lambda}(t))$.

A straightforward and often used way to implement ASP is by so-called interpolating hamiltonians of the form

$$\mathrm{H}(t) = \mathrm{H_i} + \lambda(t)(\mathrm{H_f} - \mathrm{H_i}) \tag{2.28}$$

where $\lambda(0) = 0$ and $\lambda(T) = 1$. In such cases, $\lambda(t)$ is called the *interpolation function* or *schedule*. Linear schedules with $\lambda(t) = t/T$ frequently occur in the literature; yet, as we will discuss, more clever choices exist.

2.3.2 Adiabatic theorems

In order to make ASP actually useful, we need more than just the basic argument for the adiabatic approximation. After all, what exactly is meant by "slowly enough"? Certainly, the limit $\partial_t \boldsymbol{\lambda}(t) \to 0$, which is equivalent to $T \to \infty$, is insufficient in practice. Instead, we have to accept that in finite-time dynamics, the eigenstates are not fully decoupled, and that *non-adiabatic transitions* into other eigenstates will occur; in order to define "slowly enough", we therefore require a relationship between the evolution time and the error (the overlap of the final state with the space orthogonal to the target eigenstate). To gain a bit of intuition what the dependence should look like, we can figure out the nonadiabatic coupling coefficients by differentiating the equation $\langle m(t)|\mathrm{H}(t)|n(t)\rangle = 0$ $(n \neq m)$ with respect to time[2]:

$$\begin{aligned}\partial_t\big(\langle m(t)|\mathrm{H}(t)|n(t)\rangle\big) &= \langle m(t)|\partial_t^\dagger|n(t)\rangle E_n(t) + \langle m(t)|\partial_t \mathrm{H}(t)|n(t)\rangle \\ &\quad + \langle m(t)|\partial_t|n(t)\rangle E_m(t) = 0 \\ \Rightarrow \langle m(t)|\partial_t|n(t)\rangle &= \frac{\langle m(t)|\partial_t \mathrm{H}(t)|n(t)\rangle}{E_n(t) - E_m(t)}. \end{aligned} \tag{2.29}$$

According to eq. 2.29, the magnitude of the nonadiabatic transitions depends on the rate of change in the hamiltonian, and inversely on the instantaneous energy differences between the instantaneous eigenstate we wish to follow and the other eigenstates. In the extreme case where the instantaneous target eigenstate is degenerate at some time, adiabatic preparation may even be impossible. An instance where this happens is a setting in which the ground and first excited state are decoupled (e.g. because of symmetry reasons, cf. chapter 3), and their energy levels cross: if one starts in the ground state of the initial hamiltonian, one will obtain the first excited state of the final hamiltonian, and vice versa.

Many results have been derived [AL18] which put a rigorous and precise upper bound on the error in terms of the *gap* $\omega_n(t) := \min_{m\neq n} |E_m(t) - E_n(t)|$ or *minimum gap* $\Delta_n = \min_{t\in[0,T]} \omega_n(t)$ and the time derivative of the hamiltonian. These results are known as *adiabatic theorems*. We cite the two theorems that we use in subsequent chapters of this thesis. Both of these theorems assume an ASP setting as in eq. 2.28 where the time-dependent hamiltonian is an interpolation with schedule $\lambda(t)$. The first of the theorems is due to Jansen, Seiler and Ruska [JRS07]. Their results are expressed in terms of projectors with dimensionality possibly greater than one; since we're only interested in obtaining single eigenstates, we simplify their theorem to the case of one-dimensional subspaces.

2.3.1. Theorem (Jansen et al. [JRS07], theorem 3). *Suppose that* $\mathrm{H}(\lambda)$ *is hermitian and twice differentiable, and that the spectrum of* $\mathrm{H}(\lambda)$ *restricted to* $|n(\lambda)\rangle$

[2]Depending on preference, $\langle m(t)|\partial_t^\dagger$ may be read $\langle \partial_t m(t)|$. Note that ∂_t is an antihermitian operator.

is separated by a gap $\omega_n(\lambda)$ *from the rest of the spectrum of* $\mathrm{H}(\lambda)$. *Let* $\mathrm{U}(T,0) = \mathcal{T}\exp[-\mathrm{i}\int_0^T \mathrm{d}t\,\mathrm{H}(t)]$. *Then*

$$1 - |\langle n(\lambda(T))|\mathrm{U}(T,0)|n(\lambda(0))\rangle|^2 \leq A(T)^2 \tag{2.30}$$

where

$$A(T) \leq \frac{1}{T}\frac{\|\partial_\lambda \mathrm{H}\|}{\omega_n^2}\bigg|_{b.c.} + \frac{1}{T}\int_{\lambda(0)}^{\lambda(T)} \mathrm{d}\lambda \left(\frac{\|\partial_\lambda^2 \mathrm{H}\|}{\omega_n^2} + 7\frac{\|\partial_\lambda \mathrm{H}\|^2}{\omega_n^3}\right) \tag{2.31}$$

and $f|_{b.c.}$ *is a shorthand for* $f(\lambda(0)) + f(\lambda(T))$. *(Here, b.c. stands for boundary condition.)*

In this thesis, we are interested in settings where Δ_n is small (possibly exponentially small in the system size). In such cases, the last term in eq. 2.31 dominates, and a sufficient condition on T to achieve a fidelity at least $1-\epsilon^2$ with the target eigenstate is $T \in \Omega(\Delta_n^{-3}\|\partial_\lambda \mathrm{H}\|_{\infty,2}^2 \epsilon^{-1})$. However, if we have information about $\omega_n(\lambda)$, the evolution time can be reduced. The authors report that, in cases such as unstructured search, in fact $\int \mathrm{d}\lambda\, \omega_n^{-3}(\lambda) \in O(\Delta_n^{-2})$; a further improvement $T \in \Omega(\Delta_n^{-1})$ can even be achieved by tuning $\partial_t\lambda(t)$ to the size of the gap $\omega_n(\lambda(t))$. However, such information about the gap is typically not available (in fact, determining whether a system is gapped at all turns out to be an undecidable problem [CPGW22]).

The second theorem is due to Reichardt, and achieves general improvement in the minimum gap dependency over the condition $T \in \Omega(\Delta_n^{-3})$ depending on the smoothness of the hamiltonian. The flipside is that this theorem offers no room for tuning the schedule $\lambda(t)$.

2.3.2. Theorem (Reichardt [Rei04]). *Suppose that* $\mathrm{H}(\lambda)$ *is hermitian and* $k+1$ *times differentiable, and that the spectrum of* $\mathrm{H}(\lambda)$ *restricted to* $|n(\lambda)\rangle$ *is separated by a gap* $\omega_n(\lambda)$ *from the rest of the spectrum of* $\mathrm{H}(\lambda)$. *Let* $\mathrm{U}(T,0) = \mathcal{T}\exp[-\mathrm{i}\int_0^T \mathrm{d}t\,\mathrm{H}(t)]$. *Let* $\mathrm{P}(\lambda) = |n(\lambda)\rangle\langle n(\lambda)|$. *Then there is a constant* $c(k)$ *such that, assuming* $\left\|\partial_\lambda^l \mathrm{H}(\lambda)\right\| \leq \Gamma(\lambda)$ *for all* $l \leq k+1$,

$$\|(\mathrm{I} - \mathrm{P}(\lambda(T))\,\mathrm{U}(T,0)\,\mathrm{P}(\lambda(0))\| \leq c(k)\left(\frac{1}{T}\frac{1}{\omega_n^2}\Gamma\bigg|_{b.c.} + \max_{\lambda(0)\leq\lambda\leq\lambda(T)} \frac{1}{T^k}\frac{1}{\omega_n^{2k+1}}\Gamma^{k+1}\right) \tag{2.32}$$

and $f|_{b.c.}$ *is a shorthand for* $f(\lambda(0)) + f(\lambda(T))$.

Rewriting eq. 2.32 yields a criterion $T \in \Omega(\Delta_n^{-2+1/k}\Gamma^{1+1/k}\epsilon^{-1/k})$. For large k, this asymptotic bound is generally better than that by Jansen et al.; however, the constant $c(k)$ may grow exponentially in k.

Other useful results state that a quadratic gap dependence [EH12] or logarithmic error dependence [GMC16] may be obtained if it is assumed that $\mathrm{H}(\lambda)$ belongs to a Gevrey class[3].

[3] A hamiltonian $\mathrm{H}(\lambda)$ belongs to a Gevrey class G^α if $\partial_\lambda \mathrm{H} \neq 0\ \forall \lambda \in [\lambda_\mathrm{i}, \lambda_\mathrm{f}]$ and there exist constants $C, D > 0$ such that for all $p > 1$, $\max_{\lambda\in[\lambda_\mathrm{i},\lambda_\mathrm{f}]} \|\partial_\lambda^p \mathrm{H}(\lambda)\| \leq CD^p k^{\alpha k}$.

2.3.3 Simulating adiabatic dynamics on a quantum computer

Adiabatic state preparation is naturally suited for analogue quantum simulators that implement dynamics through a continuously tunable hamiltonian. This does not mean, though, that ASP cannot be implemented on a gate-based quantum computer, as discussed in the first section. In principle, any quantum dynamics can be translated into a sequence of quantum gates through *hamiltonian simulation*. While there is a bit of overhead in going from analogue to digital simulation, this overhead is small for the most sophisticated methods. We highlight a few of these techniques here.

Let us begin with the simpler case of time-independent hamiltonians. Typically, hamiltonians are given as sums $\sum_i \mathrm{H}_i$ such that each H_i is *local*, i.e. acts on, or couples, only a few qubits. If all these terms commute, the task is easy, since the dynamics of the system then equals a product $\prod_i \exp[-\mathrm{i}\mathrm{H}_i t]$ and all the factors can be synthesised as few-qubit gates [BZC+23]. However, this is not the case for most interesting hamiltonians. In general, for two hermitian operators A and B that do not necessarily commute, the exponential $\mathrm{e}^{-\mathrm{i}(\mathrm{A}+\mathrm{B})t}$ can be split into the product $\mathrm{e}^{-\mathrm{i}\mathrm{A}t}\mathrm{e}^{-\mathrm{i}\mathrm{B}t}$ at the cost of an error $\|\mathrm{e}^{-\mathrm{i}(\mathrm{A}+\mathrm{B})t} - \mathrm{e}^{-\mathrm{i}\mathrm{A}t}\mathrm{e}^{-\mathrm{i}\mathrm{B}t}\| \in O(\|\mathrm{A}\|\,\|\mathrm{B}\|t^2)$. Clearly, if the time step t is small, so is the error. If t is large, one may divide t into r parts to obtain $\|\mathrm{e}^{-\mathrm{i}(\mathrm{A}+\mathrm{B})t} - (\mathrm{e}^{-\mathrm{i}\mathrm{A}t/r}\mathrm{e}^{-\mathrm{i}\mathrm{B}t/r})^r\| \in O(\|\mathrm{A}\|\,\|\mathrm{B}\|t^2/r)$. Here, the additional factor r in the error follows from linear propagation of errors as a result of the telescoping property of the operator norm. The parameter r can then be chosen to achieve any desired precision; namely, to upper bound the error by ϵ, r must be taken to be at least $O(t^2/\epsilon)$. This is the most elementary kind of decompositions known as Trotter formulas. More fancy Trotter decompositions exist which bring down the cost to $O\big((\sum_j \|\mathrm{H}_j\|)t^{1+o(1)}/\epsilon^{o(1)}\big)$ [CST+21a].

Another interesting technique for hamiltonian simulation is qDRIFT [Cam19]. qDRIFT is a sampling method in which the terms H_i are repeatedly sampled according to their weights (in terms of spectral norm), and with each sample, the exponential $\mathrm{e}^{-\mathrm{i}\mathrm{H}_i\delta t}$ is applied for some small δt. Specifically, the probability to choose H_i is $\|\mathrm{H}_i\|/\sum_j \|\mathrm{H}_j\|$. This method requires $O\big((\sum_j \|\mathrm{H}_j\|)t^2/\epsilon\big)$ samples.

The last category of techniques we mention is that of block encoding methods. The idea is to block encode a hamiltonian H into a unitary U with support on a larger Hilbert space:

$$\mathrm{U} = \begin{bmatrix} \mathrm{H} & \cdot \\ \cdot & \cdot \end{bmatrix}. \tag{2.33}$$

Such a block encoding can be constructed efficiently if H is provided as a linear combination of unitaries (LCU): $\mathrm{H} = \sum_{i=1}^{\ell} \beta_i \mathrm{V}_i$ with $\beta_i \in \mathbb{R}$. Namely, if we

define operators W and U such that

$$\begin{aligned} \mathrm{W} : |0\rangle \mapsto |W\rangle &= \frac{1}{\sqrt{\|\boldsymbol{\beta}\|_1}} \sum_i \sqrt{\beta_i} |i\rangle, \\ \mathrm{U} &= \sum_i |i\rangle\langle i| \otimes \mathrm{V}_i \end{aligned} \tag{2.34}$$

then a simple calculation shows that $\langle W|\mathrm{U}|W\rangle = \mathrm{H}/\|\boldsymbol{\beta}\|_1$. If each V_i can be implemented with at most C quantum gates, then W can be constructed with $O(\ell)$ elementary gates, and U with $O(C\ell)$ gates.

Given such a block encoding, techniques known as quantum signal processing [LC17], qubitisation [LC19a] and quantum singuluar value transformation [GSLW19] can be employed to implement "arithmetic" on H, including the transform $\mathrm{H} \mapsto \mathrm{e}^{-\mathrm{i}\mathrm{H}t}$. In the end, one needs $O(\|\boldsymbol{\beta}\|_1 t + \log(1/\epsilon))$ applications of W and U to implement time-independent hamiltonian simulation this way; this is also asymptotically optimal in t and ϵ [LC19a].

These techniques have, in part, been generalised to time-dependent hamiltonians as well. The "fancy" Trotter decompositions can be straightforwardly implemented with eq. 2.17 in mind; the details have been worked out by Wiebe et al. [WBHS10] and will be discussed at length in chapter 5. qDRIFT can be generalised to time-dependent settings by discretising time and sampling a term at each time step, as proposed by Berry et al. [BCS+20]. Block-encoding methods have been considered in the time-dependent setting, with attempts to replicate the optimal time complexity scaling from time-independent methods [WWRL22]. Alternatively, the Dyson series method [LW19, KSB19a] implements the approximation $\mathrm{U}(t_\mathrm{f}, t_\mathrm{i}) \approx \sum_{k=0}^{K} \frac{(-\mathrm{i})^k}{k!} \int_{t_\mathrm{i}}^{t_\mathrm{f}} \cdots \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathcal{T}[\mathrm{H}(t_k) \cdots \mathrm{H}(t_1)] \mathrm{d}^k t$. In this process, oracle access to the hamiltonian is assumed, in the form of a time-independent and a time-dependent oracle, where the time-dependent oracle is a direct sum of time-independent oracles over a finite set of times in the simulation interval (see ref. LW19 for details). While efficient methods have been developed to decompose the time-independent oracle into a sequence of quantum gates [SBM06, CW12, CMN+18, LC19b], such decompositions for the time-dependent oracle are only known in special cases [LW19]. Nonetheless, this method remains the most efficient to this day in terms of query complexity with $O(t \log(1/\epsilon))$ uses of the time-dependent oracle.

2.4 State preparation of fermionic many-body systems

2.4.1 Fermionic many-body hamiltonians

So far, we have rather abstractly described quantum systems and how to prepare certain states by adiabatic means, without specifying what these systems should actually represent. In this thesis, we will consider hamiltonians that describe systems of many interacting quantum particles. In particular, we are interested in the properties of electrons in a potential of heavy, positively charged particles, e.g. atomic nuclei. We restrict ourselves to the nonrelativistic case. The collective state of such a system can be described by a wave function $\Psi(\mathbf{R}, \mathbf{r}, \boldsymbol{\sigma})$ where $\mathbf{R} = \mathbf{R}_1, \cdots, \mathbf{R}_{n_\mathrm{N}}$ ($\mathbf{R}_I \in \mathbb{R}^3$) denotes the nuclear positions, $\mathbf{r} = \mathbf{r}_1, \cdots, \mathbf{r}_{n_\mathrm{e}}$ ($\mathbf{r}_i \in \mathbb{R}^3$) stands for the electronic positions and $\boldsymbol{\sigma} = \sigma_1, \cdots, \sigma_{n_\mathrm{e}}$ ($\sigma_i \in \{\uparrow, \downarrow\}$) are the electron spins. The hamiltonian of such a system, in its most generic form, is given by[4]

$$\mathrm{H} = \sum_I \mathrm{T}_I + \sum_i \mathrm{T}_i + \sum_{I;\, J>I} \mathrm{V}_{IJ} + \sum_{I;\, i} \mathrm{V}_{Ii} + \sum_{i;\, j>i} \mathrm{V}_{ij} \tag{2.35}$$

where the uppercase indices I, J run over the nuclei, the lowercase indices i, j run over the electrons; T is kinetic energy,

$$\mathrm{T}_I = -\frac{1}{2m_I}\frac{\partial^2}{\partial \mathbf{R}_I^2}, \quad \mathrm{T}_i = -\frac{1}{2}\frac{\partial^2}{\partial \mathbf{r}_i^2}, \tag{2.36}$$

and V is the Coulomb interaction potential,

$$\mathrm{V}_{IJ} = \frac{1}{|\mathbf{R}_I - \mathbf{R}_J|}, \quad \mathrm{V}_{Ii} = -\frac{1}{|\mathbf{R}_I - \mathbf{r}_i|}, \quad \mathrm{V}_{ij} = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{2.37}$$

Hamiltonians of this form are the central object of study in quantum chemistry and condensed matter theory, where they are used to describe molecular systems and materials that consist of multiple atoms. Nuclei are much heaver than electrons, hence their dynamics takes place on a much larger time scale than that of the electrons. For this reason, we will work in what is known as the Born-Oppenheimer approximation, which treats the nuclei as static particles of infinite mass. In this approximation, the dynamics of the electrons and nuclei are decoupled in the same sense as in the adiabatic approximation: the nuclear dynamics is demoted to a set of slowly changing external parameters, so that we can restrict our focus to the electronic dynamics. When the nuclear motion is slow, we expect the electrons to adjust (almost) adiabatically to the moving nuclei. As

[4] We work in units such that $m_e = e = 4\pi\varepsilon_0 = 1$, with m_e the mass of an electron, e the electrostatic constant and ε_0 the vacuum permittivity.

such, we split $\Psi(\mathbf{R}, \mathbf{r}, \boldsymbol{\sigma})$ into a tensor product $\chi(\mathbf{R})\psi(\mathbf{r}, \boldsymbol{\sigma})$, and will only concern ourselves with the electronic part $\psi(\mathbf{r}, \boldsymbol{\sigma})$. The behaviour of this electronic part is then governed by what is known as the *electronic structure* hamiltonian, which is the electronic portion of the full nuclear-electronic hamiltonian, parametrised by the nuclear positions:

$$\mathrm{H}(\mathbf{R}) = \sum_i \Big(\mathrm{T}_i + \sum_I \mathrm{V}_i(\mathbf{R}_I) \Big) + \sum_{ij} \mathrm{V}_{ij} \tag{2.38}$$

(where $\mathrm{V}_i(\mathbf{R}_I) = \mathrm{V}_{Ii}$).

In the end, we are interested in modelling such systems with quantum computers. Observe that, since $\psi(\mathbf{r}, \boldsymbol{\sigma})$ is a multi-electron state, it can in principle be expanded in a basis of product states

$$\psi(\mathbf{r}, \boldsymbol{\sigma}) = \sum_k c_k \Big[\bigotimes_{i=1}^{n_\mathrm{e}} \psi_i^k(\mathbf{r}_i, \sigma_i) \Big]. \tag{2.39}$$

Since a quantum computer operates on a finite-dimensional Hilbert space, the Hilbert space in which the electronic states live is typically projected onto a smaller subspace, by choosing a finite number of one-body basis functions or *modes* $\{\phi_P\}_{P=1}^L \ni \psi_i^k$, for some integer $L > 0$. If at least one of the factors ψ_i^k in a single product state equals ϕ_P, then ϕ_P is said to be *occupied* in this product state. In quantum chemistry, one-body basis functions are usually modelled after the eigenfunctions of the hydrogen atom and are hence known as *(spin) orbitals*. In condensed matter contexts, it can be more useful to work with states that are localised either in space (on the atomic nuclei, for example) or in momentum space. These basis functions are referred to as *Wannier functions*. Seeing as each σ_i can take two values, it is often convenient to define separate basis functions $\phi_{p\uparrow}$ and $\phi_{p\downarrow}$ that are functions of spatial coordinates only. In settings without spin-spatial (spin-orbit) coupling, these two functions may be chosen identically so that essentially every mode can be occupied by a spin-up and a spin-down particle. Such basis functions are called *restricted*. We write spatial modes with lowercase indices and use uppercase indices otherwise.

Product states as in eq. 2.39, however, do not form a good basis to describe many-body eigenstates, because H obeys an exchange symmetry: if we exchange electron i with electron j, the hamiltonian remains the same. In other words, the hamiltonian commutes with all exchange operators

$$\begin{aligned} \mathrm{C}_{ij} : \psi(\cdots, \mathbf{r}_i, \cdots, \mathbf{r}_j, \cdots; \cdots, \sigma_i, \cdots, \sigma_j, \cdots) \\ \mapsto \psi(\cdots, \mathbf{r}_j, \cdots, \mathbf{r}_i, \cdots; \cdots, \sigma_j, \cdots, \sigma_i, \cdots). \end{aligned} \tag{2.40}$$

Since C_{ij}^2 equals the identity, all eigenstates of H must therefore be either fully symmetrised or fully antisymmetrised:

$$\psi(\mathbf{r}, \boldsymbol{\sigma}) = \sum_k \tilde{c}_k \Big[\frac{1}{n_\mathrm{e}!} \sum_{\pi \in S_{n_\mathrm{e}}} (\mathrm{sgn}(\pi))^s \bigotimes_{i=1}^{n_\mathrm{e}} \psi_i^k(\mathbf{r}_{\pi(i)}, \sigma_{\pi(i)}) \Big] \tag{2.41}$$

so that $\mathrm{C}_{ij}\,\psi(\mathbf{r},\boldsymbol{\sigma}) = (-1)^s\psi(\mathbf{r},\boldsymbol{\sigma})$ ($s = 0$ if symmetrised, $s = 1$ if antisymmetrised). Particles described by symmetric states are *bosons*, and asymmetric particles are *fermions*. For reasons not discussed here, electrons are treated as fermions. Note that the same one-body function cannot be occupied twice in a fermionic system (that is, by two particles of the same spin), as antisymmetrisation cancels out such states. This property of fermions is known as the *Pauli exclusion principle*. A state with $\tilde{c}_1 = 1$ (i.e. a single antisymmetrised state) will be called a *free-particle state*.

Working with many-body states in this *first-quantised* formulation gets cumbersome quickly as the number of particles grows. It is therefore much more convenient to work in what is called *second quantisation*. A second-quantised state is represented as a bitstring that tracks which one-body functions are occupied and which are not. For example, the state $|1001\rangle$ can represent a two-body antisymmetrised product state in a Hilbert space where the one-body functions labelled by $P = 1$ and $P = 4$ are occupied. The Hilbert space $\mathcal{H}$ of these *occupation states* is a direct sum $\mathcal{H} = \bigoplus_{n=0}^{n_\mathrm{e}} \mathcal{H}_n$ where $\mathcal{H}_n$ is the subspace of antisymmetrised states with exactly n electrons. The state containing zero electrons is referred to as the *vacuum state* and will be written $|\mathrm{vac}\rangle$. An indispensable tool in second quantisation are the *creation* and *annihilation operators* $\mathrm{a}_P^\dagger$ and a_P which are defined for each one-body function ϕ_P. The annihilation operator a_P maps an state $\psi \in \mathcal{H}_n$ to a state $|\psi'\rangle \in \mathcal{H}_{n-1}$ by removing the electron that occupies ϕ_P, provided ϕ_P is occupied; otherwise it maps to zero. Conversely, the creation operator $\mathrm{a}_P^\dagger : \mathcal{H}_n \to \mathcal{H}_{n+1}$ adds an electron occupying ϕ_P unless it was already occupied, in which case it also maps to zero. The behaviour of $\mathrm{a}_P^\dagger$ (a_P) can be summarised by the canonical anticommutation relations

$$\{\mathrm{a}_P, \mathrm{a}_Q\} = \{\mathrm{a}_P^\dagger, \mathrm{a}_Q^\dagger\} = 0; \qquad \{\mathrm{a}_P^\dagger, \mathrm{a}_Q\} = \delta_{PQ} \tag{2.42}$$

where $\{\mathrm{A}, \mathrm{B}\} = \mathrm{AB} + \mathrm{BA}$. Like the one-body functions themselves, the creation and annihilation operators can be expressed in a split space-spin fashion, defining $\mathrm{a}_{p\uparrow}^{(\dagger)}$ and $\mathrm{a}_{p\downarrow}^{(\dagger)}$ accordingly.

It can be shown that in second quantisation, the electronic structure hamiltonian (eq. 2.38) admits the form

$$\mathrm{H} = \sum_{pq} \sum_{\sigma\in\{\uparrow,\downarrow\}} h_{pq}^{\sigma}\, \mathrm{a}_{p\sigma}^\dagger \mathrm{a}_{q\sigma} + \frac{1}{2} \sum_{pqrs} \sum_{\sigma,\tau\in\{\uparrow,\downarrow\}} g_{pqrs}^{\sigma\tau}\, \mathrm{a}_{p\sigma}^\dagger \mathrm{a}_{r\tau}^\dagger \mathrm{a}_{s\tau} \mathrm{a}_{q\sigma} \tag{2.43}$$

where the coefficients h_{pq}^{σ} and $g_{pqrs}^{\sigma\tau}$ are given by

$$h_{pq}^{\sigma} = \int \mathrm{d}\mathbf{r}_1\, \phi_{p\sigma}(\mathbf{r}_1) \Big[\mathrm{T}_1 + \sum_I \mathrm{V}_{I1}\Big] \phi_{q\sigma}(\mathbf{r}_1), \tag{2.44}$$

$$g_{pqrs}^{\sigma\tau} = \int \mathrm{d}\mathbf{r}_1\, \mathrm{d}\mathbf{r}_2\, \phi_{p\sigma}(\mathbf{r}_1)\, \phi_{q\sigma}(\mathbf{r}_1)\, \mathrm{V}_{12}\, \phi_{r\tau}(\mathbf{r}_2)\, \phi_{s\tau}(\mathbf{r}_2) \tag{2.45}$$

where the 1 and 2 subscripts on T and V refer to the two electrons. Again, in the case of restricted one-body functions, the σ and $\sigma\tau$ superscripts can be dropped from the h and g coefficients.

The electronic structure hamiltonian may be approximated further by restricting the electrons to orbitals localised at sites arranged on a lattice, and neglecting any Coulomb interaction between different sites. Taking one mode per site, one arrives at the single-band *Fermi-Hubbard* hamiltonian,

$$\mathrm{H} = j \sum_{\langle p,q\rangle,\sigma} (\mathrm{a}^\dagger_{p\sigma}\mathrm{a}_{q\sigma} + \mathrm{a}^\dagger_{q\sigma}\mathrm{a}_{p\sigma}) + U \sum_{p} \mathrm{n}_{p\uparrow}\mathrm{n}_{p\downarrow} + \mu \sum_{p\sigma} \mathrm{n}_{p\sigma} \tag{2.46}$$

where $\mathrm{n}_{p\sigma} = \mathrm{a}^\dagger_{p\sigma}\mathrm{a}_{p\sigma}$ is the occupation number operator for mode $\phi_{p\sigma}$, j is the hopping strength between two neighbouring sites, U is the on-site Coulomb interaction and μ is a chemical potential strength. The Fermi-Hubbard hamiltonian is one of the most studied systems in condensed matter physics, serving as a relatively simple model for describing solids. It constrasts with typical quantum chemistry descriptions in that its parameters (j, U and μ) are site-independent. We consider instances of a slightly generalised version of this model in chapter 3.

2.4.2 The challenge of fermionic state preparation and energy estimation

The main objective in this thesis is to provide quantum algorithmic techniques for constructing eigenstates and estimating eigenenergies of fermionic many-body hamiltonians described in the last section. Of particular interest are the lowest-energy states, or *ground states*, since most of the interesting physics of these systems tend to be dictated by their low-energy properties. For example, in molecular systems, exciting electrons to higher-energy states can be done by heating up the molecule, but this requires temperatures of thousands of degrees Kelvin above room temperature; smaller energy injections instead excite the nuclear states. Another way to excite electrons is by shining light on them; photons from a sufficiently hot body (such as the sun) carry enough energy to lift an electron to a higher-energy state. In practical situations however, these such photons typically reach only a small minority of the molecules around us. As such, we expect to find most molecules in their electronic ground states at room temperature. Knowledge about this electronic ground state then allows one to model, for instance, equilibrium geometries and reaction rates. Excited states, on the other hand, are crucial in fields such as photochemistry and spectroscopy.

Computing many-body eigenstates is, however, a difficult computational problem. To be precise, the problem of determining eigenenergies of local many-body hamiltonians to inverse-polynomial precision (in the system size) was proven to be QMA-complete by Kempe et al. [KKR06]. QMA is the quantum analogue of the classical complexity class NP, in the sense that it comprises all computational

problems whose solution (also known as "witness" or "proof certificate") can be efficiently verified by a quantum computer. Since a quantum computer can efficiently extract the energy from an eigenstate (see section 2.4.4), the corresponding eigenstate is a valid witness for this problem, meaning that preparing an eigenstate is at least QMA-hard as well. While the result of Kempe et al. applies to generic local many-body hamiltonians, finding eigenenergies of electronic structure hamiltonians has been shown to be QMA-complete as well [OIWF22]. As a result, it is expected that no efficient (i.e. polynomial-time) classical or quantum algorithm exists that can perform this task in the worst case. Nevertheless, in practice, certain systems may be considerably less difficult to describe due to additional structure.

2.4.3 Classical methods

Over the past decades, numerous classical techniques for many-body ground states have been developed, and we discuss a few of them here.

In chapter 3, we will consider approximate ground states constructed by the *mean-field approximation*. The idea here is to treat the system as if each electron in the system is assumed to move independently in the average electrostatic field produced by the other electrons. The eigenstates of such an approximate system can thus be expressed as a free-particle state. When we are interested in approximating the ground state of an electronic system, we should look for the free-particle state with the lowest possible energy: by the variational principle, the energy $\langle\psi|\mathrm{H}|\psi\rangle$ of a state $|\psi\rangle$ is lower bounded by the ground state energy E_0 of H, so the closer an approximate ground state's energy gets to E_0, the higher we expect the overlap with $|E_0\rangle$ to be. The degrees of freedom, which can be used to minimise the free-particle energy, lie in the definitions of the modes ϕ_P: rotating the modes $\phi_P \mapsto \sum_Q u_{PQ}\phi_Q$ allows one to find the lowest-energy modes which, when occupied by the present electrons, produce a minimal-energy free-particle state. This approach is known as the *Hartree-Fock* method, and the minimal-energy free-particle state is often referred to as the Hartree-Fock state. Sometimes, as in chapter 3, the minimal-energy free-particle state can be found analytically by choosing a proper ansatz; in general however, iterative computational methods such as the self-consistent field (SCF) [Har28, Sla28, Gau28] algorithm must be used. Mean-field results can be augmented with perturbation theory methods [MP34].

Tensor networks represent high-dimensional quantum states $|\psi\rangle$ using a network of low-dimensional tensors. Each node (or vertex) in the network represents a tensor, and the edges represent the contraction of shared indices between tensors, capturing correlations and entanglement between different parts of the system. The state $|\psi\rangle$ is parameterized by the tensors in the tensor network. The energy expectation value is then minimized by iteratively optimizing these tensors. Once the tensors are optimized, the computation of physical observables,

such as energy, requires contracting the tensor network. This contraction scales exponentially with system size in the general case, but can be made efficient by exploiting the network structure. For example, for MPS (matrix product states), a tensor network representation of one-dimensional systems, the contraction is linear in system size. For 2D systems using PEPS (projected entangled pair states), contraction is more challenging and typically requires approximation techniques. Tensor networks can efficiently represent quantum states with low to moderate entanglement; since ground states of local hamiltonians often satisfy an area law for entanglement entropy [ECP10] (meaning the entanglement grows only with the boundary size of a subsystem instead of with its volume), this makes tensor networks a useful technique for preparing many-body ground states.

Density functional theory (DFT) is an approach that reformulates the complex many-body problem of interacting electrons by using the electron density as the primary variable rather than the many-body wave function. The foundation of DFT rests on two theorems established by Hohenberg and Kohn in 1964, which state that (i) the ground state properties of a many-electron system are uniquely determined by its ground state electron density $\rho_0(\mathbf{r})$, and (ii) for a given hamiltonian, there exists a unique energy functional $E[\rho]$ which is minimised by ρ_0 [HK64]. Note that ρ is a function of three coordinates (instead of three times the number of particles) and as such significantly compresses the information for describing ground states. In practice, the exact functional for the total energy in terms of the electron density $E[\rho]$ is unknown. To address this, Kohn and Sham introduced a computational scheme [KS65] that reformulates the problem into a system of non-interacting electrons that generate the same electron density as the interacting system. In this formulation, a functional known as the the exchange-correlation energy encapsulates all the many-body effects missing from the non-interacting picture, including the quantum mechanical exchange interaction and electron correlation effects. This exchange-correlation functional is the only unknown part and must therefore be approximated; the nature of this approximation is then central to the accuracy of DFT. In a sense, the exchange-correlation functional is both a blessing and a curse: on the one hand, the use of approximations which capture strong correlation generally comes at an exponentially large computational cost; on the other hand, this functional is responsible for only a small part of the electronic energy, and the Kohn-Sham formulation allows for a much more precise description of kinetic energy. In the end, the compact electron density representation makes DFT suitable for large systems, although modelling strongly correlated systems remains difficult.

Lastly, *quantum Monte Carlo* (QMC) methods are a class of stochastic approaches used to address many-body quantum problems by leveraging statistical sampling techniques. In QMC, expectation values of observables with respect to a wave function are expressed in terms of integrals which are evaluated by Monte Carlo integration. Sufficiently good guesses for the ground state can be generated by, among others, variational means and imaginary time evolution techniques

[Cep10]. A major challenge for QMC methods applied to fermionic systems is the sign problem, where negative or complex weights cause destructive interference between different configurations in the Monte Carlo sampling [PM24]. This causes the variance of the sampled quantities to grow exponentially with system size or inverse temperature, which can makes accurate calculations prohibitively expensive. However, proposed techniques for mitigating this issue and the existence of sign-problem-free interacting fermionic systems [WZ05, LJY15, ADP16] renders QMC a useful technique for quantum chemistry and condensed matter theory.

2.4.4 Quantum phase estimation, the adiabatic approach and other quantum methods

While classical methods have been applied with great success in practical settings, modelling *strongly correlated* systems, in which ground states are highly entangled due to a high degree of electron interaction, remains a challenge for said methods. Quantum methods promise to advance our knowledge of strongly correlated systems by providing new techniques that harness the quantum nature of such systems. In particular, unlike in classical methods, a highly entangled state can in principle be represented by a small number of qubits (polynomial in the system size), from which properties such as energy can be extracted efficiently. Furthermore, such states can be fed into hamiltonian simulation routines which are generally considered to simulate quantum dynamics exponentially more efficiently than the best known classical algorithms.

The second-quantised formulation is especially useful for simulating many-body systems on a quantum computer. Hamiltonians in this form can be neatly deployed on a quantum computer when we use computational basis states as occupation states, where each qubit corresponds to a spin mode. The creation and annihilation operators can then be represented as

$$\mathrm{a}_P \mapsto \bigotimes_{i=1}^{P-1} \mathrm{Z}_j \otimes \frac{\mathrm{X}_P - \mathrm{i}\mathrm{Y}_P}{2}; \qquad \mathrm{a}_P^\dagger \mapsto \bigotimes_{i=1}^{P-1} \mathrm{Z}_j \otimes \frac{\mathrm{X}_P + \mathrm{i}\mathrm{Y}_P}{2}. \tag{2.47}$$

This encoding is the most elementary qubit encoding of second-quantised operators, and is known as the Jordan-Wigner transformation. Notice that the Jordan-Wigner encoding naturally produces an LCU representation of the hamiltonian (cf. section 2.3.3). Other encodings exist, most notably the Bravyi-Kitaev encoding [BK02, SRL12], which reduce the locality.

Amidst all quantum computing techniques, *quantum phase estimation* (QPE) is an important tool for estimating energies by allowing for the extraction of eigenvalues from a given quantum state [Kit96] (see also ref. dW23, chapter 4). Given a unitary U, an eigenstate ψ of U with eigenvalue $\mathrm{e}^{2\pi\mathrm{i}\varphi}$ and m ancillary qubits, QPE outputs the phase φ to m bits of precision. The algorithm works by creating a uniform superposition $2^{-m/2}\sum_{j=0}^{2^m-1}|j\rangle|\psi\rangle$ and applying the map

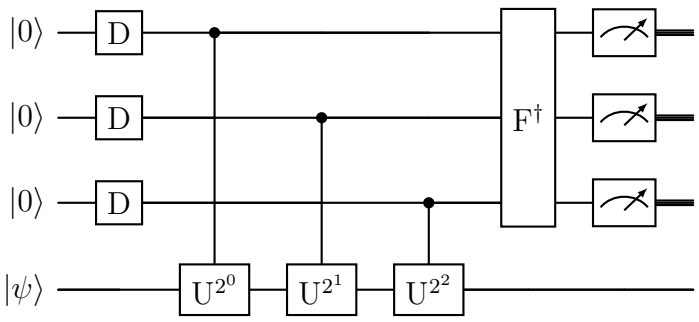


Figure 2.1. Quantum phase estimation circuit with $m = 3$ ancilla qubits. The first layer of Hadamard gates $\mathrm{D}^{\otimes m}$ creates a uniform superposition, and $\mathrm{F}^\dagger$ applies the inverse Fourier transform.

$|j\rangle|\psi\rangle \mapsto |j\rangle \mathrm{U}^j|\psi\rangle = |j\rangle \mathrm{e}^{2\pi \mathrm{i} j\varphi}|\psi\rangle$. This map can be implemented by applying ${}^C\mathrm{U}^{2^{k-1}}$, controlled on the k^{th} ancillary qubit, for $k = 1, \cdots, m$. The ancillary portion of the resulting state $2^{-m/2}\sum_j |j\rangle \mathrm{e}^{2\pi \mathrm{i} j\varphi}|\psi\rangle$ is then the Fourier transform of the state $|2^m\varphi\rangle$, which encodes an m-bit integer $2^m\varphi$ in m qubits. By applying an inverse Fourier transform and measuring the ancillary qubits, we obtain φ. The circuit is given in figure 2.1 for $m = 3$. In order to find the eigenenergy E_n corresponding to the eigenstate E_n of a hamiltonian H, we simply insert $\mathrm{U} = \mathrm{e}^{-\mathrm{i}\mathrm{H}t}$ for some small $t < \pi\|\mathrm{H}\|$ with a hamiltonian simulation technique discussed before.

The complexity of using QPE for energy estimations, however, arises from the requirement of supplying an eigenstate of the hamiltonian, since exact eigenstates are generally not available. If instead we supply a state $|\psi\rangle = \delta|E_n\rangle + \sqrt{1-\delta^2}|E_n^\perp\rangle$, the final measurement will collapse onto the correct output with probability δ^2; improvement of the energy estimation to a constant acurracy thus requires an number of repetitions that scales inverse-polynomially in δ. Without any knowledge about the system, one could always supply a random state, but it will on average have an overlap $1-\delta^2 \approx 1/d$ with $|E_n\rangle$, where d is the dimension of the fermionic Hilbert space. A better way is to apply state preparation methods to produce an input state with higher ground state overlap.

To an extent, classical methods can provide such input states: for example, a Hartree-Fock state can quickly be fed to a quantum device as a single computational basis state. However, higher overlaps can be achieved by employing adiabatic evolution or other quantum state preparation techniques. In particular, mean-field states are suitable candidates as starting points for adiabatic evolution. As long as the Hartree-Fock modes are orthogonalised, a complete orthogonal basis of the many-body Hilbert space is formed by all available free-particle states. In the mean-field basis, the mean-field hamiltonian is a diagonal operator, and can serve as the initial hamiltonian. By slowly evolving the mean-field hamiltonian towards the full interacting hamiltonian, we expect to obtain a state which has

an even larger overlap with the true ground state. This idea is explored in detail in chapter 3.

A different approach that has enjoyed popularity in recent years is the *variational quantum eigensolver* (VQE). VQE is a hybrid quantum-classical algorithm that iteratively optimizes a state $|\psi(\boldsymbol{\theta})\rangle = \mathrm{U}(\boldsymbol{\theta})|0\rangle$ produced by a parametrized quantum circuit to minimize the expectation value of the hamiltonian, thereby approximating the ground state. VQE has drawn most of its interest from being well suited for near-term quantum devices: it does not require fault-tolerant hardware (and has been claimed to be inherently noise-resilient thanks to the use of optimisation procedures) and can be deployed on small devices, thus facilitating experimental realisation with currently available setups. Nevertheless, it faces challenges of barren plateaus in large systems [LTW+24] and difficult implementation on fault-tolerant devices, which calls into question its long-term value.

Many other quantum techniques have been proposed, and we mention some of them here. *Thermal state preparation* techniques aim to prepare thermal states of the form $\rho_\beta = \mathrm{e}^{-\beta \mathrm{H}} / \operatorname{tr}[\mathrm{e}^{-\beta \mathrm{H}}]$ where β is an inverse temperature. In the zero-temperature limit ($\beta \to \infty$), ρ_β equals the ground state $|E_0\rangle\langle E_0|$. This can be powered by e.g. imaginary time evolution [MST+19] or Gibbs sampling [CKBG23], which effectively couple the system to a thermal bath. The *dissipative quantum eigensolver* (DQE) exploits different sampling style based on weak measurements [Cub23]. By repeatedly applying weak measurements that act as ground-state projections of the local ground space of each term, the initial state is slowly pushed into the direction of the ground state. Like ASP, DQE provably converges to the ground state in the infinite-time limit. Lastly, more generally applicable algorithms for estimating top eigenvectors of hermitian matrices (e.g. ref. CGdW24), which provide a polynomial advantage compared to classical algorithms in terms of the time complexity dependence on the Hilbert space dimension, can be applied to ground state preparation as well.

Chapter 3

Adiabatic ground state preparation of fermionic many-body systems from a two-body perspective

3.1 Introduction

Quantum computers are currently regarded as a prime candidate for solving problems in condensed matter physics and chemistry that are untractable for classical computers. In particular, since Feynman's observation of the potential of quantum simulation [Fey82], the pioneering work by Lloyd [Llo96] and the invention of quantum phase estimation [Kit96], interest in the deployment of quantum computers as simulators of highly correlated quantum systems has exploded. A large body of work has been established describing techniques for simulating dynamics of many-body systems on a quantum computer [WBAG11, LC19c, BWM+18, BGB+18, KGB+20, LBG+21], and these may be combined with quantum phase estimation in order to estimate eigenenergies [vBLH+21]. A critical question however, to make these methods useful, is how to prepare the states of interest – be it thermal states or eigenstates of the system under investigation – that serve as input to the algorithms that simulate dynamics or compute energies. Although experimental efforts using heuristics such as variational quantum eigensolvers [CMMS20, MS20, WLL20, TCC+22] have shown great success in preparing such states for systems of fixed size, much remains unknown with regards to "solving" highly correlated systems in general.

This chapter focusses on ground state preparation of complex fermionic many-body systems by adiabatic means. Typically, this entails preparing the system in the ground state of a simple hamiltonian H^{i}, which approximates the many-body hamiltonian H^{f}, and defining a time-dependent hamiltonian $\mathrm{H}(\lambda(t))$ such that $\mathrm{H}(\lambda(t=0)) = \mathrm{H}^{\mathrm{i}}$ and $\mathrm{H}(\lambda(t=T)) = \mathrm{H}^{\mathrm{f}}$ for some time T (cf. sec. 2.3). Theorem 2.3.1 then establishes the sufficiency condition $T \in \Omega(\Delta^{-3}\|\partial_\lambda \mathrm{H}\|^2_{\infty,2}\epsilon^{-1})$ to prepare the ground state with fidelity ϵ^2, where Δ is the ground state gap.

In principle, any adiabatic evolution may be implemented on a gate-based quantum computer with time-dependent hamiltonian simulation methods [Suz92, CST+21b, WK22, LC19a, KSB19b]. Alternative approaches approximate the evolution through a series of measurements [ATS03, LSR21] or simulations thereof [BKS09, BKS10].

The most commonly used schedule in adiabatic state preparation is a direct linear interpolation between H^{i} and H^{f} (cf. eq. 2.28), i.e.

$$\mathrm{H}(\lambda) = \mathrm{H}^{\mathrm{i}} + \lambda(\mathrm{H}^{\mathrm{f}} - \mathrm{H}^{\mathrm{i}}) \tag{3.1}$$

where $\lambda(0) = 0$ and $\lambda(T) = 1$. However, this method is rather restrictive as the evolution is controlled by only a single parameter, λ. Thus the evolution is sensitive to gap closures along the path, which cannot be avoided. An obvious solution is to increase the number of control parameters in the passage from H^{i} to H^{f}. This approach is discussed by Tomka et al. [TSRP16], who show that evolving along a geodesic path, based on the quantum metric tensor (or Fubini-Study metric) with respect to the control parameters, maximises the local fidelity along the path. In addition, they show that an increase in the number of control parameters leads to higher final fidelities. Put simply, their results rely on the fact that geodesic paths "walk around" regions of parameter space associated with small energy gaps, thus minimising diabatic errors. The problem with these approaches, however, is that their implementation becomes infeasible for large, complex systems. For the geodesic approach, the main roadblock is the inability to solve the geodesic equations, which become inaccesibly large systems of differential equations already for small many-body problems.

In this chapter, we introduce a more hands-on approach to produce new types of adiabatic paths for generic fermionic many-body hamiltonians in a second-quantised representation. Section 3.2 shows a trick for absorbing the noninteracting part into the interacting part, on which the rest of the chapter is built. The adiabatic paths we propose are based on a decomposition of the coefficient tensor of the hamiltonian (section 3.3), which defines a set of control parameters that govern the adiabatic evolution. We emphasise that such adiabatic paths can be seen as a new view on adiabatic state preparation for fermionic systems by considering many-body hamiltonians in terms of their two-body eigenstates. We demonstrate, through a set of worked examples (section 3.4), that there exist scenarios in which direct interpolation suffers from level crossings caused by (discrete) symmetries, and how the two-body decomposition may be used to explicitly break symmetries and lift crossings almost surely without making any other assumptions or choices in the design of the interpolation path. In section 3.5, we show how a description of these two-body eigenstates as superpositions of fermion pairs, following a suitable one-body transformation, leads to a worst-case adiabatic complexity in terms of the number of one-body modes L and a minimum gap Δ (section 3.5). The implications of this analysis are discussed for different systems. We summarise and conclude in section 3.6.

3.2 Two-body hamiltonian representations

Throughout this chapter, we will work with a fixed particle number (denoted N) for each hamiltonian. In a sector of fixed N, we can absorb the one-body terms of any second-quantised hamiltonian of the form in eq. 2.43 into the two-body terms, by inserting the identity as

$$\mathrm{a}_P^\dagger \mathrm{a}_Q = \frac{1}{N-1} \sum_R \mathrm{a}_P^\dagger \mathrm{a}_R^\dagger \mathrm{a}_R \mathrm{a}_Q \tag{3.2}$$

and when we define

$$w_{PQRS} := \frac{h_{PQ}\delta_{RS} + \delta_{PQ} h_{RS}}{N-1} \tag{3.3}$$

then we may write the one-body operator as

$$\sum_{PQ} h_{PQ}\, \mathrm{a}_P^\dagger \mathrm{a}_Q = \frac{1}{2} \sum_{PQRS} w_{PQRS}\, \mathrm{a}_P^\dagger \mathrm{a}_R^\dagger \mathrm{a}_S \mathrm{a}_Q. \tag{3.4}$$

Now define the combined one-body and two-body interaction tensor

$$G_{PQRS} := \frac{1}{2}(w_{PQRS} + g_{PQRS}) \tag{3.5}$$

and observe that

$$\mathrm{H} = \sum_{PQRS} G_{PQRS}\, \mathrm{a}_P^\dagger \mathrm{a}_R^\dagger \mathrm{a}_S \mathrm{a}_Q; \tag{3.6}$$

the one- and two-electron terms have now been combined into a single term. We shall refer to G as the interaction tensor of H.

Lastly, it will be convenient to express H as a sum over only the unique pairs (P, R) and (Q, S): by invoking the fermionic anticommutation relations, we may write

$$\mathrm{H} = \sum_{\substack{P<R \\ Q<S}} \tilde{G}_{PQRS}\, \mathrm{a}_P^\dagger \mathrm{a}_R^\dagger \mathrm{a}_S \mathrm{a}_Q \tag{3.7}$$

with $\tilde{G}_{PQRS} := G_{PQRS} - G_{PSRQ} - G_{RQPS} + G_{RSPQ} = 2(G_{PQRS} - G_{PSRQ})$. This tensor shall be termed the antisymmetrised interaction tensor of H.

Naturally, if the initial hamiltonian H^{i} contains all one-body terms, then there is no need to explicitly include them in the two-body terms of the residual hamiltonian H^{r} as described above. In section 3.4 we will see an example of this where H^{i} takes the form of a mean-field (also known as Hartree-Fock) approximation.

3.3 Adiabatic state preparation with two-body eigendecompositions

In this chapter, we deviate from the direct interpolation approach (eq. 3.1) by decomposing the residual hamiltonian $\mathrm{H}^{\mathrm{r}} = \mathrm{H}^{\mathrm{f}} - \mathrm{H}^{\mathrm{i}}$ into a sum of terms $\mathrm{H}^{\mathrm{r}}_k$, $k \in \{1, \ldots, M\}$, and evolving a linear combination of the terms $\mathrm{H}^{\mathrm{r}}_k$. That is,

$$\mathrm{H}(\lambda) = \mathrm{H}^{\mathrm{i}} + \sum_{k=1}^{M} \gamma_k(\lambda)\, \mathrm{H}^{\mathrm{r}}_k. \tag{3.8}$$

The decomposition defines an M-dimensional parameter space in which we consider paths $\boldsymbol{\gamma}(\lambda)$ restricted to a hypercube, starting from $\boldsymbol{\gamma}(0) = [0, 0, \ldots, 0]$ and ending at $\boldsymbol{\gamma}(1) = [1, 1, \ldots, 1]$.

We define the residual terms $\mathrm{H}^{\mathrm{r}}_k$ through a decomposition of the antisymmetrised interaction tensor of H^{r}. This is akin to known low-rank factorisation methods of the two-body part of the interaction tensor, aimed at achieving improved memory efficiency and speed-ups in quantum simulation implementations [RLB22, LBG+21]. Whereas most of these works focus on the Cholesky decomposition of the interaction tensor [KMP03, NGL21, RJ08], we use an eigendecomposition. More precisely, we regard the antisymmetrised interaction tensor $\tilde{\mathrm{G}}$ as an $L(L-1)/2 \times L(L-1)/2$ matrix F, by combining the indices (PR) and (QS):

$$F_{(PR)(QS)} := \tilde{G}_{PQRS}. \tag{3.9}$$

We point out that F represents a two-particle hamiltonian, since in the two-particle sector, the operator string $\mathrm{a}^\dagger_P \mathrm{a}^\dagger_R \mathrm{a}_S \mathrm{a}_Q$ is equivalent to the outer product $|PR\rangle\langle QS|$. For this reason, we refer to F as the two-particle matrix. Note however that the noninteracting part of F is scaled by a factor $1/(N-1)$ with respect to that of H^{r}, which arises from the insertion of identity (eq. 3.2).

Now since F is symmetric with respect to the exchange $(PR) \leftrightarrow (QS)$, we may eigendecompose F into (normalised) orthogonal eigenvectors,

$$F_{(PR)(QS)} = \sum_k \nu_k\, \phi^{(PR)}_k \phi^{(QS)}_k \tag{3.10}$$

and define

$$\mathrm{H}^{\mathrm{r}}_k := \nu_k \Big(\sum_{P<R} \phi^{(PR)}_k \mathrm{a}^\dagger_P \mathrm{a}^\dagger_R \Big) \Big(\sum_{Q<S} \phi^{(QS)}_k \mathrm{a}_S \mathrm{a}_Q \Big) =: \nu_k \Phi_k. \tag{3.11}$$

In the following, we will refer to ν_k as the two-body eigenvalues; the states $|\phi_k\rangle = \sum_{P<R} \phi^{(PR)} \mathrm{a}^\dagger_P \mathrm{a}^\dagger_R |\mathrm{vac}\rangle$ as the two-body eigenstates; and the operators Φ_k as the pseudoprojectors of F.

3.4 Lifting crossings by symmetry breaking in Fermi-Hubbard models

We will now illustrate how our method applies to cases where a discrete symmetry in the hamiltonian is influential on the course of the adiabatic evolution. In particular, we consider the situation in which the initial hamiltonian and the final hamiltonian share such a symmetry. In such a case, if the ground states of the initial and final hamiltonian belong to different symmetry sectors, then necessarily at some point the energy levels cross and an excited state is obtained at the end of the adiabatic evolution. This is a known problem that was addressed in the work by Farhi et al. [FGGS00] and also plays a role in many-body contexts [FZZ+22]. We note that, by the Von Neumann-Wigner theorem, the probability that a gap closure is encountered in a continuous set of hamiltonians parametrised by a single parameter, is extremely low unless the set conserves a symmetry [LL81, DK07]. As such, the only source of degeneracies we expect in practical settings is a symmetry in the problem instance. Typically, the solution is to add a symmetry-breaking field to the interpolation hamiltonian which is set to zero in the end. In this section, we show how symmetry-breaking behaviour emerges naturally from the formalism of the two-body eigendecomposition.

The origin of this symmetry breaking can be explained with the following theorem.

3.4.1. Theorem. *Let* U *be a unitary symmetry operator which is expressible as a product of one-body rotations. If and only if a two-particle eigenstate* $|\phi\rangle = \sum_{P<R} \phi_{(PR)} \mathrm{a}_P^\dagger \mathrm{a}_R^\dagger |\mathrm{vac}\rangle$ *is also an eigenstate of* U*, then the corresponding two-body pseudoprojector* Φ *commutes with* U*.*

Proof.
The "only if" direction is trivial, since $|\phi\rangle$ is an eigenstate of Φ. For the "if" direction, let μ be the eigenvalue of $|\phi\rangle$, and observe that

$$\begin{aligned}
\mu|\phi\rangle &= \mathrm{U}\Big(\sum_{P<R} \phi_{(PR)}\, \mathrm{a}_P^\dagger \mathrm{a}_R^\dagger |\mathrm{vac}\rangle\Big) \\
&= \sum_{P<R} \phi_{(PR)} \mathrm{U}\mathrm{a}_P^\dagger \mathrm{U}^\dagger \mathrm{U}\mathrm{a}_R^\dagger \mathrm{U}^\dagger |\mathrm{vac}\rangle \\
&= \sum_{P<R} \phi_{(PR)} \Big(\sum_M U_{PM} \mathrm{a}_M^\dagger\Big)\Big(\sum_N U_{RN} \mathrm{a}_N^\dagger\Big)|\mathrm{vac}\rangle \\
&= \sum_{M<N} \overbrace{\sum_{P<R} \phi_{(PR)} \big[U_{PM}U_{RN} - U_{PN}U_{RM}\big]}^{=\mu\,\phi_{(MN)}} \times \\
&\qquad \times \mathrm{a}_M^\dagger \mathrm{a}_N^\dagger |\mathrm{vac}\rangle. \qquad (3.12)
\end{aligned}$$

The last line and the fact that $|\mu| = 1$ then imply that

$$\mathrm{U}\Phi\mathrm{U}^\dagger = \mu\Phi\mu^* = \Phi \tag{3.13}$$

so that indeed $[\Phi, \mathrm{U}] = 0$. □

This theorem has the following implication: if all two-body eigenvalues ν_k are distinct, then all two-particle states $|\phi_k\rangle$, being eigenstates of the residual hamiltonian H^r, are also eigenstates of the symmetry operator U, and thus all hamiltonian terms H^r_k commute with U. In such a case, no symmetry is broken. However, if two two-body eigenstates with different symmetry are degenerate, then these two-body eigenstates may be mixed to produce new hamiltonian terms which do not commute with U and therefore break the symmetry. Note that for any discrete symmetry, only a finite number of mixings produce two-body eigenstates that respect the symmetry, and thus keep the symmetry of the interpolating hamiltonian intact; as such, one may simply pick a Haar random mixing to ensure that the symmetry of the interpolating hamiltonian is broken almost surely.

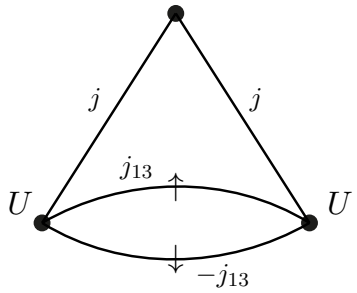
With these considerations in mind, we will focus on the following class of systems: (i) the initial and final hamiltonian share a (discrete) symmetry; (ii) the ground states of the initial and final hamiltonian belong to different symmetry sectors; (iii) the resulting residual hamiltonian has at least two degenerate two-body eigenstates in different symmetry sectors. For this class of systems, the two-body eigendecomposition method provides an out-of-the-box strategy for breaking the symmetry and opening up a gap along the path, without making any arbitrary choices regarding the symmetry-breaking field (apart from the choice of the two-body mixing, which can be randomised). We emphasise that in particular the third condition may be checked in a time scaling at most quadratically in L (the number of one-body modes).

We demonstrate this idea with two simple examples belonging to the class defined above. We will show that, taking a mean-field (Hartree-Fock) approximation as the initial hamiltonian, all three conditions are satisfied. In such a situation, straightforward adiabatic following of the Hartree-Fock state cannot avoid a crossing into an excited state. A multi-step adiabatic procedure, along the lines presented in this chapter, will avoid the crossing altogether and allow a convergence on the true ground state.

3.4.1 Fermi-Hubbard trimer

First, consider the following two-particle, three-site Fermi-Hubbard model:

H =



$$= \sum_{\sigma} j(\mathrm{a}^{\dagger}_{1\sigma}\mathrm{a}_{2\sigma} + \mathrm{a}^{\dagger}_{3\sigma}\mathrm{a}_{2\sigma}) + j_{\sigma}\,\mathrm{a}^{\dagger}_{1\sigma}\mathrm{a}_{3\sigma} + \mathrm{h.c.}$$
$$+ U(\mathrm{n}_{1\uparrow}\mathrm{n}_{1\downarrow} + \mathrm{n}_{3\uparrow}\mathrm{n}_{3\downarrow}) \tag{3.14}$$

where $\mathrm{n}_{i\sigma} = \mathrm{a}^{\dagger}_{i\sigma}\mathrm{a}_{i\sigma}$; j, $j_{\uparrow}$, $j_{\downarrow}$ and $U \leq 0$ are real constants; and we set $j_{\uparrow} = j_{13} = -j_{\downarrow}$. The discrete symmetry we will keep track of is the reflection of sites $1 \leftrightarrow 3$. We assume $j > 0$ throughout.

The case $U = 0$

Let us first consider $U = 0$. For $j_{13} = j$, the one-body kinetic energy terms for spin up have two degenerate ground states, at energy $E_{\mathrm{kin}\uparrow} = -j$. For $j_{13} < j$, the unique ground state is symmetric (S), while for $j_{13} > j$ the ground state is anti-symmetric (A). This makes clear that the ground state for one spin-up and one spin-down particle is symmetric for $j_{13} < j$ but antisymmetric for $j_{13} > j$. Turning on $U < 0$ will shift the S-A transition to lower values of j_{13}. The Hartree-Fock mean field solution follows this trend but we will see that there are values of j_{13} where Hartree-Fock places the ground state in the wrong symmetry sector.

The case $U < 0$: tracing the ground state of H

Turning on a negative U will change the nature of the two-body ground state.

For small $|U| \ll j$ and $j_{13} = j + \delta$ the energies of the symmetric (S) and anti-symmetric (A) states split as (in first order perturbation theory in U, δ)

$$E_{\mathrm{S}} = -3j + U/9 - \delta/3, \quad E_{\mathrm{A}} = -3j + U/3 - 5\delta/3 \tag{3.15}$$

implying that the S-A crossing (as a function of j_{13}) shifts to $j_{13} = j + U/6$, that is to a smaller value of j_{13}.

For U large and negative, the two electrons will tend to form a local pair at site 1 or 3, with energy U. In second order perturbation theory, taking into account processes with two hops (of strength j or j_{13}) connecting the pair states

with unpaired states at energy 0, the on-site energies of these pairs are adjusted to

$$\epsilon_1 = \epsilon_3 = U + 2j^2/U + 2j_{13}^2/U \tag{3.16}$$

while the pair hopping amplitude becomes

$$t_{13} = -2j_{13}^2/U. \tag{3.17}$$

This leads to S and A ground state energies

$$E_{\mathrm{S}}^{(2)} = U + 2j^2/U, \qquad E_{\mathrm{A}}^{(2)} = U + 2j^2/U + 4j_3^2/U. \tag{3.18}$$

Including terms of order j^4/U^3, $j^2 j_{13}^2/U^3$ and j_{13}^4/U^3 we find

$$E_{\mathrm{S}}^{(4)} = U + 2\frac{j^2}{U} + 8\frac{j^4}{U^3} + \frac{14}{3}\frac{j^2 j_{13}^2}{U^3} + O\Big(\frac{(j^2 + j_{13}^2)^3}{U^5}\Big), \tag{3.19}$$

$$E_{\mathrm{A}}^{(4)} = U + 2\frac{j^2}{U} + 4\frac{j_{13}^2}{U} - 4\frac{j^4}{U^3} + 2\frac{j^2 j_{13}^2}{U^3} - 16\frac{j_{13}^4}{U^3} + O\Big(\frac{(j^2 + j_{13}^2)^3}{U^5}\Big). \tag{3.20}$$

This puts the S-A crossing (in an expansion in terms of j/U) at $j_{13} = \sqrt{3}j^2/|U|$.

The case $U < 0$: Hartree-Fock approximation

To write a mean field (Hartree-Fock) ansatz, we should first decide on the symmetry sector. For an overall antisymmetric ansatz, we have

$$|\mathrm{HF, A}\rangle = \frac{1}{\sqrt{2}}(\mathrm{a}_{1\uparrow}^\dagger - \mathrm{a}_{3\uparrow}^\dagger)\frac{1}{\sqrt{2 + x^2}}(\mathrm{a}_{1\downarrow}^\dagger + x\mathrm{a}_{2\downarrow}^\dagger + \mathrm{a}_{3\downarrow}^\dagger)|\mathrm{vac}\rangle. \tag{3.21}$$

The expectation value becomes

$$\begin{aligned}\langle \mathrm{H}\rangle_{\mathrm{HF,A}} &= \frac{1}{2 + x^2}\left(4jx - 2j_{13} - (2 + x^2)j_{13} + U\right) \\ &= \frac{1}{2 + x^2}\left((U - 4j_{13}) + 4jx - j_{13}x^2\right).\end{aligned} \tag{3.22}$$

This expression is minimised for (keeping the leading terms in an expansion in terms of j/U, j_{13}/U)

$$\begin{aligned} x = 4j/U \quad &\Rightarrow \\ \langle \mathrm{H}\rangle_{\mathrm{HF,A}}^{\min} = \frac{U}{2} - 2j_{13} + 4\frac{j^2}{U} &+ O\Big(\frac{(j + j_{13})^3}{U^2}\Big).\end{aligned} \tag{3.23}$$

The competing ansatz is symmetric in both the up and the down factors,

$$|\mathrm{HF,S}\rangle = \frac{1}{\sqrt{2+y^2}}(\mathrm{a}^\dagger_{1\uparrow} + y\mathrm{a}^\dagger_{2\uparrow} + \mathrm{a}^\dagger_{3\downarrow}) \times \\ \times \frac{1}{\sqrt{2+x^2}}(\mathrm{a}^\dagger_{1\downarrow} + x\mathrm{a}^\dagger_{2\downarrow} + \mathrm{a}^\dagger_{3\downarrow})|\mathrm{vac}\rangle, \tag{3.24}$$

leading to

$$\begin{aligned}\langle \mathrm{H}\rangle_{\mathrm{HF,S}} &= \frac{1}{(2+x^2)(2+y^2)}\big[4jx(2+y^2) + 4jy(2+x^2) \\ &\quad - 2(2+y^2)j_{13} + 2(2+x^2)j_{13} + 2U\big] \\ &= \frac{1}{(2+x^2)(2+y^2)}\big[8j(x+y) + 4j(xy^2+yx^2) \\ &\quad + 2(x^2-y^2)j_{13} + 2U\big].\end{aligned} \tag{3.25}$$

In leading order, the minimum energy is reached for

$$\begin{aligned} x = y = 4j/U \quad &\Rightarrow \\ \langle \mathrm{H}\rangle^{\min}_{\mathrm{HF,S}} &= \frac{U}{2} + 8\frac{j^2}{U} + O\Big(\frac{j^4}{U^3}\Big).\end{aligned} \tag{3.26}$$

Comparing the expressions in the S and A sectors, we conclude that, in mean field and to leading order in j/U, the S-A crossing happens at $j_{13} = 2j^2/|U|$.

Adiabatic procedure

Suppose now that we consider the Fermi-Hubbard trimer with $j > 0$, $U < 0$, $|U| \gg j$ and $\sqrt{3}j^2/|U| < j_{13} < 2j^2/|U|$ and try to identify the ground state with a single spin-up and spin-down particle through adiabatic following. We have just demonstrated that in this situation, the HF solution is in the S sector, while the true ground state is in the A sector. This means that the adiabatic procedure will fail altogether and end up in a symmetric state, which is an excited state of H.

Let us now consider how the stepwise procedure works out in this example. The adiabatic procedure starts from the HF hamiltonian, with mean field parameters (called x, y in the above) optimised for our choice of U, j and j_{13}. It is obtained from H by replacing

$$\mathrm{n}_{i\uparrow}\mathrm{n}_{i\downarrow} \to \mathrm{n}_{i\uparrow}\langle \mathrm{n}_{i\downarrow}\rangle + \langle \mathrm{n}_{i\uparrow}\rangle \mathrm{n}_{i\downarrow} - \langle \mathrm{n}_{i\uparrow}\rangle\langle \mathrm{n}_{i\downarrow}\rangle. \tag{3.27}$$

This implies that the exact two-body hamiltonian $\mathrm{H}^{(2)}$ differs from the two-body HF hamiltonian via diagonal terms only, which directly correspond to the two-body eigenvalues ν_k of the stepwise adiabatic following from $\mathrm{H}^{(2)}$ to $\mathrm{H}^{(2)}_{\mathrm{HF}}$. The

system has two degenerate two-body ground states; the symmetry-adapted forms are

$$|\pm\rangle := \frac{1}{\sqrt{2}}(|1_\uparrow 1_\downarrow\rangle \pm |3_\uparrow 3_\downarrow\rangle) \tag{3.28}$$

and the corresponding eigenvalue is

$$\nu = \frac{U}{2} + \frac{32j^4}{U^3} + O\Big(\frac{j^6}{U^5}\Big). \tag{3.29}$$

As explained before, we can pick a Haar random mixing of these two two-body eigenstates to appropriately break the symmetry and allow an adiabatic passage, with unit probability. For concreteness, we pick the states $|1_\uparrow 1_\downarrow\rangle$ and $|3_\uparrow 3_\downarrow\rangle$, corresponding to the number operators at sites 1 and 3 respectively. We can then design the stepwise adiabatic procedure as follows. Defining

$$\begin{aligned} \mathrm{H}_1^{\mathrm{r}} &= \nu\, \mathrm{a}_{1\uparrow}^\dagger \mathrm{a}_{1\downarrow}^\dagger \mathrm{a}_{1\downarrow} \mathrm{a}_{1\uparrow} \\ \mathrm{H}_3^{\mathrm{r}} &= \nu\, \mathrm{a}_{3\uparrow}^\dagger \mathrm{a}_{3\downarrow}^\dagger \mathrm{a}_{3\downarrow} \mathrm{a}_{3\uparrow} \\ \mathrm{H}^{\mathrm{rest}} &= \mathrm{H} - \mathrm{H}_{\mathrm{HF}} - \mathrm{H}_1^{\mathrm{r}} - \mathrm{H}_3^{\mathrm{r}} \end{aligned} \tag{3.30}$$

we interpolate thus:

$$\begin{aligned} \mathrm{H}^{\mathrm{i}} &\to \mathrm{H}^{\mathrm{i}} + \mathrm{H}_1^{\mathrm{r}} \\ &\to \mathrm{H}^{\mathrm{i}} + \mathrm{H}_1^{\mathrm{r}} + \mathrm{H}^{\mathrm{rest}} \\ &\to \mathrm{H}^{\mathrm{i}} + \mathrm{H}_1^{\mathrm{r}} + \mathrm{H}^{\mathrm{rest}} + \mathrm{H}_3^{\mathrm{r}} = \mathrm{H}. \end{aligned} \tag{3.31}$$

With this, the discrete symmetry is broken along all steps of the path, and the S-A crossing, which derails the direct adiabatic interpolation from the HF to the exact hamiltonian, is avoided. The resulting development of the instantaneous gap can be seen in figure 3.1. Through numerical solutions to the Schrödinger equation, we obtain final infidelities with the exact ground state for different sweep times, as seen in figure 3.2.

3.4.2 Fermi-Hubbard model on four sites with alternating hopping

As a second example, we present a variation on the same theme: a simple model for correlated electrons where a mean field (Hartree Fock) solution is unable to correctly incorporate two-body correlations and as a result puts the ground state in the wrong symmetry sector, derailing adiabatic interpolation with the HF state as starting point. The stepwise adiabatic procedure based on two-body eigenspaces cures this situation.

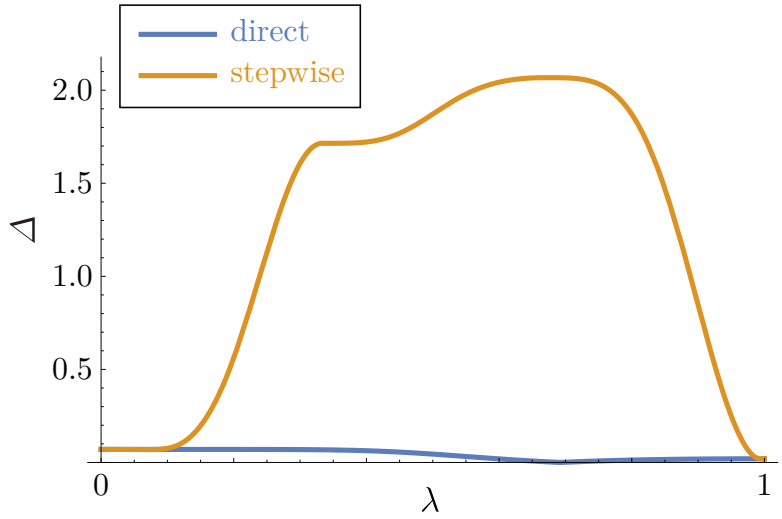


Figure 3.1. Instantaneous ground state energy gaps in the adiabatic ground state preparation of the Fermi-Hubbard trimer, with $U = -5$, $j = 1$ and $j_{13} = 0.37$. In the direct interpolation, $\mathrm{H}(\lambda) = (1-\lambda)\mathrm{H}_{\mathrm{HF}} + \lambda\mathrm{H}$, a gap closure occurs around $\lambda = 0.69$. The stepwise procedure is carried out as in eq. 3.31, with each step taking a third of the total time. Through symmetry breaking, a gap is visibly opened.

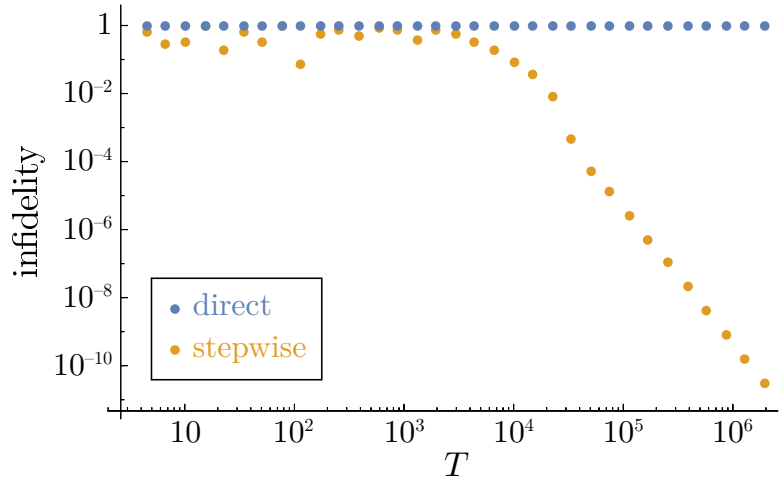
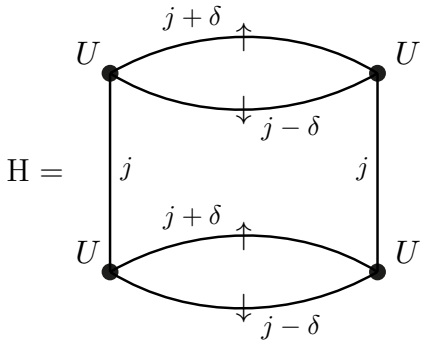


Figure 3.2. Dependence of final infidelity after an adiabatic sweep on total sweep time T for the Fermi-Hubbard trimer, with $U = -5$, $j = 1$ and $j_{13} = 0.37$. Clearly, direct interpolation produces a state orthogonal to the ground state (namely, in the wrong symmetry sector) regardless of T. On the other hand, the stepwise procedure allows for asymptotic convergence onto the exact ground state.

We consider a Fermi-Hubbard model on four sites,



$$= \sum_i \sum_\sigma j_{i\sigma}(\mathrm{a}^\dagger_{i\sigma}\mathrm{a}_{i+1,\sigma} + \mathrm{h.c.}) + U\sum_i \mathrm{n}_{i\uparrow}\mathrm{n}_{i\downarrow}, \tag{3.32}$$

with a uniform $U < 0$ but spin-dependent, non-uniform hoppings

$$j_{1\sigma} = j_{3\sigma} = j, \quad j_{2\uparrow} = j_{4\uparrow} = j + \delta, \quad j_{2\downarrow} = j_{4\downarrow} = j - \delta. \tag{3.33}$$

Note that we assume periodic boundary conditions, identifying site $i = 5$ with $i = 1$. We assume half filling, $N_\uparrow = N_\downarrow = 2$.

The hamiltonian H is invariant under a reflection $1 \leftrightarrow 4$, $2 \leftrightarrow 3$. Assuming $j > 0$ and $0 < \delta < j$, it is quickly found that for $U = 0$ the spin up particles have a symmetric (S) ground state, while the ground state for particles with spin down is antisymmetric (A). This renders the overall ground state antisymmetric (A).

Tracing the exact ground state

Turning on $U < 0$ will change the nature of the ground state, as the particles will tend to form two local pairs. For $|U| \gg j$, there is an effective description in terms of two such pairs with induced pair hopping of order j^2/U. It is quickly checked that this leads to a symmetric (S) ground state. The symmetry sector of the many-body ground state will thus change from A to S at a critical value $U_c(j, \delta) < 0$.

For a quick estimate of the cross-over point from A to S, we can follow, in (degenerate) first order perturbation theory in U and δ, in the lowest two energy eigenstates (here labeled by their symmetry A or S)

$$\langle \mathrm{H} \rangle_\mathrm{A} = -4j - 2\delta + U, \qquad \langle \mathrm{H} \rangle_\mathrm{S} = -4j + \frac{5}{4}U, \tag{3.34}$$

putting the A-S cross-over at $U_c = -8\delta$.

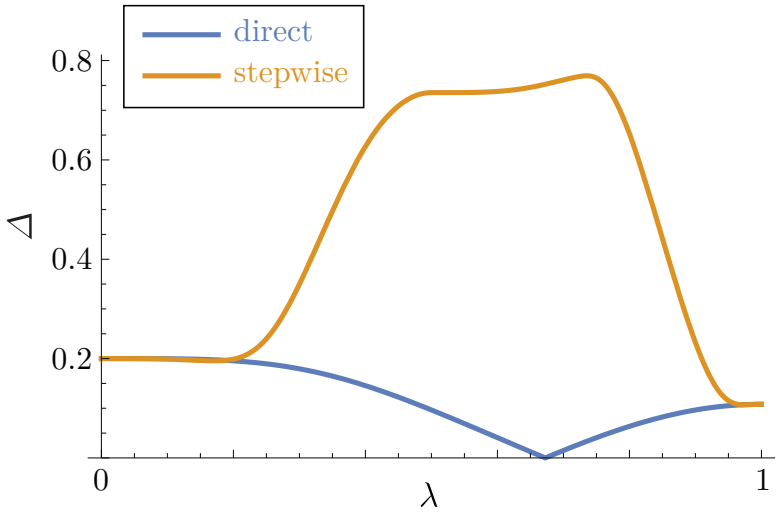


Figure 3.3. Instantaneous gaps in the direct and stepwise adiabatic ground state preparation of the four-site Fermi-Hubbard model with alternating spins, with $U = -2$, $j = 1$ and $\delta = 0.1$. The direct interpolation causes a level crossing around $\lambda = 0.67$. The stepwise interpolation is carried out by adding a single projector onto one of the sites in the first half, and the rest of $\mathrm{H} - \mathrm{H}_{\mathrm{HF}}$ in the second half. It is observed that the stepwise method avoids the level crossing.

Mean field (Hartree-Fock, HF) and adiabatic procedure

As in our previous example, a HF state is unable to accommodate the correlations induced by $U < 0$. In fact, since all four sites are on equal footing, self-consistent mean fields for both the up and down spins will be uniform on all four sites and will not affect the one-particle states that make up the HF ground state. The mean field ground state will thus be the same as that of the non-interacting problem, and it will be antisymmetric (A). This means that, for $|U|$ larger than $|U_c|$, direct adiabatic interpolation from the Hartree-Fock hamiltonian H_{HF} (into which we absorb the uniform mean-field energy shift) to H will result in an excited state of H. The stepwise procedure based on two-particle eigenstates of $\mathrm{H}_{\mathrm{HF}} - \mathrm{H}$ avoids this problem. It has four non-zero two-body eigenvalues ν_k connected to projectors on each of the sites. As in the three-site example, the two-body eigenstates corresponding to the number operators at these sites are symmetry-broken states. Adding one of these in the first step and the other three in the second step gives an adiabatic path that breaks the left-right symmetry and thereby avoids the crossing, allowing a correct interpolation from the antisymmetric HF state to the symmetric ground state of H. This can be seen in figure 3.3. We observe that in the same model with $U > 0$ the stepwise adiabatic following similarly avoids the problem of an exact A-S crossing, but in that case the gaps coming with the stepwise procedure are smaller and tend to decrease with δ. This is seen in figure 3.5. As with the trimer example, we

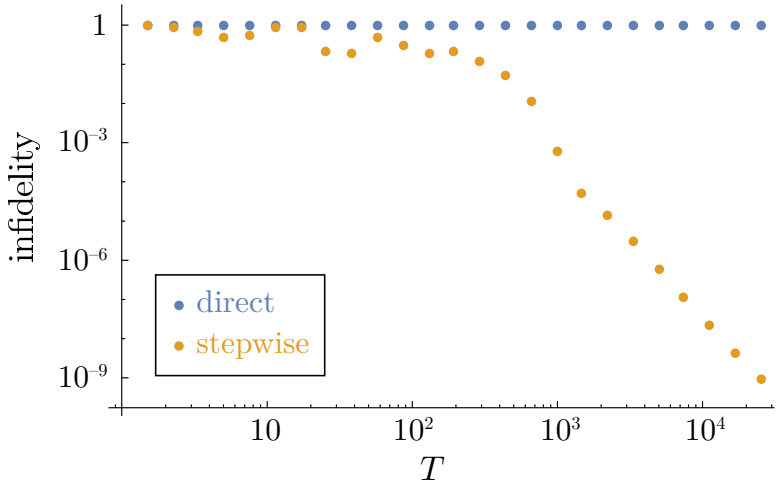


Figure 3.4. Dependence of final infidelity after an adiabatic sweep on total time T for the four-side Fermi-Hubbard model, with $U = -2$, $j = 1$ and $\delta = 0.1$. As expected, the state after direct interpolation remains orthogonal to the true ground state, whereas the stepwise procedure produces a state with fidelity that increases with T.

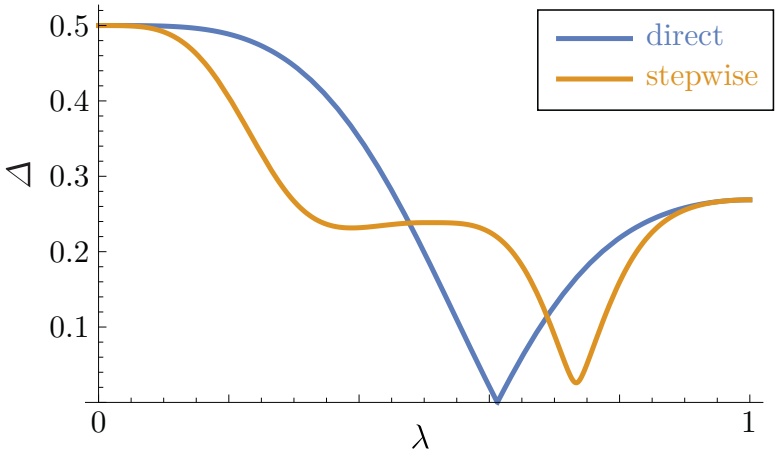


Figure 3.5. Instantaneous gaps for the four-site Fermi-Hubbard model, with positive U ($U = +2$, $j = 1$ and $\delta = 0.25$). In the same way as in figure 3.3, the direct interpolation causes a gap closure around $\lambda = 0.61$ while the stepwise procedure keeps the gap open. Note however a small gap in the stepwise procedure around $\lambda = 0.67$.

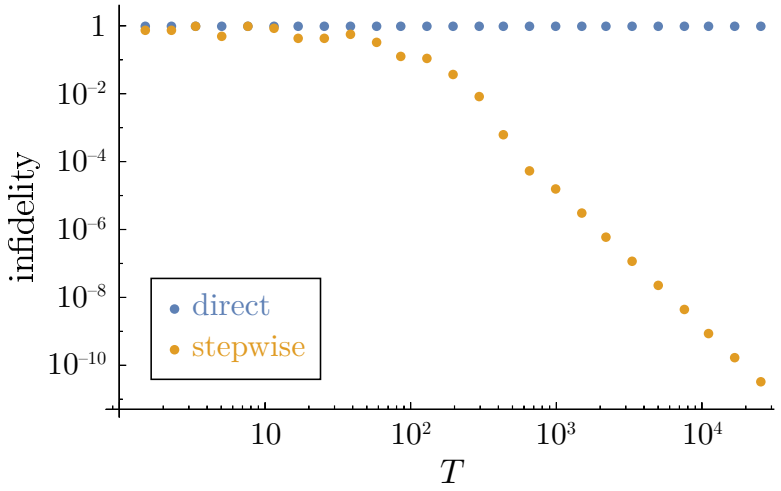


Figure 3.6. Relation between final infidelity on total time T for the four-side Fermi-Hubbard model, with $U = +2$, $j = 1$ and $\delta = 0.25$. Again, direct interpolation remains stuck at zero fidelity while the stepwise procedure approaches unit fidelity in the large time limit.

numerically integrate the Schrödinger equation to obtain fidelity profiles for direct and stepwise interpolation for both parameter settings. As seen in figures 3.4 and 3.6, direct interpolation fails to produce a state with any overlap with the true ground states, while stepwise interpolation succeeds in doing so.

3.5 Complexity considerations

Having seen the potential of a (partially) piecewise interpolation to lift gap closures, we will now study more generally the worst-case complexity of direct and piecewise paths, in view of the adiabatic complexity bound of eq. 2.31. For simplicity, we replace the gap $\Delta(\lambda)$ by its minimum $\Delta = \min_{\lambda\in[0,1]} \Delta(\lambda)$, and consider complexity in terms of this parameter. What remains then is to determine the scaling of the numerators

$$I_n := \int_0^1 \|\partial_\lambda^n \mathrm{H}(\sigma)\|^{2/n} d\lambda \quad (n \in \{1, 2\}) \tag{3.35}$$

in the system size. Since for typical systems of interest, the number of particles N scales proportionally with the number of one-particle modes L, we take L as the system size scaling parameter.

We focus on the extreme case where all terms from the two-body eigen-decomposition are interpolated separately, one by one. Let there be M such terms with nonzero eigenvalue. (Generally, M will be a function of L, and often, $M = L(L-1)/2$.) Furthermore, we assume that the path functions $\gamma_k(\lambda)$ are chosen "reasonably", meaning monotonically increasing, with vanishing

derivatives at the boundaries and bounded second derivatives. Now, from eqs. 3.8 and 3.11,

$$\begin{aligned}\|\partial_\lambda^n \mathrm{H}(\lambda)\| &\leq \sum_{k=1}^{M} |\partial_\lambda^n \gamma_k(\lambda)| \cdot |\nu_k| \cdot \|\Phi_k\| \\ &\leq \nu_{|\max|}\, \Phi_{\|\max\|} \sum_{k=1}^{M} |\partial_\lambda^n \gamma_k(\lambda)| \end{aligned} \tag{3.36}$$

where $\nu_{|\max|} := \max_k |\nu_k|$ and $\Phi_{\|\max\|} := \max_k \|\Phi_k\|$. Inserting this back into eq. 3.35 yields

$$\begin{aligned} I_1 &\leq \nu_{|\max|}^2\, \Phi_{\|\max\|}^2 \int_0^1 \mathrm{d}\lambda \left(\sum_{k=1}^{M} |\partial_\lambda \gamma_k(\lambda)| \right)^2 \\ &= \nu_{|\max|}^2\, \Phi_{\|\max\|}^2 \int_0^1 \mathrm{d}\lambda \, (\partial_\lambda \gamma_k(\lambda))^2 \end{aligned} \tag{3.37}$$

where we used that $\partial_\lambda \gamma_k(\lambda) \partial_\lambda \gamma_l(\lambda) = \delta_{kl} (\partial_\lambda \gamma_k(\lambda))^2$ for separately interpolated two-body terms. Each γ_k increases from 0 to 1 in a λ interval of size $1/M$, so that $|\partial_\lambda \gamma_k(\lambda)|^2 \in O(M^2)$ in this interval and is zero otherwise; hence

$$I_1 \in O(\nu_{|\max|}^2\, \Phi_{\|\max\|}^2 \cdot M \cdot \frac{1}{M} \cdot M^2) = O(\nu_{|\max|}^2\, \Phi_{\|\max\|}^2\, M^2). \tag{3.38}$$

At the same time,

$$I_2 \leq \nu_{|\max|}\, \Phi_{\|\max\|} \sum_{k=1}^{M} \int_0^1 \mathrm{d}\lambda \, |\partial_\lambda^2 \gamma_k(\lambda)|; \tag{3.39}$$

since $\int_0^1 \mathrm{d}\lambda \, |\partial_\lambda^2 \gamma_k(\lambda)| \in O(\max_{\lambda \in [0,1]} \partial_\lambda \gamma_k(\lambda)) = O(M)$ we find

$$I_2 \in O(\nu_{|\max|}\, \Phi_{\|\max\|} \cdot M \cdot \frac{1}{M} \cdot M) = O(\nu_{|\max|}\, \Phi_{\|\max\|}\, M). \tag{3.40}$$

Therefore, since we expect that $\nu_{|\max|}\, \Phi_{\|\max\|} \in \Omega(1)$, the first derivative numerator I_1 is dominant over I_2 in L scaling, and it suffices to consider only I_1. The two-body eigenvalues ν_k then, being the eigenvalues of F, are bounded by the energy scale of a two-particle system which does not grow with the system size. Therefore the maximum pseudoprojector norm $\Phi_{\|\max\|}$ is the only remaining meaningful quantity to be upper bounded.

We will universally upper bound the operator norm of any pseudoprojector Φ_k in the following, and henceforth drop the subscript k. Afterwards, we discuss some implications of this bound for different choices of paths and systems.

Upper bound to the pseudoprojector operator norm

Define b such that $\Phi = \mathrm{b}^\dagger \mathrm{b}$ for some operator Φ from the decomposition. When we define the $L \times L$ antisymmetric matrix $\tilde{\phi}$ with entries

$$\tilde{\phi}^{PR} := \phi^{(PR)} - \phi^{(RP)} \tag{3.41}$$

we may write

$$\mathrm{b}^\dagger = \sum_{P<R} \phi^{(PR)} \mathrm{a}_P^\dagger \mathrm{a}_R^\dagger = \frac{1}{2} \sum_{PR} \tilde{\phi}^{PR} \mathrm{a}_P^\dagger \mathrm{a}_R^\dagger. \tag{3.42}$$

Next, apply a Youla decomposition $\tilde{\phi} = \mathrm{V} \Xi \mathrm{V}^\top$ where V is an $L \times L$ orthogonal matrix and

$$\Xi = \bigoplus_{m=1}^{L/2} \begin{bmatrix} 0 & \xi_m \\ -\xi_m & 0 \end{bmatrix} \tag{3.43}$$

if L is even; if L is odd, Ξ has an additional row and column of zeros. This then yields

$$\mathrm{b}^\dagger = \sum_{m=1}^{\lfloor L/2 \rfloor} \xi_m \tilde{\mathrm{a}}_{2m-1}^\dagger \tilde{\mathrm{a}}_{2m}^\dagger \tag{3.44}$$

where we defined the rotated fermionic operators $\tilde{\mathrm{a}}_K^{(\dagger)} = \sum_P V_{PK} \mathrm{a}_P^{(\dagger)}$. Note that since the vector with entries $\phi^{(PR)}$ is a normalised eigenvector, the squares of ξ_m sum to unity.

Since $\|\Phi\| = \|\mathrm{b}\|^2 = \max_{|\Psi\rangle} \|\mathrm{b}|\Psi\rangle\|^2$, in order to obtain the spectral norm of Φ it is sufficient to find the state whose norm is maximised under the application of b. Now, from eq. 3.44 we observe that b defines a set of pairs $\{(2m-1, 2m)\}_{m=1}^{\lfloor L/2 \rfloor}$, and only annihilates particles from a product state $\prod_P \tilde{\mathrm{a}}_P^\dagger |\mathrm{vac}\rangle$ if they appear together in these pairs. As such, it makes sense to describe a product state in terms of its fermion pairs and its unpaired fermions. We shall denote a product state as a ket $|\mathcal{P}, \mathcal{U}\rangle$ where $\mathcal{P}$ is the set of filled pairs and $\mathcal{U}$ is the set of remaining unpaired fermions; in other words,

$$|\mathcal{P}, \mathcal{U}\rangle = \Bigg(\prod_{i \in \mathcal{U}}{}' \tilde{\mathrm{a}}_i^\dagger \Bigg) \Bigg(\prod_{m \in \mathcal{P}} \tilde{\mathrm{a}}_{2m-1}^\dagger \tilde{\mathrm{a}}_{2m}^\dagger \Bigg) |\mathrm{vac}\rangle \tag{3.45}$$

where the prime on the leftmost product symbol indicates that a certain order of the unpaired modes is assumed, in order to fix the sign of $|\mathcal{P}, \mathcal{U}\rangle$. In this notation, the matrix elements of Φ are given by

$$\langle \mathcal{P}, \mathcal{U} | \Phi | \mathcal{P}', \mathcal{U}' \rangle = \delta_{\mathcal{U}\mathcal{U}'} \begin{cases} \sum_{m \in \mathcal{P}} \xi_m^2 & \text{if } \mathcal{P} = \mathcal{P}' \\ \xi_m \xi_n & \text{if } \mathcal{P} \setminus \{m\} = \mathcal{P}' \setminus \{n\} \\ 0 & \text{otherwise.} \end{cases} \tag{3.46}$$

From eq. 3.46, it is clear that $\|\mathrm{b}|\mathcal{P},\mathcal{U}\rangle\|$ is maximised when $|\mathcal{P},\mathcal{U}\rangle$ lies in the sector with a minimal number of unpaired fermions (zero if N is even, one if odd). Furthermore, $\mathrm{b}^\dagger\mathrm{b}$ preserves the unpaired fermions, and therefore the state $|\Psi\rangle$ that maximises the norm must lie in this sector.

If N is even, all particles in this sector are paired up. Such paired fermions, then, are equivalent to what are known as *hardcore bosons* (HCBs): particles whose operator algebra commutes at different sites, but which may only singly occupy any given site. In this sense, the set $\{\mathrm{b}_m^\dagger|\mathrm{vac}\rangle\}_{m=1}^{\lfloor L/2\rfloor}$ with $\mathrm{b}_m^\dagger = \tilde{\mathrm{a}}_{2m-1}^\dagger\tilde{\mathrm{a}}_{2m}^\dagger$ may be viewed as a single-particle HCB basis, and $\mathrm{b}^\dagger$, to which we shall add a subscript $\mathrm{b}^\dagger = \mathrm{b}_{\boldsymbol{\xi}}^\dagger$, is then a rotated HCB creation operator. A universal upper bound on the spectral norm of a pseudoprojector Φ is now given by

$$\|\Phi\| \leq \max_{\|\boldsymbol{\xi}\|=1} \max_{|\Psi\rangle\in\mathcal{H}^{\mathrm{HCB}}_{\lfloor L/2\rfloor,N/2}} \langle\Psi|\mathrm{b}_{\boldsymbol{\xi}}^\dagger\mathrm{b}_{\boldsymbol{\xi}}|\Psi\rangle \tag{3.47}$$

where $\mathcal{H}^{\mathrm{HCB}}_{l,n}$ denotes an l-site, n-particle HCB Hilbert space. An expression for the right-hand side of eq. 3.47 was found by Tennie et al. [TVS17, theorem 1]; the upper bound that follows is

3.5.1. Theorem. *For even particle number N, a universal upper bound on the operator norm of a pseudoprojector Φ is given by*

$$\|\Phi\| \leq \frac{N/2}{\lfloor L/2\rfloor}(\lfloor L/2\rfloor - N/2 + 1). \tag{3.48}$$

For N odd, a similar result may be found through the observation that any eigenstate of Φ with maximum eigenvalue must lie in the sector where all fermions except one are paired up, for the same reason as discussed above. The problem of upper bounding the operator norm of $\|\Phi\|$ is then equivalent to the even case in a Hilbert space with one less HCB site available. This is formalised in the following theorem.

3.5.2. Theorem. *For odd N, $\|\Phi\|$ is upper bounded by*

$$\|\Phi\| \leq \frac{\lfloor N/2\rfloor}{\lfloor L/2\rfloor - 1}(\lfloor L/2\rfloor - \lfloor N/2\rfloor). \tag{3.49}$$

Proof.
The proof largely follows that of Tennie et al. [TVS17, theorem 1].

We seek to maximise the expectation value $\langle\Psi|\mathrm{b}_{\boldsymbol{\xi}}^\dagger\mathrm{b}_{\boldsymbol{\xi}}|\Psi\rangle$ with respect to $|\Psi\rangle$ and $\boldsymbol{\xi}$. The hard-core boson (HCB) operator $\mathrm{b}_{\boldsymbol{\xi}}$ is defined as $\mathrm{b}_{\boldsymbol{\xi}} = \sum_m \xi_m \mathrm{b}_m = \sum_m \xi_m \tilde{\mathrm{a}}_{2m}\tilde{\mathrm{a}}_{2m-1}$; in the following, we will use m both as an index of HCB sites and as a shorthand for the (equivalent) pair of fermionic sites $(2m-1, 2m)$. Additionally, we use the *flattening* symbol ϕ to convert between sets of pairs and sets of fermionic sites,

$$\phi(\{(\mu_1,\mu_2),\ldots\}) = \{\mu_1,\mu_2,\ldots\}. \tag{3.50}$$

In the subspace of maximally paired N-particle states (with a single fermion left unpaired), $|\Psi\rangle$ may be expanded as

$$|\Psi\rangle = \sum_{I,i} A_{I,i} |I, \{i\}\rangle \tag{3.51}$$

where I is a set of fermionic pairs and $|I, \{i\}\rangle = \tilde{\mathrm{a}}_i^\dagger \prod_{m\in I} \mathrm{b}_m^\dagger |\mathrm{vac}\rangle$, in line with eq. 3.45. Furthermore, we *define* $A_{I,i} = 0$ if $i \in \phi(I)$ or $|I| \neq \lfloor N/2 \rfloor$. The desired expectation value may then be expressed as

$$\begin{aligned}
\langle\Psi|\mathrm{b}_{\boldsymbol{\xi}}^\dagger \mathrm{b}_{\boldsymbol{\xi}}|\Psi\rangle &= \sum_{\substack{I,i\\J,j}} \sum_{mn} A_{I,i} A_{J,j}\, \xi_m \xi_n \,\langle I, \{i\}|\mathrm{b}_m^\dagger \mathrm{b}_n|J, \{j\}\rangle \\
&= \sum_{\substack{I,i\\J,j}} A_{I,i} A_{J,j} \sum_{\substack{m\in I\\n\in J}} \xi_m \xi_n\, \delta_{I\setminus\{m\},J\setminus\{n\}}\, \delta_{ij} \\
&= \sum_{I',i} \sum_{mn} A_{I'\cup\{m\},i}\, A_{I'\cup\{n\},i}\, \xi_m \xi_n \\
&= \sum_{I',i} \Big(\sum_m A_{I'\cup\{m\},i}\, \xi_m \Big)^2 .
\end{aligned} \tag{3.52}$$

In the third line, the I' indexes all pair sets of cardinality $\lfloor N/2 \rfloor - 1$.

The way to get to the desired upper bound of this expression is to insert the indicator functions $\mathbb{1}_{m\notin I'}$ and $\mathbb{1}_{i\notin\phi(I'\cup\{m\})}$ into the final line of eq. 3.52. While these indicator functions are already incorporated in the definition of the coefficients $A_{I,i}$ and hence may seem redundant, they allow for clever use of the Cauchy-Schwarz inequality in two different ways, which leads to a system of inequalities from which we can obtain the upper bound. The first of these is nearly identical to that presented by Tennie et al. [TVS17, appendix A, eq. A3], and is as follows,

$$\begin{aligned}
\langle\Psi|\mathrm{b}_{\boldsymbol{\xi}}^\dagger \mathrm{b}_{\boldsymbol{\xi}}|\Psi\rangle &= \sum_{I',i} \Big(\sum_m A_{I'\cup\{m\},i}\, \xi_m \mathbb{1}_{m\notin I'} \Big)^2 \\
&\leq \sum_{I',i} \sum_l A_{I'\cup\{l\},i}^2 \sum_k (\xi_k \mathbb{1}_{k\notin I'})^2 \\
&= \sum_{I',i} \sum_l A_{I'\cup\{l\},i}^2 \sum_{k\notin I'} \xi_k^2 \\
&= \sum_{I,i} \sum_{l\in I} A_{I,i}^2 \sum_{k\notin I\setminus\{l\}} \xi_k^2 \\
&= \sum_{I,i} A_{I,i}^2 \sum_{l\in I} \Big(\xi_l^2 + \sum_{k\notin I} \xi_k^2 \Big)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{I,i} A_{I,i}^2 \Big(\lfloor N/2 \rfloor \sum_{k \notin I} \xi_k^2 + \sum_{l \in I} \xi_l^2 \Big) \\
&= \sum_{I,i} A_{I,i}^2 \Big(\lfloor N/2 \rfloor \sum_{k \notin I} \xi_k^2 + 1 - \sum_{l \notin I} \xi_l^2 \Big) \\
&= 1 + (\lfloor N/2 \rfloor - 1) \sum_{I,i} A_{I,i}^2 \sum_{k \notin I} \xi_k^2 .
\end{aligned} \tag{3.53}
$$

In the penultimate line, we used the normalisation of $\boldsymbol{\xi}$, and in the last line that of $|\Psi\rangle$.

For the second way, it becomes important that the unpaired fermion takes away a site pair that could otherwise be occupied by a pair of fermions:

$$
\begin{aligned}
\langle \Psi | \mathrm{b}_{\boldsymbol{\xi}}^\dagger \mathrm{b}_{\boldsymbol{\xi}} | \Psi \rangle &= \sum_{I',i} \Big(\sum_m A_{I' \cup \{m\},i} \, \xi_m \mathbb{1}_{m \notin I'} \mathbb{1}_{i \notin \phi(I' \cup \{m\})} \Big)^2 \\
&\leq \sum_{I',i} \sum_l A_{I' \cup \{l\},i}^2 \, \xi_l^2 \sum_k \mathbb{1}_{k \notin I'} \mathbb{1}_{i \notin \phi(I' \cup \{k\})} \\
&= \sum_{I',i} \sum_l A_{I' \cup \{l\},i}^2 \, \xi_l^2 \, (\lfloor L/2 \rfloor - \lfloor N/2 \rfloor) \\
&= (\lfloor L/2 \rfloor - \lfloor N/2 \rfloor) \sum_{I,i} \sum_{l \in I} A_{I,i}^2 \, \xi_l^2 \\
&= (\lfloor L/2 \rfloor - \lfloor N/2 \rfloor) \Big(1 - \sum_{I,i} \sum_{l \notin I} A_{I,i}^2 \, \xi_l^2 \Big).
\end{aligned} \tag{3.54}
$$

Again, in the last line we used the normalisation of $\boldsymbol{\xi}$ and $|\Psi\rangle$.

When we take the appropriate linear combination of inequalities 3.53 and 3.54, the sum $\sum_{I,i} A_{I,i}^2 \sum_{k \notin I} \xi_k^2$ cancels out,

$$
\begin{aligned}
&\frac{\lfloor L/2 \rfloor - \lfloor N/2 \rfloor}{\lfloor N/2 \rfloor - 1} \langle \Psi | \mathrm{b}_{\boldsymbol{\xi}}^\dagger \mathrm{b}_{\boldsymbol{\xi}} | \Psi \rangle + \langle \Psi | \mathrm{b}_{\boldsymbol{\xi}}^\dagger \mathrm{b}_{\boldsymbol{\xi}} | \Psi \rangle \\
&\qquad \leq (\lfloor L/2 \rfloor - \lfloor N/2 \rfloor) \Big(1 + \frac{1}{N-1} \Big),
\end{aligned} \tag{3.55}
$$

and we find

$$
\langle \Psi | \mathrm{b}_{\boldsymbol{\xi}}^\dagger \mathrm{b}_{\boldsymbol{\xi}} | \Psi \rangle \leq \frac{\lfloor N/2 \rfloor (\lfloor L/2 \rfloor - \lfloor N/2 \rfloor)}{\lfloor L/2 \rfloor - 1} \tag{3.56}
$$

as desired. □

The bounds of theorems 3.5.1 and 3.5.2 are also tight.

3.5.3. Theorem. *Let $\ell = \lfloor L/2 \rfloor$. The upper bound in the even case, theorem 3.5.1, is saturated by taking $\xi_m = 1/\sqrt{\ell}$ $\forall m$, and taking $|\Psi\rangle$ to be the maximally*

symmetric state

$$|\Psi\rangle = \binom{\ell}{N/2}^{-1/2} \sum_{\mathcal{P}:|\mathcal{P}|=N/2} |\mathcal{P},\emptyset\rangle \tag{3.57}$$

where the sum runs over all sets $\mathcal{P}$ of $N/2$ HCB sites.

The upper bound of in the odd case, theorem 3.5.2, is saturated by $\xi_m = 1/\sqrt{\ell-1}\;\forall m \neq \ell$, $\xi_\ell = 0$ and

$$|\Psi\rangle = \binom{\ell-1}{\lfloor N/2\rfloor}^{-1/2} \sum_{\substack{\mathcal{P}:|\mathcal{P}|=\lfloor N/2\rfloor \\ \ell\notin\mathcal{P}}} |\mathcal{P},\{2\ell-1\}\rangle. \tag{3.58}$$

Proof.
For the even case, when we take $\xi_m = 1/\sqrt{\ell}\,\forall m$ and $|\Psi\rangle$ the maximally symmetric state

$$|\Psi\rangle = \binom{\ell}{N/2}^{-1/2} \sum_{\mathcal{P}:|\mathcal{P}|=N/2} |\mathcal{P},\emptyset\rangle \tag{3.59}$$

where $|\mathcal{P},\emptyset\rangle = \prod_{k\in\mathcal{P}} \mathrm{b}_k^\dagger|\mathrm{vac}\rangle$, then a straightforward calculation shows that

$$\begin{aligned}
\langle\Psi|\mathrm{b}_{\boldsymbol{\xi}}^\dagger \mathrm{b}_{\boldsymbol{\xi}}|\Psi\rangle &= \frac{1}{\ell}\binom{\ell}{N/2}^{-1} \sum_{\mathcal{P},\mathcal{P}'}\sum_{mn} \langle\mathrm{vac}|\Big(\prod_{k\in\mathcal{P}}\mathrm{b}_k\Big)\mathrm{b}_m^\dagger \mathrm{b}_n\Big(\prod_{l\in\mathcal{P}'}\mathrm{b}_l^\dagger\Big)|\mathrm{vac}\rangle \\
&= \frac{1}{\ell}\binom{\ell}{N/2}^{-1} \sum_{\mathcal{P},\mathcal{P}'}\sum_{\substack{m\in\mathcal{P}\\ n\in\mathcal{P}'}} \delta_{\mathcal{P}\setminus\{m\},\mathcal{P}'\setminus\{n\}} \\
&= \frac{1}{\ell}\binom{\ell}{N/2}^{-1}\binom{\ell}{N/2}(N/2)(\ell-N/2+1) \\
&= \frac{N/2}{\lfloor L/2\rfloor}(\lfloor L/2\rfloor - N/2+1).
\end{aligned} \tag{3.60}$$

In the odd case then, where $\xi_m = 1/\sqrt{\ell-1}\;\forall m \neq \ell$, $\xi_\ell = 0$ and

$$|\Psi\rangle = \binom{\ell-1}{\lfloor N/2\rfloor}^{-1/2} \sum_{\substack{\mathcal{P}:|\mathcal{P}|=\lfloor N/2\rfloor \\ \ell\notin\mathcal{P}}} |\mathcal{P},\{2\ell-1\}\rangle \tag{3.61}$$

where $|\mathcal{P},\{i\}\rangle = \tilde{\mathrm{a}}_i^\dagger \prod_{k\in\mathcal{P}} \mathrm{b}_k^\dagger|\mathrm{vac}\rangle$, we find

$$\begin{aligned}\langle\Psi|\mathrm{b}_{\boldsymbol{\xi}}^\dagger \mathrm{b}_{\boldsymbol{\xi}}|\Psi\rangle &= \frac{1}{\ell-1}\binom{\ell-1}{\lfloor N/2\rfloor}^{-1} \sum_{\substack{\mathcal{P},\mathcal{P}'\\ \ell\notin\mathcal{P},\mathcal{P}'}} \sum_{m,n\neq\ell} \langle\mathrm{vac}|\Big(\prod_{k\in\mathcal{P}}\mathrm{b}_k\Big)\tilde{\mathrm{a}}_{2\ell-1}\mathrm{b}_m^\dagger \\ &\qquad\qquad \times \mathrm{b}_n\tilde{\mathrm{a}}_{2\ell-1}^\dagger\Big(\prod_{l\in\mathcal{P}'}\mathrm{b}_l^\dagger\Big)|\mathrm{vac}\rangle \\ &= \frac{1}{\ell-1}\binom{\ell-1}{\lfloor N/2\rfloor}^{-1} \sum_{\substack{\mathcal{P},\mathcal{P}'\\ \ell\notin\mathcal{P},\mathcal{P}'}} \sum_{m,n\neq\ell} \langle\mathrm{vac}|\Big(\prod_{k\in\mathcal{P}}\mathrm{b}_k\Big)(\underbrace{\delta_{m\ell}\tilde{\mathrm{a}}_{2m}^\dagger}_{=0} + \mathrm{b}_m^\dagger\tilde{\mathrm{a}}_{2\ell-1}) \\ &\qquad\qquad \times (\underbrace{\delta_{n\ell}\tilde{\mathrm{a}}_{2n}}_{=0} + \tilde{\mathrm{a}}_{2\ell-1}^\dagger\mathrm{b}_n)\Big(\prod_{l\in\mathcal{P}'}\mathrm{b}_l^\dagger\Big)|\mathrm{vac}\rangle \\ &= \frac{1}{\ell-1}\binom{\ell-1}{\lfloor N/2\rfloor}^{-1} \sum_{\substack{\mathcal{P},\mathcal{P}'\\ \ell\notin\mathcal{P},\mathcal{P}'}} \sum_{m,n\neq\ell} \langle\mathrm{vac}|\Big(\prod_{k\in\mathcal{P}}\mathrm{b}_k\Big)\mathrm{b}_m^\dagger\mathrm{b}_n\Big(\prod_{l\in\mathcal{P}'}\mathrm{b}_l^\dagger\Big)|\mathrm{vac}\rangle \\ &= \frac{N/2}{\lfloor L/2\rfloor - 1}(\lfloor L/2\rfloor - N/2). \end{aligned} \tag{3.62}$$

In the last line, we directly used the result of eq. 3.60. □

We note that instead of the ℓ-th HCB site, the unpaired fermion could occupy any HCB site.

In typical systems of interest, the particle number N will scale proportionally to the number of modes L; we have thus shown that the operator norm of each pseudoprojector Φ_k scales at most linearly in L.

Implications

Let us now think about how the above result can be used to reason about the adiabatic complexity of a choice of system or path, in terms of the numerator I_1 in eq. 3.35. We set a baseline with the following bound which applies to direct interpolation. Define the residual hamiltonian H^{r} in a generic fashion as in eq. 3.7, and observe that

$$\begin{aligned}\|\mathrm{H}^{\mathrm{r}}\| &\leq \sum_{\substack{P<R\\ Q<S}} \|F_{(PR)(QS)}\,\mathrm{a}_P^\dagger\mathrm{a}_R^\dagger\mathrm{a}_S\mathrm{a}_Q\| \\ &\leq \sum_{\substack{P<R\\ Q<S}} |F_{(PR)(QS)}| \leq \sqrt{L(L-1)/2}\,\|\mathrm{F}\|_{\mathrm{F}} \end{aligned} \tag{3.63}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm. Given the dimensionality of F, it is clear that $\|\mathrm{F}\|_{\mathrm{F}} \leq c\sqrt{L(L-1)/2}$ for some nonnegative constant c, and thus $\|\mathrm{H}^{\mathrm{r}}\| \leq O(L^2)$. The resulting numerator, for direct interpolation, from eq. 3.35 then scales as $O(L^4)$.

In comparison, consider a fully stepwise scheme where we "adiabatically add" every term $\mathrm{H}^{\mathrm{r}}_k$ from the eigendecomposition, eq. 3.11, separately, i.e. we evolve

$$\mathrm{H}^{\mathrm{i}} \to \mathrm{H}^{\mathrm{i}} + \mathrm{H}^{\mathrm{r}}_{k_1} \to \mathrm{H}^{\mathrm{i}} + \mathrm{H}^{\mathrm{r}}_{k_1} + \mathrm{H}^{\mathrm{r}}_{k_2} \to \cdots \to \mathrm{H} \tag{3.64}$$

where $\mathrm{A} \to \mathrm{B}$ denotes a direct adiabatic interpolation between A and B. In this scheme, at any point in the evolution, exactly one of the γ_k (cf. eq. 3.8) must increase at a rate scaling in the number of terms (which is $O(L^2)$), with the rest staying constant (being either 0 or 1). Combining the universal result that $\|\Phi_k\| \leq O(L)$ with eq. 3.38 (taking $M = L(L-1)/2$), we obtain

$$I_1 \in O(L^2 \cdot L^4) = O(L^6). \tag{3.65}$$

This indicates that the fully stepwise procedure is unfavourable as compared to direct interpolation. This is no surprise: with the work of Tomka et al [TSRP16] in mind, the direct path is a geodesic if the gap is held constant, whereas the fully stepwise approach is a walk along the corners of a hypercube in parameter space.

However, we emphasise that these bounds are worst case and can be improved in certain settings. Consider, for example, the standard Fermi-Hubbard model (see eq. 2.46) with the hopping part plus the chemical potential (which is proportional to the identity when N is fixed) taken as the initial hamiltonian. The spectral norm of the residual hamiltonian H^{r} is easily found to be $UN/2 \in O(L)$, leading to $I_1 \leq O(L^2)$. Furthermore, since H^{r} is diagonal in the two-particle position basis, all its pseudoprojectors are of the form $\Phi_k = \mathrm{a}^\dagger_{I_k}\mathrm{a}^\dagger_{J_k}\mathrm{a}_{J_k}\mathrm{a}_{I_k}$ and thus have unit spectral norm. Since H^{r} contains $L/2$ such terms, a fully stepwise procedure yields the same numerator bound, $I_1 \leq O(L^2)$.

On the other hand, the paired fermion formalism may be used to find examples which saturate the bounds of both eq. 3.63 and 3.65. To this end, we define a residual hamiltonian H^{r} and the corresponding operators Φ_k through its two-body eigenstates. As we have seen, a fully paired state $|\Psi\rangle = (\xi_1 \mathrm{a}^\dagger_1\mathrm{a}^\dagger_2 + \cdots + \xi_{\lfloor L/2 \rfloor}\mathrm{a}^\dagger_{L-1}\mathrm{a}^\dagger_L)|\mathrm{vac}\rangle$, where $\xi_m = (\lfloor L/2 \rfloor)^{-1/2}$ for all m, gives rise to a pseudoprojector with a maximal norm that is $O(L)$. Since the one-body basis is free to choose, we can permute the one-body modes in $L-1$ ways such that all resulting two-body states are fully paired and mutually orthogonal[1]. Furthermore, we

[1] One way to see this is to draw a complete graph of L nodes where nodes $2, \ldots, L$ are drawn in a circle around node 1. A full pairing (also known as perfect matching) may then be found by selecting an edge from node 1 to any other node and pairwise connecting the other nodes through edges orthogonal to the first edge. In this way, we find $L-1$ pairings without drawing any parallel edges, guaranteeing that that resulting two-body states are mutually orthogonal.

make use of the fact that the mapping $\xi_m \mapsto -\xi_m$ preserves the spectrum of any pseudoprojector Φ. After all, from eq. 3.46 we see that ξ_m appears in an off-diagonal element $\langle \mathcal{P}, \mathcal{U}|\Phi|\mathcal{P}', \mathcal{U}\rangle$ only if $m \in \mathcal{P}$ and $m \notin \mathcal{P}'$ or vice versa. Therefore this mapping is realised by the transformation $\Phi_k \mapsto \Sigma\Phi_k\Sigma$ where Σ is a diagonal matrix with a -1 entry in those columns corresponding to the product state $|\mathcal{P}, \mathcal{U}\rangle$ where $m \in \mathcal{P}$, and a $+1$ entry elsewhere. Now we can use this to vary the signs of the terms in a fully paired two-body state; if L is a power of two, we can construct $L/2$ vectors $\boldsymbol{\xi}$ that are the (normalised) columns of an Hadamard matrix, so that the resulting $L/2$ two-body states are mutually orthogonal. As such, we have defined the $L(L-1)/2$ two-body eigenstates necessary to describe an interacting hamiltonian, each of which is fully paired.

Now, since we have $L(L-1)/2$ pseudoprojectors with maximal norm, the inequality of eq. 3.65 is automatically saturated (for any choice of the two-body eigenvalues), and the adiabatic numerator is maximised for the stepwise procedure. Furthermore, this construction also attains a maximally scaling numerator in the case of direct interpolation, if we set all two-body eigenvalues to 1. Indeed, consider the sum over all pseudoprojectors Φ_k which carry the same pairing (and therefore only differ in their $\boldsymbol{\xi}$ vectors):

$$\begin{aligned}\sum_{k\,:\,\text{same pairing}} \Phi_k &= \sum_k \sum_{mn} \xi_m^k \xi_n^k \, \mathrm{a}_{2m-1}^\dagger \mathrm{a}_{2m}^\dagger \mathrm{a}_{2n} \mathrm{a}_{2n-1} \\ &= \sum_m \mathrm{n}_{2m-1} \mathrm{n}_{2m}. \end{aligned} \tag{3.66}$$

In other words, this particular sum of pseudoprojectors is an operator that counts all pairs of fermions in a product state that coincide with the mode pairs that define the pseudoprojectors. As a result, the sum over *all* pseudoprojectors defined in this example is an operator that counts all possible pairs of fermions, and is therefore simply equal to $N(N-1)/2$ times the identity. This operator saturates the bound $\|\mathrm{H}^\mathrm{r}\| \leq O(L^2)$ (under the assumption that $L/N = O(1)$) and therefore realises a maximal adiabatic numerator scaling of $O(L^4)$. This analysis establishes a condition, expressed in the paired fermion formalism, on the two-body eigenstates that yields worst-case numerator scaling for both direct and fully stepwise interpolation. In addition, it shows that direct interpolation indeed outperforms a fully stepwise protocol in this sense.

3.6 Discussion

In this chapter, we have proposed a new protocol for adiabatic preparation of fermionic many-body ground states based on the eigendecomposition of the (combined one- and two-body) coefficient tensor of the residual hamiltonian, being the difference between the initial and final hamiltonian, in second quantisation.

The eigenvectors in this decomposition are equivalent to two-body eigenstates of the residual hamiltonian. The method decomposes the residual hamiltonian into a sum of simpler terms, each of which corresponds to an eigenvalue and eigenvector from the eigendecomposition. In the adiabatic scheme, every point along the evolution path is then a linear combination of these terms.

We have demonstrated how this idea may be applied to generalised Fermi-Hubbard models, through a few small worked examples. Our finding is that a level crossing occurring in a direct interpolation from a mean-field hamiltonian, which arises from a discrete one-body symmetry, can be cured with the two-body decomposition approach. Although this is not a general superiority result, it shows the existence of scenarios in which the use of (partially) piecewise paths resulting from a two-body decomposition is advantageous as compared to direct interpolation. More precisely, in this approach, one can design a procedure which explicitly breaks the symmetry by interpolating through an intermediate hamiltonian which contains only a subset of the hamiltonian terms from the decomposition. As a result, a gap is seen to be opened. The conditions for this to occur are specific: while the initial hamiltonian must share a symmetry with the target hamiltonian and place the ground state in the incorrect symmetry sector, the two-particle matrix of the residual hamiltonian must have degenerate eigenvalues in order to mix two-particle eigenstates from the relevant symmetry sectors. However, if these conditions are met, the symmetry is broken with unit probability. As such, the need for (sometimes arbitrary) choices in the design of the path, or an expensive approximation of a geodesic path or counterdiabatic terms, is eliminated.

We wish to point out an additional benefit that comes with the two-body decomposition method, namely that it could be useful also in a setting where the symmetry in question is unknown. After all, in such a case, it is nontrivial to explicitly construct an additional symmetry-breaking hamiltonian term by hand. But using the two-body eigendecomposition method we can simply evolve in a stepwise fashion with a separate step for all degenerate two-body eigenvectors and, under the assumption that the problem instance satisfies the three applicability criteria, be guaranteed that the symmetry is broken. A caveat here is that while one can easily check whether there exist degenerate two-body eigenvectors, it is not straightforward to check whether they lie in different sectors with respect to the unknown symmetry. Nonetheless, the method can serve a black-box trial-and-error purpose simply by choosing random mixings of all degenerate two-body eigenvectors and separately interpolating the corresponding many-body hamiltonian terms, in a stepwise fashion. In this sense, the two-body decomposition may be viewed as a heuristic in a situation where otherwise complex methods (such as an NP-hard search over paths) would be required.

Having established this gap opening potential of the two-body decomposition method, we proceeded to analyse the adiabatic complexity of piecewise paths more broadly, by examining how the two-body decomposition influences the

numerator part of the complexity of many-body adiabatic state preparation. This numerator is primarily dependent on the operator norm of each term from the residual hamiltonian decomposition. We have found that a description in terms of fermion pairs (or equivalently, hard-core bosons) is key to understanding the scaling, in terms of the number of single-particle modes L, of this operator norm and therefore the adiabatic numerator. The main result is that each residual hamiltonian term scales at most as $O(L)$, for a typical system where the number of particles N scales proportionally with L. This result has different implications, depending on the system under investigation and the chosen evolution path. For example, for the Fermi-Hubbard model with the interaction part taken as the residual hamiltonian, the adiabatic complexity scales as $O(L^2/\Delta^3)$ both in a direct interpolation, and when following a fully piecewise path. This is due to the fact that the norm of each residual hamiltonian term from the decomposition scales as $O(1)$, and there are only L nonzero terms. On the other hand, for a situation in which all two-body eigenstates are uniformly weighted superpositions of distinct fermion pairs, each term attains the maximal scaling of $O(L)$; as a result, the time complexity of direct interpolation in this case scales as $O(L^4/\Delta^3)$, whereas under a fully stepwise path we find an $O(L^6/\Delta^3)$ scaling. Both these scalings are worst case. This finding agrees with the statement by Tomka et al. that a geodesic path in parameter space is generally beneficial in terms of time complexity [TSRP16]. The result suggests that one should be selective when choosing which of the hamiltonian terms from the decomposition to interpolate in a piecewise fashion, and which to interpolate directly. Namely, by piecewise interpolating only those terms which (are expected to) break any relevant symmetries, one retains the power to lift level crossings, while avoiding potentially unfavourable scaling in L. We note however that the situation of maximal scaling is a case where the residual hamiltonian is particularly dense. In large chemical systems, for example, the two-electron part of the hamiltonian is typically sparse, so it is expected that a lower adiabatic numerator can be achieved for such systems.

All in all, our examples show that the two-body eigendecomposition method can outperform direct interpolation through symmetry breaking, and we expect that the method can be helpful in situations beyond a single reflection symmetry. An example is the nonrelativistic treatment of molecular electronic structure, which maintains a $SU(2)$ spin symmetry. Another approach could be the use of a two-body eigendecomposition as a black box if there is a hidden symmetry and the precise cause of a gap closure is not straightforward to determine.

Chapter 4

Virtual mitigation of coherent non-adiabatic transitions by echo verification

4.1 Introduction

The study of quantum many-body systems requires the precise estimation of observables. Quantum state preparation is naturally a prerequisite to this end, which is the rationale behind quantum computers or quantum simulators. The adiabatic algorithm has demonstrated large success in a variety of platforms [BSL+16, EKC+22]. Still, the performance of current devices is hindered by noise, which cannot be error corrected, yet. Therefore, error mitigation techniques have been explored both theoretically and experimentally and can significantly improve the estimation of observables [CBB+23, OAG+23, KEA+23]. Surprisingly, there have been few synergies jointly considering error mitigation for the adiabatic algorithm.

Any quantum circuit can be efficiently simulated by the adiabatic algorithm [AvDK+07]. In adiabatic quantum computation, the system is initialized in the ground state of a trivial hamiltonian and one seeks to prepare the ground state of the final hamiltonian by slowly interpolating between the two. The success of the algorithm is determined by the speed of the adiabatic passage and spectral properties of the hamiltonians [Mes62, FGGS00]. More precisely, the total evolution time, or circuit depth, depends inverse polynomially on the minimum spectral gap between the ground state and the first excited state along the adiabatic path. These relations are quantified by the adiabatic theorem and versions thereof [JRS07, Ami09, WB12]. The adiabatic algorithm is especially suited for devices that implement dynamics natively without any Trotter overhead [GB17, SSW+21, BLS+22, KRL+23].

To address the restrictions in current hardware, various error mitigation techniques have been explored in recent years to improve the usefulness of a noisy

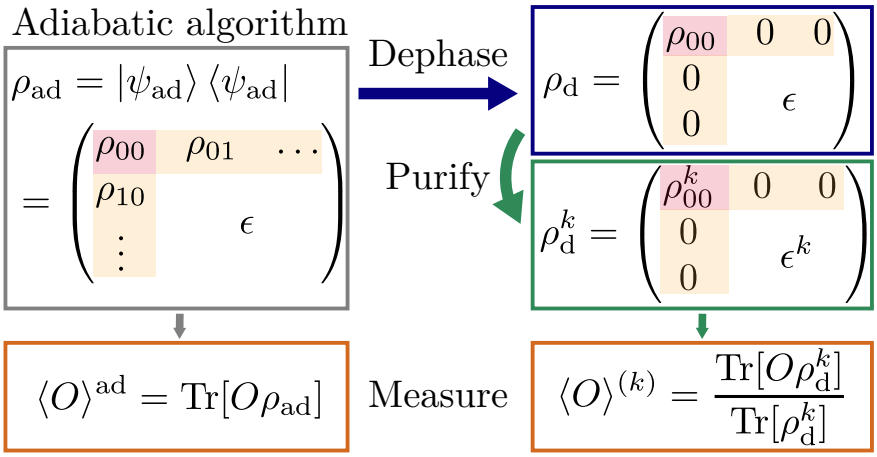


Figure 4.1. Schematic overview of the method. Density matrices are expressed in the energy eigenbasis of the target hamiltonian. The pure state after the adiabatic evolution (ρ_{ad}) approximates the true ground state. Via a dephasing operation, the coherent error is promoted to an incoherent error in ρ_{d} such that error mitigation techniques can be applied. This allows measuring the k^{th} degree purified observable $\langle \mathrm{O}\rangle^{(k)}$ which yields a lower bias than evaluating the state directly after the adiabatic preparation $\left[\langle \mathrm{O}\rangle^{\mathrm{ad}}\right]$.

quantum computation [CBB+23]. These methods include zero-noise extrapolation, exploiting symmetry or purity constraints, and several other approaches. Here, we focus on purity methods, which aim to suppress stochastic errors by projecting the noisy state ρ onto the closest pure state, given by the dominant eigenvector of ρ.

The purification can in general be achieved by collective measurements of several copies of ρ, known as virtual state distillation [HMO+21] or error suppression by derangement [Koc21]. Echo verification (EV) achieves this using two copies of ρ multiplexed in time, rather than in space [Cai21, HL22, OPR+21]. In EV, a desired state is prepared, an observable is measured controlled by an auxiliary qubit, and the state is then uncomputed. This allows to access expectation values of the so-called 2^{nd}-degree purified state of ρ: $\langle \mathrm{O}\rangle_{\mathrm{EV}} = \mathrm{tr}[\mathrm{O}\rho^2]/\,\mathrm{tr}[\rho^2]$. Recently, purification-based error mitigation has been tested experimentally in the context of the variational quantum eigensolver [OAG+23]. Error mitigation methods tailored specifically to the adiabatic algorithm have been explored considerably less in the literature. Few exceptions consider error suppression and correction [YSBK13] or symmetry protection for Trotter dynamics [TSCT21].

In this chapter, we present a mitigation technique for estimating observables on quasi-adiabatically-prepared states, in the spirit of echo verification. Along with stochastic device noise, our method seeks to suppress the coherent error due to non-adiabatic transitions. Our method which we denote *adiabatic echo verification* (AEV) relies on dephasing operations to promote the coherent errors to random errors, which can then be mitigated. We illustrate the combination of

dephasing and purification on a pure state in fig. 4.1. Similar to the original echo verification technique, the leading order error in the ground state expectation value of an observable is suppressed quadratically. In particular, we consider an imperfect implementation of the dephasing operation using random-time evolution. Related random-time dynamics have been successfully used in the context of Zeno-type protocols [BKS09]. The overhead from this dephasing operation is only poly-logarithmic in the accuracy of the dephasing operation for estimating observables of states within gapped phases. We discuss how the protocol compares favorably against doubling the total evolution time in the standard adiabatic algorithm. A key feature of our technique is that hardware noise is also mitigated naturally through the EV method. Our protocol only requires implementing positive-time evolution and applying the operator of interest in a controlled way. Hence, the protocol is not only suitable for purely gate-based quantum devices but also for hybrid quantum simulators, e.g. using neutral Rydberg atoms [BLS$^+$22].

4.2 Background

4.2.1 The adiabatic algorithm

In order to be able to measure observables on the ground state $|E_0\rangle$ of a target hamiltonian H_T, a state approximating $|E_0\rangle$ with sufficient precision needs to be prepared. The quantum adiabatic algorithm (QAA) is a suitable algorithm for this task. At the heart of the QAA is the adiabatic theorem, which states that a system remains in an instantaneous eigenstate if the hamiltonian is changed sufficiently slowly and the eigenstate is separated from other eigenstates by a minimum spectral gap $\Delta_{\min}$ throughout the transition [BF28]. Hence, the desired ground state $|\psi_T\rangle$ of a hamiltonian of interest H_T can be prepared by interpolating from a suitable hamiltonian H_0 with a trivial ground state $|\psi_0\rangle$ as

$$\mathrm{H}(\lambda) = (1-\lambda)\mathrm{H}_0 + \lambda \mathrm{H}_T. \tag{4.1}$$

where $\lambda = t/T$ is the parametrized time. The adiabatic theorem 2.3.1 gives a bound $T \in \Omega\left(\Delta_{\min}^{-3}\epsilon^{-1}\right)$ if $\mathrm{H}(\lambda)$ is twice differentiable [JRS07, Ami09]. Given a finite coherence time, the QAA prepares an approximation to the target state $|\psi_{\mathrm{ad}}\rangle = \sqrt{1-\epsilon}|E_0\rangle + \sqrt{\epsilon}|E_0^{\perp}\rangle$ where $\langle E_0|E_0^{\perp}\rangle = 0$. Measuring an observable O, we obtain an approximation to the true value $\mathrm{tr}\left[\mathrm{O}|\psi_{\mathrm{ad}}\rangle\langle\psi_{\mathrm{ad}}|\right] = (1-\epsilon)\langle \mathrm{O}\rangle_{|E_0\rangle} + O(\sqrt{\epsilon})$.

4.2.2 Purification-based error mitigation

Purification methods such as echo verification (EV) or virtual state distillation improve the quality of an expectation value measurement on a noisy (incoherent) approximation ρ of a pure state $|\psi\rangle\langle\psi|$. This is achieved by effectively measuring

the expectation value $\mathrm{tr}[\mathrm{O}\rho^k]$ of O on the k^{th} power of the density matrix. Raising ρ to the k^{th} power suppresses the eigenvectors with smaller eigenvalues, increasing the relative weight of the dominant eigenvector which, for small enough noise, should be close to $|\psi\rangle$. As ρ^k is non-normalized, purification methods prescribe to independently measure $\mathrm{tr}[\rho^k]$ to calculate the desired estimator

$$\langle \mathrm{O} \rangle^{(k)} \coloneqq \mathrm{tr}[\mathrm{O}\rho^k] / \mathrm{tr}[\rho^k]. \tag{4.2}$$

If ρ has an eigenstate $|\phi\rangle$ with large weight $c_0 = 1 - \epsilon$ (small positive ϵ), we can write the density matrix as $\rho = c_0 |\phi\rangle\langle\phi| + \epsilon\rho_\perp$ where $\rho_\perp$ is a density matrix orthogonal to $|\phi\rangle$ (i.e., $\rho_\perp |\phi\rangle = 0$). The k^{th} degree purified estimator is then

$$\langle \mathrm{O} \rangle^{(k)} = \frac{c_0^k \langle\phi|\mathrm{O}|\phi\rangle + \epsilon^k \,\mathrm{tr}[\rho_\perp^k \mathrm{O}]}{c_0^k + \epsilon^k \,\mathrm{tr}[\rho_\perp^k]} \tag{4.3}$$

$$= \langle\phi|\mathrm{O}|\phi\rangle + O(\epsilon^k \,\mathrm{tr}[\rho_\perp^k] \|\mathrm{O}\|). \tag{4.4}$$

Echo verification implements purification for $k = 2$ using a single register by multiplexing two state-(un)preparation oracles in time. The method suppresses the error contributions to the observable estimator such that the leading order becomes $O(\epsilon^2)$.

4.3 Mitigating coherent errors in adiabatic state preparation

Our main contribution is to propose a method where the echo verification technique is applied to coherent errors. We focus on an application where the coherent error arises in the adiabatic algorithm due to finite algorithm runtimes. However, as the state prepared by a noiseless implementation of the adiabatic algorithm is pure, naive purification will not have any effect.

To recover the error mitigation power on $\rho_{\mathrm{ad}} = |\psi_{\mathrm{ad}}\rangle\langle\psi_{\mathrm{ad}}|$, we introduce an ideal dephasing channel that turns coherent errors into incoherent noise,

$$\underset{\mathrm{H}}{\mathrm{deph}}[\rho] \coloneqq \sum_j |E_j\rangle\langle E_j|\rho|E_j\rangle\langle E_j| = \mathrm{diag}[\rho], \tag{4.5}$$

where we sum over an eigenbasis $\{|E_j\rangle\}_j$ of the target hamiltonian H_T.

We note that this formulation assumes a nondegenerate spectrum. If the spectrum of H contains degeneracies, the dephasing channel will project ρ to a block-diagonal operator, where the blocks are defined by the (degenerate) eigenspaces of H:

$$\underset{\mathrm{H}}{\mathrm{deph}}[\rho] = \sum_j \Pi_j \rho \Pi_j \tag{4.6}$$

where $\Pi_j = \delta(\mathrm{H} - E_j)$ is the projector on the eigenspace of H with eigenvalue E_j.

For AEV, we are only interested in dephasing the ground state with respect to the rest of the spectrum; we only require the ground state of H to be non-degenerate for eq. 4.7 and the subsequent analysis to be valid. This is anyway a typical requirement in adiabatic state preparation.

The dephasing channel projects a density matrix onto its diagonal in the energy eigenbasis, removing the off-diagonal coherences. Applying the channel to the state prepared by the adiabatic algorithm yields

$$\rho_{\mathrm{d}} \coloneqq \underset{\mathrm{H}}{\mathrm{deph}}[\rho_{\mathrm{ad}}] = c_0|E_0\rangle\langle E_0| + \epsilon\rho_\perp \tag{4.7}$$

$$= \begin{pmatrix} c_0 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \epsilon\rho_\perp & \\ 0 & & & \end{pmatrix} \tag{4.8}$$

with $\rho_\perp = \sum_{j\neq 0} \rho_{jj}\epsilon^{-1}|E_j\rangle\langle E_j|$. In this setting, the high-weight eigenstate $|\phi\rangle$ of $\mathrm{deph}_{\mathrm{H}}[\rho_{\mathrm{ad}}]$ coincides with the ground state $|E_0\rangle$ of the hamiltonian. Then, using the echo verification technique on the dephased state, which is a mixed state, we obtain the following result for the observable O:

$$\frac{\mathrm{tr}[\mathrm{O}\rho_{\mathrm{d}}^k]}{\mathrm{tr}[\rho_{\mathrm{d}}^k]} = (1-\gamma)\langle E_0|\mathrm{O}|E_0\rangle + \gamma\frac{\mathrm{tr}[\mathrm{O}\rho_\perp^k]}{\mathrm{tr}[\rho_\perp^k]}, \tag{4.9}$$

with

$$\gamma = \left[1 + c_0^k / \left(\epsilon^k \,\mathrm{tr}[\rho_\perp^k]\right)\right]^{-1} \in O\left(\epsilon^k \,\mathrm{tr}[\rho_\perp^k]\right). \tag{4.10}$$

To implement echo verification, typically, an inverse pair of unitaries $(\mathrm{U}_\rightarrow, \mathrm{U}_\rightarrow^\dagger)$ would be required [OPR+21]. The quantum circuit that implements adiabatic echo verification for an operator O and a depasing operation is shown in fig. 4.2(a). A priori, the unpreparation $\mathrm{U}_\rightarrow^\dagger$ would use negative-time dynamics, which is generally not available in analog simulators. For our purposes, however, we can consider the two states $\rho_{\mathrm{ad}} = \mathrm{U}_\rightarrow|\psi_0\rangle\langle\psi_0|\mathrm{U}_\rightarrow^\dagger$ and $\sigma_{\mathrm{ad}} = \mathrm{U}_\leftarrow^\dagger|\psi_0\rangle\langle\psi_0|\mathrm{U}_\leftarrow$, where $\mathrm{U}_\leftarrow$ is a positive-time adiabatic evolution with an inverted schedule from $\lambda = 1$ to $\lambda = 0$. Both states have the same guaranteed fidelity with the target state $|E_0\rangle$ from the adiabatic theorem and ground state coherences are surpressed after the dephasing operation. This allows to use positive-time dynamics for the unpreparation step in AEV. We illustrate the inverted schedule in fig. 4.2(b), where a red line indicates the hamiltonian in the dynamics.

Next, we consider the implementation of the dephasing channel. Importantly, we observe that a channel that dephases only the ground state would also be sufficient to achieve our goal, producing a state of the form eq. 4.7 with a more general, non-diagonal $\rho_\perp$, provided that c_0 still dominates. In the following part, we analyze such an approximate dephasing operation using positive-time dynamics.

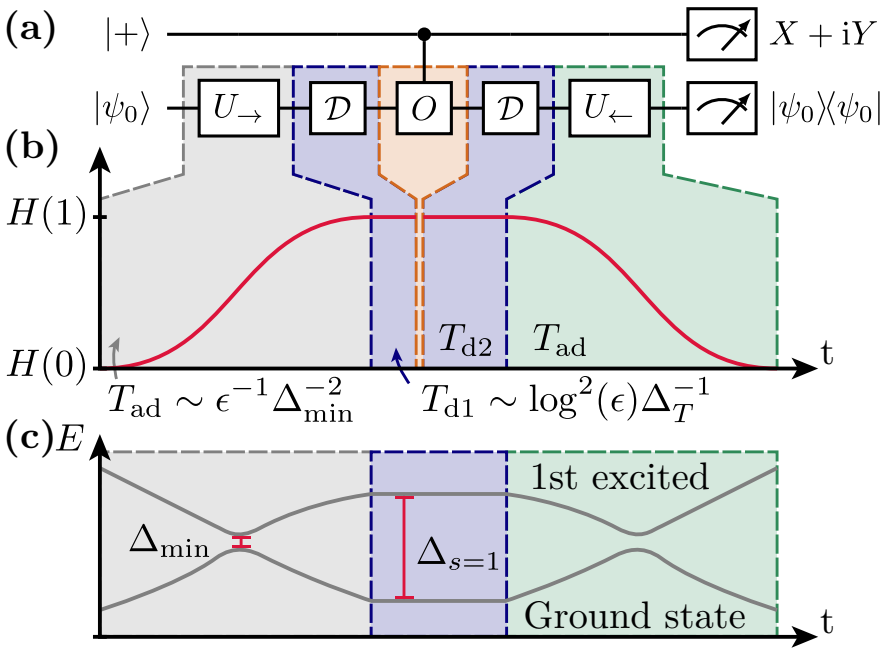


Figure 4.2. **(a)** Quantum circuit for adiabatic echo verification to estimate an observable $\langle O\rangle$. A quasi-adiabatic sweep $U_\rightarrow$ is followed by an approximate ground state dephasing operation $\mathcal{D}$. After the controlled application of a unitary observable O and dephasing again, the sweep is performed backward $U_\leftarrow(\neq U_\rightarrow^\dagger)$. Postprocessing the measurement result, including the success information of the ground state projection, allows to extract an improved expectation value. **(b)** Schematic of the hamiltonian dynamics. Approximate dephasing is implemented by evolving with the target hamiltonian at $\lambda = 1$ for a random time. Typically, this time is much smaller than the time required for the adiabatic algorithm as depicted in **(c)**, where we sketch a corresponding low-energy spectrum.

4.3.1 Implementation and cost of the dephasing

We can implement an approximation of the dephasing channel (eq. 4.5) by a random-time evolution $\exp(-iH_T\tau)$, with τ sampled from a probability distribution $P(\tau)$, as follows. We limit the support of P to the interval $\tau \in [0, T_d]$. This ensures the dephasing can be realized naturally in quantum simulators and limits the time overhead of the dephasing operation to $2\,T_d$ for the AEV circuit. We define the approximate dephasing channel

$$\underset{H,P}{\mathrm{deph}}[\rho] \coloneqq \int_0^{T_d} d\tau\, P(\tau)\, e^{-iH_T\tau}\rho e^{iH_T\tau} \tag{4.11}$$

$$= \sum_{j,k} \mathcal{F}_{jk}|E_j\rangle\langle E_j|\rho|E_k\rangle\langle E_k|, \tag{4.12}$$

where $\mathcal{F}_{jk} \coloneqq \mathcal{F}[P](E_j - E_k)$ is the Fourier transform of the random-time distribution at the transition energies. We will make use of the shorthand $\mathcal{D}[\rho] \coloneqq$

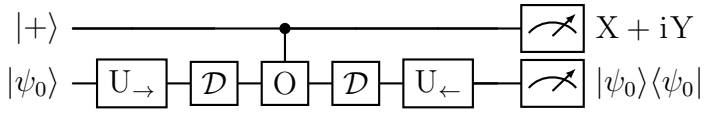
$\mathrm{deph}_{\mathrm{H},P}[\rho]$.

The echo verification (EV) expectation value estimator [PAO23, OAG+23] is defined as

$$\langle \mathrm{O} \rangle^{\mathrm{EV}} := \frac{\mathbb{E}[\text{verified Hadamard test circuit (VHT)}]}{\mathbb{E}[\text{echo circuit (Echo)}]}. \tag{4.13}$$

In our case, the verified Hadamard test circuit is

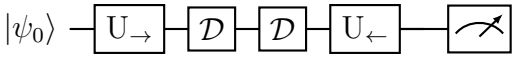
VHT :=



, (4.14)

where $\mathrm{U}_{\rightarrow}$ is the adiabatic state preparation, $\mathrm{U}_{\leftarrow}$ is the state unpreparation, and $\mathcal{D} := \mathrm{deph}_{\mathrm{H},P}$ is the approximate dephasing. At the end of the circuit we need to measure $\mathrm{X} \otimes |\psi_0\rangle\langle\psi_0|$ and $\mathrm{Y} \otimes |\psi_0\rangle\langle\psi_0|$ to recover the real $\mathrm{Re}[\langle \mathrm{O} \rangle^{\mathrm{EV}}]$ and imaginary $\mathrm{Im}[\langle \mathrm{O} \rangle^{\mathrm{EV}}]$ parts of the expectation value, respectively. The output of a single sample of the circuit will be the result of the Pauli X (± 1) or Y ($\pm \mathrm{i}$) on the control qubit if the system register returns to the state $|\psi_0\rangle$, and 0 otherwise. Our notation supposes that O is a unitary operator, and its application is controlled by the state of the control qubit. If O is not unitary, we can rewrite it as a decomposition $\mathrm{O} = \sum_x a_x \mathrm{Re}[\mathrm{U}_x] + b_x \mathrm{Im}[\mathrm{U}_x]$ and measure the terms of the decomposition separately [PAO23]. The echo circuit, which is used to compute the normalization of $\langle \mathrm{O} \rangle^{\mathrm{EV}}$, is given as

Echo :=



(4.15)

and obtained by substituting the operator O with the identity in the previous circuit.

The adiabatic preparation and the adiabatic unpreparation are defined as

$$\begin{aligned} \mathrm{U}_{\rightarrow} &= \mathcal{T} \exp\left[-\mathrm{i} \int_0^{T_{\mathrm{ad}}} \mathrm{H}[\lambda(t)]\, \mathrm{d}t \right]; \\ \mathrm{U}_{\leftarrow} &= \mathcal{T} \exp\left[-\mathrm{i} \int_0^{T_{\mathrm{ad}}} \mathrm{H}[\lambda(T_{\mathrm{ad}} - t)]\, \mathrm{d}t \right], \end{aligned} \tag{4.16}$$

where $\mathcal{T}\exp$ notates the time-ordered exponential, $\mathrm{H}(\lambda)$ is the adiabatic hamiltonian (eq. 4.1), $\lambda(t)$ the adiabatic schedule with $\lambda(0) = 0$ and $\lambda(T_{\mathrm{ad}}) = 1$, and T_{ad} is the total evolution time of the adiabatic algorithm. Note that $\mathrm{d}t$ is always positive, thus negative-time evolution is not required to implement $\mathrm{U}_{\rightarrow}$ and $\mathrm{U}_{\leftarrow}$. Typically, in an EV circuit, if the preparation unitary is U, then the unpreparation is performed with its conjugate transpose $\mathrm{U}^\dagger$ such that $\mathrm{U}\mathrm{U}^\dagger = \mathrm{I}$. Here, however, this is not the case for the two operations $\mathrm{U}_{\leftarrow}$ and $\mathrm{U}_{\rightarrow}$. We will show that the added dephasing indeed removes this requirement for our purposes.

In our calculations, we only assume that $\mathrm{U}_{\rightarrow}$ ($\mathrm{U}_{\leftarrow}$) implement an approximate state (un)preparation of $|E_0\rangle$ with a fidelity of at least $1-\epsilon$ with small $\epsilon > 0$. Concretely, we define

$$|\langle E_0|\mathrm{U}_{\rightarrow}|\psi_0\rangle|^2 = 1-\epsilon_{\leftarrow}\,, \quad |\langle\psi_0|\mathrm{U}_{\leftarrow}|E_0\rangle|^2 = 1-\epsilon_{\rightarrow}, \tag{4.17}$$

such that $\epsilon = \max\{\epsilon_{\leftarrow}, \epsilon_{\rightarrow}\}$. It is reasonable to assume the two adiabatic processes will have a similar error, as any adiabatic theorem bounds both in the same way.

Writing ${}^{C}\mathrm{O}$ for controlled-O, the expectation value of circuit eq. 4.14 is

$$\begin{aligned}
\mathbb{E}[\mathrm{VHT}] &= \mathrm{tr}\Big\{\mathrm{U}_{\leftarrow}\,\mathcal{D}\Big[{}^{C}\mathrm{O}\;\mathcal{D}\big[\mathrm{U}_{\rightarrow}(|+\rangle\langle+|\otimes|\psi_0\rangle\langle\psi_0|)\mathrm{U}_{\rightarrow}^{\dagger}\big]\;{}^{C}\mathrm{O}^{\dagger}\Big]\mathrm{U}_{\leftarrow}^{\dagger} \\
&\qquad \times\big(|\psi_0\rangle\langle\psi_0|\otimes 2|0\rangle\langle 1|\big)\Big\} \\
&= \mathrm{tr}\Big\{\underbrace{\mathrm{U}_{\leftarrow}^{\dagger}|\psi_0\rangle\langle\psi_0|\mathrm{U}_{\leftarrow}}_{\sigma}\;\mathcal{D}\Big[\mathrm{O}\;\mathcal{D}\big[\underbrace{\mathrm{U}_{\rightarrow}|\psi_0\rangle\langle\psi_0|\mathrm{U}_{\rightarrow}^{\dagger}}_{\rho}\big]\Big]\Big\} \\
&= \sum_{jk}\mathcal{F}_{jk}\,\mathrm{tr}\Big\{\mathrm{U}_{\leftarrow}^{\dagger}|\psi_0\rangle\langle\psi_0|\mathrm{U}_{\leftarrow}\;\mathcal{D}\Big[\mathrm{O}|E_j\rangle\underbrace{\langle E_j|\mathrm{U}_{\rightarrow}|\psi_0\rangle\langle\psi_0|\mathrm{U}_{\rightarrow}^{\dagger}|E_k\rangle}_{\rho_{jk}}\langle E_k|\Big]\Big\} \\
&= \sum_{jkl}\mathcal{F}_{jk}\mathcal{F}_{lk}\sigma_{kl}O_{lj}\rho_{jk}
\end{aligned} \tag{4.18}$$

where we expand the dephasing channels, and we define the density matrices ρ and σ, corresponding respectively to the pure states

$$\mathrm{U}_{\rightarrow}|\psi_0\rangle = \sqrt{1-\epsilon}|E_0\rangle + \sqrt{\epsilon}\sum_{j>0}\alpha_j|E_j\rangle\,, \quad \sum_{j>0}|\alpha_j|^2 = 1; \tag{4.19}$$

$$\langle\psi_0|\mathrm{U}_{\leftarrow} = \sqrt{1-\epsilon}\langle E_0| + \sqrt{\epsilon}\sum_{j>0}\beta_j^{*}\langle E_j|\,, \quad \sum_{j>0}|\beta_j|^2 = 1. \tag{4.20}$$

We can then absorb the dephasing coefficients into $\tilde{\rho}_{jk} = \mathcal{F}_{jk}\rho_{jk}$ and $\tilde{\sigma}_{kl} = \mathcal{F}_{lk}\sigma_{kl} = \mathcal{F}_{lk}^{*}\sigma_{kl}$, simplifying

$$\mathbb{E}[\mathrm{VHT}] = \mathrm{tr}[\tilde{\rho}\tilde{\sigma}\mathrm{O}]\,, \quad \mathbb{E}[\mathrm{Echo}] = \mathrm{tr}[\tilde{\rho}\tilde{\sigma}]. \tag{4.21}$$

Evaluating the adiabatic echo verification circuit with the *approximate* dephasing channel $\mathcal{D}[\rho]$ therefore yields an estimator with expectation

$$\langle\mathrm{O}\rangle^{\mathrm{AEV}} = \frac{\mathrm{tr}[\mathrm{O}\tilde{\rho}\tilde{\sigma}]}{\mathrm{tr}[\tilde{\rho}\tilde{\sigma}]} \tag{4.22}$$

where $\tilde{\rho}_{jk} = \mathcal{F}_{jk}[\rho_{\mathrm{ad}}]_{jk}$ and $\tilde{\sigma}_{kl} = \mathcal{F}_{kl}^{*}[\sigma_{\mathrm{ad}}]_{kl}$, expressed as matrix elements in the eigenbasis of the target hamiltonian. Comparing this result to the standard purification estimator eq. 4.2, we see that the ρ^2 is substituted by $\tilde{\rho}\tilde{\sigma}$.

4.3.2 Evaluation of the AEV estimator with approximate dephasing

The accuracy of the AEV estimator is dependent on the Fourier coefficients $\mathcal{F}_{jk}$. Specifically, since we are only interested in expectation values with respect to the ground state, the only requirement on the dephasing channel is that the Fourier coefficients $\max_{j>0} |\mathcal{F}_{0j}|$ are upper bounded by a constant δ. In other words, we require that the coherences between the ground state and any other eigenstate are suppressed by a factor smaller than δ. This is sufficient to establish the following bound on the deviation of the AEV estimator from the ground state expectation value.

4.3.1. Theorem. *Let the AEV estimator $\langle \mathrm{O} \rangle^{AEV}$ be as in eq. 4.22. Let the coherent error from the adiabatic (un)preparation $\mathrm{U}_{\rightarrow}$ and $\mathrm{U}_{\leftarrow}$ be at most $\epsilon < \sqrt{3/2} - 1$, and let the approximate dephasing channel $\mathcal{D}$ be such that the corresponding Fourier coefficients satisfy $|\mathcal{F}_{0j}| \leq \delta$. Then the estimator bias w.r.t. the true ground state expectation value is upper bounded by*

$$\Big| \langle \mathrm{O} \rangle^{AEV} - \langle E_0 | \mathrm{O} | E_0 \rangle \Big| \in O\big(\|\mathrm{O}\| (\epsilon^{1/2} \delta + \epsilon^2) \big). \tag{4.23}$$

Proof.
We bound the error of $\langle \mathrm{O} \rangle^{\mathrm{EV}}$ with respect to the target $\langle E_0 | \mathrm{O} | E_0 \rangle$,

$$\text{error} \coloneqq \left| \frac{\mathrm{tr}[\mathrm{O}\tilde{\rho}\tilde{\sigma}]}{\mathrm{tr}[\tilde{\rho}\tilde{\sigma}]} - \langle E_0 | \mathrm{O} | E_0 \rangle \right| = \big| \mathrm{tr}[\mathrm{O}\tilde{\rho}\tilde{\sigma}] - O_{00} \,\mathrm{tr}[\tilde{\rho}\tilde{\sigma}] \big| \cdot \big| \mathrm{tr}[\tilde{\rho}\tilde{\sigma}] \big|^{-1}, \tag{4.24}$$

by explicitly writing out the matrix elements of $\tilde{\rho}\tilde{\sigma}$ in the eigenbasis of H_T:

$$\begin{aligned}
\tilde{\rho}\tilde{\sigma} =& \Big[(1-\epsilon)^2 + \epsilon(1-\epsilon) \sum_{j>0} \alpha_j^* \beta_j \mathcal{F}_{j0}^* \mathcal{F}_{0j} \Big] |E_0\rangle\langle E_0| \\
&+ \sqrt{\epsilon}\sqrt{1-\epsilon} \sum_{j>0} \Big[(1-\epsilon)\alpha_j \mathcal{F}_{j0} + \epsilon \sum_{l>0} \alpha_j \alpha_l^* \beta_l \mathcal{F}_{jl} \mathcal{F}_{l0}^* \Big] |E_j\rangle\langle E_0| \\
&+ \sqrt{\epsilon}\sqrt{1-\epsilon} \sum_{j>0} \Big[(1-\epsilon)\beta_j^* \mathcal{F}_{j0} + \epsilon \sum_{l>0} \beta_j^* \alpha_l^* \beta_l \mathcal{F}_{jl} \mathcal{F}_{l0}^* \Big] |E_0\rangle\langle E_j| \\
&+ \epsilon \sum_{j,k>0} \Big[(1-\epsilon)\alpha_j \beta_k^* \mathcal{F}_{j0} \mathcal{F}_{0k}^* + \epsilon \sum_{l>0} \alpha_j \alpha_l^* \beta_l \beta_k^* \mathcal{F}_{jl} \mathcal{F}_{lk}^* \Big] |E_j\rangle\langle E_k| \qquad (4.25) \\
=& [\tilde{\rho}\tilde{\sigma}]_{00} |E_0\rangle\langle E_0| + \sum_{j>0} [\tilde{\rho}\tilde{\sigma}]_{j0} |E_j\rangle\langle E_0| + \sum_{j>0} [\tilde{\rho}\tilde{\sigma}]_{0j} |E_0\rangle\langle E_j| \\
&+ \sum_{j,k>0} [\tilde{\rho}\tilde{\sigma}]_{jk} |E_j\rangle\langle E_k|. \qquad (4.26)
\end{aligned}$$

In this basis, the relevant terms read

$$\mathrm{tr}[\mathrm{O}\tilde{\rho}\tilde{\sigma}] = [\tilde{\rho}\tilde{\sigma}]_{00}O_{00} + \sum_{j>0}[\tilde{\rho}\tilde{\sigma}]_{j0}O_{0j} + \sum_{k>0}[\tilde{\rho}\tilde{\sigma}]_{0k}O_{k0} + \sum_{j,k>0}[\tilde{\rho}\tilde{\sigma}]_{jk}O_{kj}, \tag{4.27}$$

$$\mathrm{tr}[\tilde{\rho}\tilde{\sigma}] = [\tilde{\rho}\tilde{\sigma}]_{00} + \sum_{j>0}[\tilde{\rho}\tilde{\sigma}]_{jj}. \tag{4.28}$$

We focus first on bounding the first factor on the right-hand side of eq. 4.24,

$$\left|\mathrm{tr}[\mathrm{O}\tilde{\rho}\tilde{\sigma}] - O_{00}\,\mathrm{tr}[\tilde{\rho}\tilde{\sigma}]\right| = \left|\mathrm{tr}[(\mathrm{O} - O_{00})\tilde{\rho}\tilde{\sigma}]\right|. \tag{4.29}$$

We separate this expression through a triangle inequality,

$$\begin{aligned}\left|\mathrm{tr}[\mathrm{O}\tilde{\rho}\tilde{\sigma}] - O_{00}\,\mathrm{tr}[\tilde{\rho}\tilde{\sigma}]\right| \leq{}& \left|\sum_{j>0}[\tilde{\rho}\tilde{\sigma}]_{j0}O_{0j}\right| + \left|\sum_{k>0}[\tilde{\rho}\tilde{\sigma}]_{0k}O_{k0}\right| \\ &+ \left|\sum_{j,k>0}[\tilde{\rho}\tilde{\sigma}]_{jk}O_{kj} - \sum_{j>0}[\tilde{\rho}\tilde{\sigma}]_{jj}O_{00}\right|.\end{aligned} \tag{4.30}$$

To bound the first term, we first use that, by the definition of the spectral norm,

$$\left|\sum_{j>0}[\tilde{\rho}\tilde{\sigma}]_{0j}O_{j0}\right| \leq \|O\| \cdot \|[\tilde{\rho}\tilde{\sigma}]_{\cdot 0}\| \tag{4.31}$$

where $[\tilde{\rho}\tilde{\sigma}]_{\cdot 0}$ is the vector with entries $[\tilde{\rho}\tilde{\sigma}]_{j0}$ for $j > 0$ (and zero for $j = 0$). Let us explicitly write out an upper bound on $\|[\tilde{\rho}\tilde{\sigma}]_{\cdot 0}\|^2$:

$$\begin{aligned}\|[\tilde{\rho}\tilde{\sigma}]_{\cdot 0}\|^2 &= \epsilon(1-\epsilon)\sum_{j>0}\left|(1-\epsilon)\alpha_j\mathcal{F}_{j0} + \epsilon\sum_{l>0}\alpha_j\alpha_l^*\beta_l\mathcal{F}_{jl}\mathcal{F}_{l0}^*\right|^2 \\ &\leq \epsilon(1-\epsilon)\sum_{j>0}|\alpha_j|^2\left((1-\epsilon)|\mathcal{F}_{j0}| + \epsilon\left|\sum_{l>0}\alpha_l^*\beta_l\mathcal{F}_{jl}\mathcal{F}_{l0}^*\right|\right)^2 \\ &\leq \epsilon(1-\epsilon)\sum_{j>0}|\alpha_j|^2\left((1-\epsilon)|\mathcal{F}_{j0}| + \epsilon\sum_{l>0}|\alpha_l^*||\beta_l||\mathcal{F}_{jl}||\mathcal{F}_{l0}^*|\right)^2 \\ &\leq \epsilon(1-\epsilon)\|\boldsymbol{\alpha}\|^2\left((1-\epsilon)\delta + \epsilon\delta\,\|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|\right)^2 \\ &= \epsilon(1-\epsilon)\delta^2\end{aligned} \tag{4.32}$$

where we note that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are normalized by definition, $\max_{k>0}|\mathcal{F}_{0k}| = \delta$ and $\mathcal{F}_{jk} \leq 1$; Cauchy-Schwarz was used in the fourth line. Hence, the first term in eq. 4.30 is upper bounded by $\|O\|\sqrt{\epsilon}\sqrt{1-\epsilon}\delta$. The same bound applies to the second term $\left|\sum_{k>0}[\tilde{\rho}\tilde{\sigma}]_{k0}O_{0k}\right|$.

The last term of eq. 4.30 can be rewritten as

$$\sum_{j,k>0}[\tilde{\rho}\tilde{\sigma}]_{jk}O_{kj} - \sum_{j>0}[\tilde{\rho}\tilde{\sigma}]_{jj}O_{00} = \mathrm{tr}[\Pi_{>}\,\tilde{\rho}\tilde{\sigma}\,\Pi_{>}(\mathrm{O} - O_{00}\mathrm{I})] \tag{4.33}$$

where $\Pi_> = \mathrm{I} - |E_0\rangle\langle E_0|$ is the projector on the subspace orthogonal to $|E_0\rangle$. We can then use the Von Neumann inequality to bound

$$\left|\mathrm{tr}[\Pi_> \,\tilde{\rho}\tilde{\sigma}\,\Pi_>(\mathrm{O} - O_{00}\mathrm{I})]\right| \leq \|\mathrm{O} - O_{00}\mathrm{I}\| \cdot \|\Pi_> \,\tilde{\rho}\tilde{\sigma}\,\Pi_>\|_1. \tag{4.34}$$

Now, by virtue of the triangle inequality, and the fact that $\||u\rangle\langle v|\|_1 = \|u\|\|v\|$, for any vectors $|u\rangle, |v\rangle$, we have

$$\begin{aligned}
\|\Pi_>\tilde{\rho}\tilde{\sigma}\Pi_>\|_1 &\leq \epsilon(1-\epsilon)\left\|\sum_{j>0}\alpha_j\mathcal{F}_{j0}|E_j\rangle\right\|\left\|\sum_{k>0}\beta_k\mathcal{F}_{0k}|E_k\rangle\right\| \\
&\quad + \epsilon^2\sum_{l>0}|\alpha_l||\beta_l|\left\|\sum_{j>0}\alpha_j\mathcal{F}_{jl}|E_j\rangle\right\|\left\|\sum_{k>0}\beta_k\mathcal{F}_{lk}|E_k\rangle\right\| \\
&\leq \epsilon(1-\epsilon)\delta^2 + \epsilon^2\,(\max_{j,k>0}|\mathcal{F}_{jk}|)^2\sum_{l>0}|\alpha_l||\beta_l| \\
&\leq \epsilon(1-\epsilon)\delta^2 + \epsilon^2
\end{aligned} \tag{4.35}$$

where we used Cauchy-Schwarz in the last line together with the normalisation of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

Combining the bounds from eqs. 4.32 and (4.35), and using that $\|\mathrm{O} - O_{00}\mathrm{I}\| \leq 2\|\mathrm{O}\|$, we get

$$\left|\mathrm{tr}[\mathrm{O}\tilde{\rho}\tilde{\sigma}] - O_{00}\,\mathrm{tr}[\tilde{\rho}\tilde{\sigma}]\right| \leq 2\|\mathrm{O}\|\left[(1-\epsilon)^{1/2}\,\epsilon^{1/2}\delta + \epsilon(1-\epsilon)\delta^2 + \epsilon^2\right]. \tag{4.36}$$

Next, to bound the factor $|\mathrm{tr}[\tilde{\rho}\tilde{\sigma}]|^{-1}$ in eq. 4.24, we apply the reverse triangle inequality to $|\mathrm{tr}[\tilde{\rho}\tilde{\sigma}]|$:

$$\begin{aligned}
\left|\mathrm{tr}[\tilde{\rho}\tilde{\sigma}]\right| &= \left|(1-\epsilon)^2 + 2\epsilon(1-\epsilon)\mathrm{Re}\Big(\sum_{j>0}\alpha_j\beta_j^*\mathcal{F}_{j0}^2\Big) + \epsilon^2\sum_{j,l>0}\alpha_j\alpha_l^*\beta_l\beta_j^*\mathcal{F}_{jl}^2\right| \\
&\geq \left|(1-\epsilon)^2 - 2\epsilon(1-\epsilon) - \epsilon^2\right|
\end{aligned} \tag{4.37}$$

where we used that $\delta \leq 1$. For $\epsilon < \sqrt{3/2} - 1$, the argument in eq. 4.37 is strictly positive, so we can remove the absolute value symbol. In this case, we have $\left|\mathrm{tr}[\tilde{\rho}\tilde{\sigma}]\right|^{-1} \in O(1)$. The dominant terms in the error eq. 4.24 are then

$$\text{error} \in O\big(\|\mathrm{O}\|(\epsilon^{1/2}\delta + \epsilon^2)\big) \tag{4.38}$$

as desired. □

We can verify that for $\delta \to 0$ we recover the error scaling with ϵ^2, as expected from perfect dephasing. To achieve the same scaling, it is in fact sufficient to choose $\delta \in O(\epsilon^{3/2})$.

4.3.3 Dephasing time for a smooth probability distribution

An upper bound on the absolute coefficients $|\mathcal{F}_{0j}|$ can be obtained as a functional of the distribution $P(\tau)$. We can thus redefine

$$\delta \coloneqq \max_{\Delta > \Delta_T} \left|\mathcal{F}[P](\Delta)\right| \tag{4.39}$$

where $\Delta_T < E_1 - E_0$ is a lower bound on the target hamiltonian ground state gap. In principle, different distributions can be chosen, as long as they have support on $[0, T_\mathrm{d}]$. This ensures we only need to evolve for positive times and the maximal dephasing time is T_d.

As an example, we might simply choose a uniform distribution $P(\tau) = 1/T_\mathrm{d}$ for $\tau \in [0, T_\mathrm{d}]$. As its Fourier transform is the cardinal sine function $\sin(x)/x$, we obtain $\delta \in O\big((\Delta T_\mathrm{d})^{-1}\big)$. However, discontinuities in P or its derivatives limit the asymptotic decay of $\mathcal{F}[P]$ to a polynomial. We can improve upon this without increasing the maximal evolution time by choosing a mollifier, i.e. a smooth distribution supported on $[0, T_\mathrm{d}]$.

To obtain the best possible asymptotic decay of the Fourier transform of a function $\mathcal{F}[f]$, we should choose f to be smooth. In fact, requiring the Fourier transform of f to decay as $\mathcal{F}[f](k) \in O\big(|k|^{-(r+1+\epsilon)}\big)$ (for any choice of $\epsilon > 0$) implies that

$$\exists L_r > 0 : \forall x \quad \left|\frac{\mathrm{d}^r f(x)}{\mathrm{d}x^r}\right| = \left|\int \mathrm{d}k\, \mathrm{e}^{\mathrm{i}kt} \underbrace{k^r \mathcal{F}[f](k)}_{\lesssim |k|^{-(1+\epsilon)}}\right| < L_r, \tag{4.40}$$

because $|k|^{-(1+\epsilon)}$ is absolutely integrable away from 0 and $\mathcal{F}[f](k)$ is bounded. This implies that f and all its derivatives up to order $r-1$ are Lipschitz continuous. Thus, to achieve a Fourier transform decaying faster than any polynomial $\mathcal{F}[f](k) = o(1/\mathrm{poly}(k))$, we have to choose $f(x) \in C_\infty$ a smooth function.

One smooth function with compact support is the bump function

$$f(x) = \begin{cases} \mathrm{e}^{-(1-x^2)^{-1}} & \text{if} \quad -1 < x < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.41}$$

We define its norm $\mathcal{N} \coloneqq \int_{-1}^{1} f(x)\,\mathrm{d}x \approx 2.25$. Based on this function, we define the probability distribution

$$P_{T_\mathrm{d}}(\tau) = \frac{2}{T_\mathrm{d}\mathcal{N}} f\left(2\frac{\tau}{T_\mathrm{d}} - 1\right) = \begin{cases} \frac{2}{T_\mathrm{d}\mathcal{N}} \exp\left(\left[4(\frac{\tau}{T_\mathrm{d}} - 1)\frac{\tau}{T_\mathrm{d}}\right]^{-1}\right) & \text{if} \quad 0 < \tau < T_\mathrm{d}, \\ 0 \quad \text{otherwise} \end{cases} \tag{4.42}$$

which is normalized, smooth, and has support on $[0, T_\mathrm{d}]$. The Fourier transform of this function can be estimated through the saddle point approximation. We build on the results of ref. Joh15 which provide a bound on the Fourier transform $\mathcal{F}[\mathcal{N}f]$ of the normalized $f(x)$:

$$\mathcal{F}[\mathcal{N}f](k) \approx 2\mathrm{Re}\left[\sqrt{\frac{-\mathrm{i}\pi}{\sqrt{2\mathrm{i}}}}\mathrm{e}^{\mathrm{i}k-\frac{1}{4}-\mathrm{i}\sqrt{k}}\right] k^{-\frac{3}{4}}\mathrm{e}^{-\sqrt{k}}. \tag{4.43}$$

We construct a monotonic envelope for this oscillating function by substituting the real part for an absolute value, and we perform a change of variables obtaining the (approximate) bound

$$\delta = \max_{\Delta>\Delta_T} |\mathcal{F}[P_{T_\mathrm{d}}](\Delta)| \approx \tilde{\delta} = \sqrt{\frac{8\pi}{\sqrt{\mathrm{e}}}} \left(T_\mathrm{d}\Delta_T\right)^{-3/4} \mathrm{e}^{-\sqrt{T_\mathrm{d}\Delta_T/2}}. \tag{4.44}$$

The validity of this bound is also verified numerically. This translates to a statement on the dephasing time T_d required to achieve a target dephasing performance δ for a given gap Δ_T between the ground state and the first excited state of H_T. Note that the inverse is defined in terms of the principal branch W_0 of the Lambert W function. We obtain

$$T_\mathrm{d} \approx \tilde{T}_\mathrm{d} = \frac{9}{2\,\Delta_T} W_0\left[\frac{2\sqrt{2}\,\pi^{1/3}}{3\,\mathrm{e}^{1/6}\,\delta^{2/3}}\right]^2 \in O\big(\log^2(\delta_{\mathrm{target}}^{-1})\Delta_T^{-1}\big). \tag{4.45}$$

Since we established earlier that $\delta \in O(\epsilon^{3/2})$ suffices to suppress the overall error quadratically, we have thus shown the following result.

4.3.2. Theorem. *Let the requirements of theorem 4.3.1 be satisfied. If the AEV estimator is chosen with the dephasing time probability distribution in eq. 4.42, then a dephasing time $T_d \approx \tilde{T}_d \in O(\Delta_T^{-1}\log^2[\epsilon^{-1}])$ is sufficient to achieve an estimator bias of at most $O(\epsilon^2)$.*

Often, one is interested in observables of states in gapped phases [Sac99], such that only the poly-logarithmic term contributes to a non-constant overhead. We include the dephasing time in the illustration in fig. 4.2(b). In fig. 4.2(c), we show a cartoon of a the lowest two eigenenergies of the instantaneous hamiltonian $\mathrm{H}(\lambda)$, showing a small adiabatic gap $\Delta_{\min}$ during the quasi-adiabatic sweep and a larger, possibly constant gap $\Delta_{\lambda=1}$ of the final hamiltonian $\mathrm{H}(1)$.

4.4 Comparison with standard adiabatic algorithm

We seek to compare the method proposed here with the trivial alternative for improving the performance of the adiabatic algorithm, which simply consists of

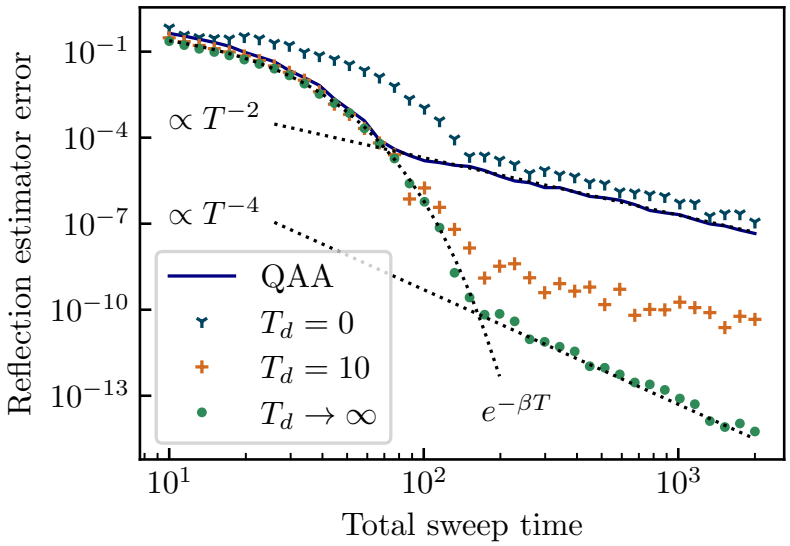


Figure 4.3. Comparing the QAA and AEV for different dephasing times T_d, as a function of the total sweep time T (QAA: $T = T_\mathrm{ad}$, AEV: $T = 2\,T_\mathrm{ad}$). AEV improves over simply doubling the QAA sweep time in the regime of polynomial error dependence. The respective estimator bias from $\langle E_0|\mathrm{O}|E_0\rangle$ is shown, where $\mathrm{O} = \mathrm{I} - 2|E_0\rangle\langle E_0|$ is a reflection on the target state. We perform density-matrix simulations; time-dependent evolution is implemented by Euler integration, ensuring a sufficiently small error when discretizing the sweep. Approximate dephasing is implemented with $P(\tau)$ as in eq. 4.42.

doubling the evolution time in the QAA. In the standard adiabatic theorem, there is a polynomial relationship between the accuracy and the evolution time [JRS07]. In principle, the adiabatic theorem can be improved towards an exponential error dependence by assuming a sufficiently smooth schedule with vanishing derivatives at the beginning and end of the schedule [WB12]. However, this is at the cost of passing the minimal spectral gap at a faster rate, which, in general, leads to more transitions.

Regarding our method, we therefore conclude that if the error dependence was indeed exponential, as in a Landau-Zener problem, the AEV would yield a performance comparable to the QAA with double the evolution time. Compared to the standard theorems with a polynomial dependence, our method improves up to quadratically.

4.4.1 Numerical simulation: ground state reflection in an Ising model

For the sake of concreteness, we include numerical benchmarks in fig. 4.3 showing an advantage for preparing the ground state of a transverse field Ising model. We consider the reflection operator $\mathrm{O} = \mathrm{I} - 2|E_0\rangle\langle E_0|$ on the ground state

$|E_0\rangle$ and show the error from the expectation value $\langle E_0|\mathrm{O}|E_0\rangle$ as a function of the total sweep time. In appendix 4.4.2, we include additional numerics that consider the average magnetization of the state as another observable with practical relevance. For the QAA, the sweep time is simply the time of one forward sweep. We compare this with the error we obtain using the AEV for three different scenarios: no dephasing ($T_d = 0$), approximate dephasing ($T_d = 10$) and perfect dephasing $T_d \to \infty$. Here, the total sweep time is the sum of the forward and the backward sweeps. The hamiltonian chosen for the benchmark is an Ising model with a transverse and a longitudinal field, which is a non-integrable model. The quasi-adiabatic sweeps linearly interpolate from $\mathrm{H}_0 = \sum_{j=1}^{5} \mathrm{X}_j$ to $\mathrm{H}_T = 0.2 \sum_{j=1}^{5} \mathrm{Z}_j - \sum_{j=1}^{4} \mathrm{Z}_j \mathrm{Z}_{j+1}$. We observe an exponential scaling, in the so-called Landau-Zener regime, transitioning into an inverse-quadratic scaling for longer times (cf. ref. RPL10). For these longer times, the AEV improves up to quadratically (error $\propto T^{-4}$) over the QAA (error $\propto T^{-2}$) if the detuning time is sufficiently large. The source code for the numerical simulations is published on Zenodo [Pol24].

4.4.2 Numerical simulation: magnetization in an Ising model

In addition, we include simulations where we consider the total magnetization

$$M = \sum_{i=1}^{n} \mathrm{Z}_i \tag{4.46}$$

as the observable in the AEV protocol. The considered model is the same Ising model as for fig. 4.3, where the reflection along the target ground state was probed instead of the magnetization. We observe a strong similarity in the behaviour of the error between both observables.

Note that the error on the measured observable in AEV with approximate dephasing (see sec. 4.3.1) can be positive or negative, thus the results can overshoot variational bounds. This is because the result of AEV is obtained as a fraction between expectation value of two circuits, either of which having independent variationally-bounded errors. In practice, in fig. 4.4 we observe a value of the magnetization overshooting the bound of $\langle M \rangle \leq 5$ only for $T_d = 0$. A modest amount of dephasing suppresses the error in the echo circuit, suppressing the oscillations in the sign of the error and recovering a variationally-bounded result.

4.5 Discussion and practical considerations

In this chapter, we have introduced *adiabatic echo verification* (AEV), a scheme to mitigate the coherent errors that characterize adiabatic state preparation.

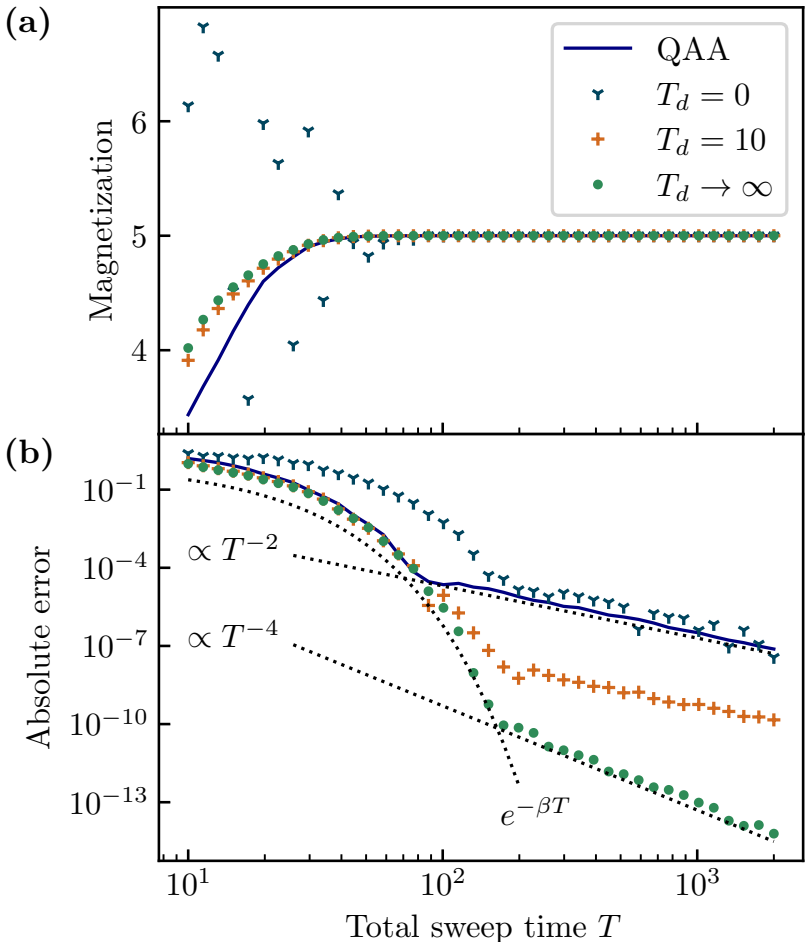


Figure 4.4. Numerical simulation of the AEV protocol on Ising chain of five spins; this is the same model as in fig. 4.3, but here the measured observable is the total magnetization $M = \langle \sum_{j=1}^{5} \mathrm{Z}_j \rangle$. **(a)** Expectation values for the magnetization M, comparing the quantum adiabatic algorithm (QAA) with the AEV for different dephasing times. While all scenarios converge asymptotically towards the true value for sufficiently long sweep times, they differ in their convergence behavior. **(b)** Absolute error of the different scenarios with the true value. The convergence of the magnetization behaves qualitatively as the reflection operation in fig. 4.3.

Our method is tailored to current quantum devices, which lack the possibility to correct errors. AEV requires doubling the circuit time compared to standard adiabatic state preparation, but improves up to quadratically in the estimator bias. The additional features of the protocol are the following. First, in order to implement the verification part of the circuit, the path of the quasi-adiabatic evolution is simply reversed. Moreover, we show how the dephasing operation can be approximately implemented with positive-time dynamics. Hence, only positive-time evolution is required in AEV. This makes our method suitable for

quantum computers that operate in a hybrid mode of digital gates and analog simulation. Rydberg atom arrays have recently demonstrated such capabilities [BLS+22, EBK+23].

Additionally, AEV naturally mitigates non-coherent hardware noise through echo verification. While this chapter focuses on the theoretical analysis of the suppression of algorithmic noise, the hardware error mitigation power of echo verification has been previously demonstrated in literature. Numerical studies suggest that echo verification can effectively mitigate the effect of any single error, obtaining a quadratic improvement on the error on an observable from $O(p)$ to $O(p^2)$ with p the error probability [OPR+21]. The effectiveness of echo verification for hardware error mitigation has also been demonstrated for a 10-qubit chemical simulation problem on a superconducting gate-based device [OAG+23].

We note that our method is compatible with arbitrary sweep profiles in the QAA. This is especially helpful as it is well known that slowing down the adiabatic sweep at the position of the minimum spectral gap mitigates transitions out of the ground state [RC02, STC22]. More generally, our technique can be applied to other coherent approximate state preparation approaches, such as variational quantum algorithms (VQAs) [CAB+21]. This applies to VQAs that prepare a pure state heuristically by a parametrized operation $\mathrm{U}(\boldsymbol{\theta})$ aiming at approximating the desired ground state. By dephasing the prepared state and using the echo verification technique, unpreparing the state with $\mathrm{U}(\boldsymbol{\theta})^\dagger$, we expect that the performance of VQAs can be improved.

We note that the control-free versions of echo verification [OPR+21, OAG+23], which employ a reference state instead of a control qubit, are not naively available for AEV. This is due the dephasing channel annihilating coherences between the reference state and the state of interest. Recently, a method for rescaling survival probabilities was considered that has similarities with control-free echo verification [YCCV+23]. While their unnormalized estimator $\mathrm{tr}[\rho \mathrm{O} \rho \mathrm{O}^\dagger]$ differs from the echo verification counterpart $\mathrm{tr}[\mathrm{O}\rho^2]$, it would still allow for mitigating errors for certain interesting observables such as out-of-time-order correlators (OTOCs) [MRQ+21]. Not requiring the implementation of a controlled operation can significantly simplify experiments. This is why an extension of AEV without a control qubit is an interesting direction for future work; a related idea has been recently proposed [SNKM24]. Another promising research direction is the combination of AEV with other purification-based error mitigation methods such as virtual state distillation [Koc21, HMO+21]. Using multiple copies of the quasi-adiabatically prepared state, further improvements for surpressing errors seem possible.

Chapter 5

Gate-based counterdiabatic driving with complexity guarantees

5.1 Introduction

Adiabatic quantum computing (AQC) is a paradigm of quantum computation that leverages the principles of adiabatic processes in quantum mechanics to solve computational problems [AL18]. AQC relies on the adiabatic theorem, which states that a quantum system approximately remains in its instantaneous eigenstate if a given time-dependent hamiltonian that governs its energy levels is changed slowly enough [Kat50]. In this sense, AQC is a conceptually straightforward way to prepare ground states or other eigenstates of complex systems, by starting from an eigenstate of a simple hamiltonian (e.g. a product state) and slowly time evolving towards the complex hamiltonian.

Contrary to AQC, counterdiabatic driving (CD) aims to prepare target states through a fast time evolution, actively suppressing excitations that would typically accompany such rapid driving. Traditional adiabatic processes require slow evolution to maintain the system in its instantaneous eigenstate, which can be impractically long. CD involves adding auxiliary, non-adiabatic control fields to counteract the diabatic transitions, thereby simulating the effect of a slow adiabatic process in a (much) shorter time [KSMP17]. It is a specific method within a broader class of techniques known as "shortcuts to adiabaticity" [GORK$^+$19], designed to achieve the same goal of mimicking adiabatic evolution on a small time scale. Its original formulation is due to Demirplak & Rice [DR03, DR05, DR08]; in this formalism, the central object is the adiabatic gauge potential (AGP), which is an auxiliary field that is added to the system hamiltonian and cancels what is akin to a Coriolis force in the rotating system eigenbasis, thereby supressing all excitations. The issue however is that determining an AGP exactly for a given system is computationally difficult – possibly as difficult, if not more, as preparing the eigenstate of interest – and analytical expressions are only known for very specific systems [KSMP17]. This is due to the fact that, typically, the AGP is

highly nonlocal and contains high-rank (two-body, three-body etc.) interactions [dCRZ12, SOMdC14, Tak13, Dam14]; moreover, if the spectrum of the system is gapless, the AGP does not exist [Jar95].

For this reason, much effort has been focussed on computational methods that approximate the AGP. In this context, a pivotal role has been played by variational methods [SP17]; these methods assume a parametrised ansatz that is optimised with respect to an action function that approaches zero as the ansatz approaches the exact AGP. This approach has enabled approximate CD in a large number of different contexts, leading to improved fidelities as compared to standard AQC [ZJN$^+$20, PCFL20, XSY22, HML22, ML22, CPDD23, SB23]. Among these variational methods, gate-based (digitised) approaches have been proposed as well [HPD$^+$21], where the dynamics is trotterised for implementation on a gate-based quantum computer. However, the limit to the variational approach lies in the fact that evaluating the action function itself becomes harder when higher-rank (two-body, three-body etc.) interactions are taken into account in the ansatz; therefore only ansätze with low-rank interactions are computationally feasible for variational methods.

Further improvements to this scheme were made through the invention of a nested commutator ansatz to the AGP [CPSP19] and subsequent Krylov space methods to solve for the optimal coefficients in this ansatz [Bha23, TdC24]. Here, Krylov spaces are subspaces of operator space that are specifically chosen such that the coefficients can be found by solving a linear equation, which avoids optimisation heuristics with unpredictable running times. At the same time, this approach sheds more light on the computational complexity of computing the AGP. Numerical experiments [Bha23] suggest that the relevant search space for weakly interacting and integrable systems is a small subspace in operator space, so that the AGP is easily approximated with an ansatz containing few coefficients; on the other hand, for strongly interacting, chaotic systems, this tends not to be the case, and the whole operator space may need to be explored. In this sense, the complexity of CD is tightly connected to quantum Krylov complexity [PCA$^+$19, BRSS19, BCNP22, RSGSS21, RSGSS22, NMRMA$^+$24]. More recently, interesting suggestions have been put forward to apply tensor network approaches for approximating the gauge potential by expressing it as a matrix product operator [KFS24, MKL24].

The main issue, however, in these developments, is that a precise description of the computational complexity of CD has so far been lacking. In particular, while the implementation of the aforementioned variational methods is straightforward, none of these have been accompanied with any rigorous computational complexity statement. The reason for this is twofold: (i) these approaches do not present any clear-cut quantification of the accuracy to which an AGP for a given system must be approximated in order to achieve a desired fidelity; (ii) the use of optimisation schemes makes complexity analysis of several of the methods prohibitively difficult.

In this chapter, we address the complexity issue by asking the following question: does the computational effort we put ifnto determining the AGP, in order to boost the fidelity of the output state with the target state, compare favourably to that of AQC with simply a longer evolution time? Following up on this question, how should we define computational complexity so that a fair comparison is made between the two methods? After all, the complexity of AQC is typically measured in terms of physical evolution time, whereas that of CD, in the light of aforementioned variational and Krylov space methods, boils down to the computing time a classical machine requires to solve a (linear) equation for the ansatz coefficients. To address this question, we place AQC and CD on an equal footing by considering the complexity of both approaches in terms of the number of quantum gates as required by a fault-tolerant quantum computer. To this end, we develop a gate-based quantum algorithm for CD (sections 5.3 and 5.4) that can be compared to gate-based version AQC (section 5.6). This method is nonvariational, and differs from the variational digitised approach previously put forward [HPD+21] in that it evades the difficulty of evaluating complexity that comes with variational methods. We give upper bounds on the number of quantum gates required for both gate-based CD and gate-based AQC. In this way, we contribute to the answering of the posed questions.

A crucial insight here is that the complexity of AQC is state dependent, in the sense that it is known to scale with the inverse of the spectral gap around the instantaneous eigenstate of interest, while that of currently known approaches to CD is not. So far, CD approaches have aimed to approximate the exact AGP, which cancels excitations on the entire spectrum; however, if the objective is to prepare only a single eigenstate, this is inefficient, since only the excitations from the eigenstate of interest need to be suppressed. We remedy this by basing our CD algorithm on an approximate AGP that effectively ignores excitations below a certain energy gap cutoff; this cutoff is then chosen small enough so that at least the transitions from the relevant eigenstate are suppressed. In this way, the gap around this eigenstate enters into the complexity description of CD. While the idea of an energy cutoff to suppress large-gap transitions itself is not new, and appears for example in the theory of quasi-adiabatic continuation [HW05, Has10, MAAV14], it has so far been overlooked in studies on the computational complexity of implementing counterdiabatic driving.

The main contribution of this chapter is theorem 5.4.3, which states the gate-based CD algorithm and an upper bound on its worst-case quantum gate complexity. The gate complexity bound we find is $\tilde{O}(\Delta^{-(3+o(1))}\epsilon^{-(1+o(1))})$, where Δ is the minimum energy gap along the path, and ϵ is the square root of the infidelity between the output state and the target state. Since Δ often decreases exponentially in the system size, especially in complex systems, the Δ dependence is usually the dominant factor in the complexity of AQC. The formalism we introduce allows for the evaluation of CD in terms of Δ as well.

This chapter is structured as follows. In section 5.2, we discuss the formalism

of counterdiabatic driving and detail the Lie-Suzuki-Trotter (LTS) decomposition of a path-ordered exponential that translates a physical time description to a gate-based description. Section 5.3 introduces the regularised truncated AGP, which brings the aforementioned gap cutoff into play and is at the heart of the gate-based CD algorithm. In section 5.4, we construct the gate-based CD algorithm, which is a simulation, in the form of a LTS decomposition, of the path-ordered exponentiation of the regularised truncated AGP. Section 5.5 discusses a randomised approach to gate-based CD which, depite its inferior complexity upper bound, may be of independent interest. In section 5.6, we compare the CD result to gate-based AQC by applying LTS simulation to standard adiabatic theorems. We conclude in section 5.7.

5.2 Preliminaries

5.2.1 Counterdiabatic driving

Consider the time-dependent Schrödinger equation for a system parametrised by a time-dependent parameter $\lambda(t)$:

$$\mathrm{i}\,\mathrm{d}_t|\psi(t)\rangle = \mathrm{H}\big(\lambda(t)\big)|\psi(t)\rangle. \tag{5.1}$$

Clearly, the solution $|\psi(t)\rangle$ will be dependent on $\lambda(t)$. Let $\mathrm{U}\big(\lambda(t),\lambda(0)\big)$ be the unitary that connects the instantaneous eigenstates $|n\big(\lambda(t)\big)\rangle$ of $\mathrm{H}\big(\lambda(t)\big)$ with those of $\mathrm{H}\big(\lambda(0)\big)$ – that is, let $\mathrm{H}\big(\lambda(t)\big)$ admit the spectral decomposition

$$\begin{aligned}\mathrm{H}\big(\lambda(t)\big) &= \sum_n E_n\big(\lambda(t)\big)\,|n\big(\lambda(t)\big)\rangle\langle n\big(\lambda(t)\big)| \\ &= \sum_n E_n\big(\lambda(t)\big)\,\mathrm{U}\big(\lambda(t),\lambda(0)\big)|n\big(\lambda(0)\big)\rangle\langle n\big(\lambda(0)\big)|\mathrm{U}^\dagger\big(\lambda(t),\lambda(0)\big).\end{aligned} \tag{5.2}$$

with $E_n\big(\lambda(t)\big)$ the corresponding energies (eigenvalues). We can view this system in a rotating (λ-dependent) basis by writing $|\psi(t)\rangle = \mathrm{U}\big(\lambda(t),\lambda(0)\big)|\tilde{\psi}\rangle$, where $|\tilde{\psi}\rangle$ is some reference state that is independent of λ. In the rotating basis then, the time-dependent Schrödinger equation becomes

$$\mathrm{i}\,\mathrm{d}_t|\tilde{\psi}\rangle = (\mathrm{U}^\dagger\mathrm{H}\mathrm{U} - \mathrm{i}\dot{\lambda}\mathrm{U}^\dagger\partial_\lambda\mathrm{U})|\tilde{\psi}\rangle. \tag{5.3}$$

Since $\mathrm{U}^\dagger\mathrm{H}\mathrm{U}$ is diagonal in the rotating eigenbasis of H, any nonadiabatic transitions can only be generated by the second term of the effective hamiltonian in eq. 5.3. Thus we can suppress all nonadiabatic transitions by time evolving with a different hamiltonian that cancels out this second term in the rotating frame:

$$\mathrm{U}^\dagger\mathrm{H}_{\mathrm{CD}}\mathrm{U} = \mathrm{U}^\dagger\mathrm{H}\mathrm{U} + \mathrm{i}\dot{\lambda}\mathrm{U}^\dagger\partial_\lambda\mathrm{U}, \tag{5.4}$$

or, going back to the fixed (λ-independent) basis – i.e. conjugating with U on both sides, we find

$$\mathrm{H_{CD}}(t) = \mathrm{H}\big(\lambda(t)\big) + \dot{\lambda}(t)\mathrm{A}\big(\lambda(t)\big) \tag{5.5}$$

where the operator $\mathrm{A} = \mathrm{i}\partial_\lambda \mathrm{U}\mathrm{U}^\dagger$ is called the *adiabatic gauge potential* (AGP). Since evolution under $\mathrm{H_{CD}}$, known as the *counterdiabatic hamiltonian*, causes no transitions into excited states, one may carry out a perfect evolution at any speed:

$$\mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) = \mathcal{T} \exp\left[-\mathrm{i}\int_0^T \mathrm{d}t\, \mathrm{H_{CD}}(t)\right] \qquad \text{for any } T \in [0, \infty). \tag{5.6}$$

The adiabatic gauge potential, then, can be viewed simply as the derivative operator $\mathrm{i}\partial_\lambda$:

$$\begin{aligned}\langle m(\lambda)|\mathrm{A}(\lambda)|n(\lambda)\rangle &= \mathrm{i}\langle m(\lambda)|\partial_\lambda \mathrm{U}(\lambda, \lambda_\mathrm{i})|n(\lambda_\mathrm{i})\rangle \\ &= \mathrm{i}\langle m(\lambda)|\partial_\lambda|n(\lambda)\rangle,\end{aligned} \tag{5.7}$$

and therefore generates motion in λ space,

$$\mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) = \mathcal{T} \exp\left[-\mathrm{i}\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\, \mathrm{A}(\lambda)\right] \tag{5.8}$$

which is simply a special case of eq. 5.6, namely the limit $T \to 0$. From the Hellmann-Feynman theorem, A may be expressed in the eigenbasis of $\mathrm{H}(\lambda)$, and has the off-diagonal matrix elements

$$\langle m(\lambda)|\mathrm{A}(\lambda)|n(\lambda)\rangle = \frac{\langle m(\lambda)|\partial_\lambda \mathrm{H}(\lambda)|n(\lambda)\rangle}{\mathrm{i}\omega_{mn}(\lambda)} \tag{5.9}$$

where $\omega_{mn} = E_m(\lambda) - E_n(\lambda)$. The diagonal elements are arbitrary since they depend on a gauge choice (namely, the phase of the eigenstates of $\mathrm{H}(\lambda)$) – this makes the AGP nonunique for a given hamiltonian. Lastly, the form of the AGP in eq. 5.9 may be expressed as a time integral [CPSP19, PCC+20]

$$\mathrm{A}(\lambda) = \frac{1}{2}\lim_{\eta\to 0}\int_{-\infty}^{\infty} \mathrm{d}\tau\, \mathrm{e}^{-\eta|\tau|}\,\mathrm{sgn}(\tau)\,\mathrm{e}^{-\mathrm{i}\mathrm{H}(\lambda)\tau}\partial_\lambda \mathrm{H}(\lambda)\mathrm{e}^{\mathrm{i}\mathrm{H}(\lambda)\tau} \tag{5.10}$$

since

$$\begin{aligned}\langle m(\lambda)|\mathrm{A}(\lambda)|n(\lambda)\rangle &= \frac{1}{2}\langle m(\lambda)|\partial_\lambda \mathrm{H}(\lambda)|n(\lambda)\rangle \lim_{\eta\to 0}\int_{-\infty}^{\infty} \mathrm{d}\tau\, \mathrm{e}^{-\eta|\tau|}\,\mathrm{sgn}(\tau)\,\mathrm{e}^{-\mathrm{i}\,\omega_{mn}(\lambda)\tau} \\ &= \langle m(\lambda)|\partial_\lambda \mathrm{H}(\lambda)|n(\lambda)\rangle \lim_{\eta\to 0}\frac{\omega_{mn}(\lambda)}{\mathrm{i}(\omega_{mn}(\lambda)^2 + \eta^2)} \\ &= \frac{\langle m(\lambda)|\partial_\lambda \mathrm{H}(\lambda)|n(\lambda)\rangle}{\mathrm{i}\omega_{mn}}.\end{aligned} \tag{5.11}$$

Note that the limit $\eta \to 0$ must be taken after integration for the integral to converge.

5.2.2 Lie-Trotter-Suzuki decompositions

Since the main objective of this chapter is to describe a counterdiabatic driving protocol that runs on a digital quantum computer, a significant portion of the chapter will be dedicated to simulation – that is, the approximation as a sequence of quantum gates – of ordered exponentials. The problem of implementing ordered exponentials of hermitian operators is far from new, and has been discussed in chapter 2. In this chapter, we focus our attention on Lie-Trotter-Suzuki (LTS) decompositions [Suz92, WBHS10]. These methods are competitive in terms of efficiency in comparison to Dyson series [LW19, KSB19a] or clock construction methods [WWRL22] while being independent of any time-dependent oracular access to the hamiltonian. LTS formulae approximate a time-ordered exponential $\mathrm{U}(t_\mathrm{f}, t_\mathrm{i}) = \mathcal{T}\exp[-\mathrm{i}\int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}t\mathrm{H}(t)]$ as a product $\mathrm{U}(t_\mathrm{f}, t_\mathrm{i}) \approx \prod_j \mathrm{e}^{-\mathrm{i}\mathrm{H}_{i_j}(t_j)\delta t_j}$. Each factor in this product may then be simulated through time-*independent* hamiltonian simulation methods such as qubitisation [LC19b], or in the case of LCU, as a multi-qubit rotation gate.

LTS formulae build on the idea of trotterisation, but form more refined approximations to the operator exponential (in the time-independent case) or the path-ordered exponential (in the time-dependent case) whose implementation is more efficient in terms of the simulation time $t_\mathrm{f} - t_\mathrm{i}$ and the precision ϵ. These decompositions may be defined for any order $k \in \mathbb{N}_{>0}$, with every order representing a more fine-grained approximation, and achieve $O(\delta t^{2k+1})$ error scaling. Again dividing a large δt into r smaller segments, this results in an r scaling $O(\delta t^{1+1/2k}\epsilon^{1/2k})$ to achieve precision ϵ.[1]

While LTS formulae were originally defined only for time-independent settings [Suz92], the concept of LTS formulae naturally extends to time-dependent operator exponentials. Nonetheless, general error bounds were found only years after the original formulation, by Wiebe et al. [WBHS10]. Their result is very similar to the path-independent case, but imposes conditions on the differentiability of the operator exponents. We base the analysis in this section on their work, accurately describing the decomposition of the ordered exponential $\mathrm{U}(t_\mathrm{f}, t_\mathrm{i}) = \mathcal{T}\exp[-\mathrm{i}\int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}t\,\mathrm{H}(t)]$ and the requirements to achieve a given precision.

The first-order single-segment LTS approximation $\tilde{\mathrm{U}}_{1,1}(t_\mathrm{f}, t_\mathrm{i})$ of $\mathrm{U}(t_\mathrm{f}, t_\mathrm{i})$, for the operator $\mathrm{H}(t) = \sum_{i=1}^{\ell} \mathrm{H}_i(t)$, is given by

$$\tilde{\mathrm{U}}_{1,1}(t_\mathrm{f}, t_\mathrm{i}) = \left(\prod_{i=1}^{\ell} \exp[-\mathrm{i}\,\mathrm{H}_i(t_\mathrm{i} + \delta t_1/2)\,\delta t_1/2]\right)\left(\prod_{i=\ell}^{1} \exp[-\mathrm{i}\,\mathrm{H}_i(t_\mathrm{i} + \delta t_1/2)\,\delta t_1/2]\right) \tag{5.12}$$

where $\delta t_1 = t_\mathrm{f} - t_\mathrm{i}$. Higher-order LTS product formulas are built by repeatedly

[1] The Trotter decomposition $\mathrm{e}^{-\mathrm{i}(\mathrm{O}+\mathrm{Q})\delta t} = (\mathrm{e}^{-\mathrm{i}\mathrm{O}\delta t/r}\mathrm{e}^{-\mathrm{i}\mathrm{Q}\delta t/r})^r + O(\delta t^2/r)$ may be viewed as analogous to a LTS formula of order $k = 1/2$.

applying the following recursive process to $\tilde{\mathrm{U}}_{1,1}$:

$$\begin{aligned}\tilde{\mathrm{U}}_{k+1,1}(t+\delta t,t) = \tilde{\mathrm{U}}_{k,1}(t+\delta t,t+[1-s_k]\delta t)\,\tilde{\mathrm{U}}_{k,1}(t+[1-s_k]\delta t,t+[1-2s_k]\delta t)\\ \times \tilde{\mathrm{U}}_{k,1}(t+[1-2s_k]\delta t,t+2s_k\delta t)\,\tilde{\mathrm{U}}_{k,1}(t+2s_k\delta t,t+s_k\delta t)\\ \times \tilde{\mathrm{U}}_{k,1}(t+s_k\delta t,t)\end{aligned} \tag{5.13}$$

with $s_k = (4-4^{1/(2k+1)})^{-1}$. Now if every $\mathrm{H}_i(t)$ is $2k$ times differentiable in t, then the error is bounded as in the path-independent case:

$$\|\mathrm{U}(t_\mathrm{f},t_\mathrm{i}) - \tilde{\mathrm{U}}_{k,1}(t_\mathrm{f},t_\mathrm{i})\| \in O(\delta t^{2k+1}). \tag{5.14}$$

This means that if δt is small, then the approximation error decreases monotonically with the order k. If δt is not small, we divide it into r segments and apply a LTS product formula for each segment:

$$\tilde{\mathrm{U}}_{k,r}(t+\delta t,t) = \prod_{s=1}^{r} \tilde{\mathrm{U}}_k(t+s\,\delta t/r, t+(s-1)\delta t/r). \tag{5.15}$$

We refer to this decomposition as a k^{th}-order r-segment LTS formula. The work of Wiebe et al. [WBHS10] now provides a condition on r that is sufficient to upper bound the spectral distance between an ordered exponential of a sum of operators and its LTS simulation.

5.2.1. Lemma (Wiebe et al. [WBHS10], theorem 1). *Let* $\mathrm{H}(t) = \sum_i \mathrm{H}_i(t)$ *be defined on the interval* $[t_\mathrm{i}, t_\mathrm{f}]$ *such that each* $\mathrm{H}_i(t)$ *is hermitian and* $2k$ *times differentiable on the entire interval. Let* $\delta t = t_\mathrm{f} - t_\mathrm{i}$. *Furthermore, suppose*

$$\max_{p=0,1,\ldots,2k}\left[\max_{t\in[t_\mathrm{i},t_\mathrm{f}]}\left(\sum_i \left\|\frac{\partial^p}{\partial t^p}\mathrm{H}_i(t)\right\|\right)^{1/(p+1)}\right] \leq \Lambda. \tag{5.16}$$

If $\epsilon \leq \min\{(9/10)(5/3)^k \Lambda\delta t, 1\}$, *then the spectral distance between the ordered exponential* $\mathcal{T}\exp[-\mathrm{i}\int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}t\,\mathrm{H}(t)]$ *and its* k^{th}*-order* r*-segment LTS decomposition is at most* ϵ, *provided that*

$$r \geq 5k\Lambda\delta t\left(\frac{5}{3}\right)^k\left(\frac{\Lambda\delta t}{\epsilon}\right)^{1/2k}. \tag{5.17}$$

5.3 Regularised truncated gauge potential

In the following sections, we will describe our approach to a gate-based counterdiabatic driving algorithm. We will base our methods on the time-integral expression for the adiabatic gauge potential from eq. 5.10. However, in its stead, we will use its regularised truncated version $\mathrm{A}_{\eta,a}$. Here, regularised means that

we fix an $\eta > 0$ instead of taking the limit $\eta \to 0$; furthermore, we restrict the integration to a bounded interval $[-a, a]$. That is:

$$\mathrm{A}_{\eta,a}(\lambda) = \frac{1}{2}\int_{-a}^{a} \mathrm{d}\tau\, \mathrm{e}^{-\eta|\tau|}\, \mathrm{sgn}(\tau) \mathrm{e}^{-\mathrm{i}\mathrm{H}(\lambda)\tau} \partial_\lambda \mathrm{H}(\lambda) \mathrm{e}^{\mathrm{i}\mathrm{H}(\lambda)\tau}. \tag{5.18}$$

One may view this approximation choice as a straightforward way to get rid of unworkable infinities. But there is a more intuitive reason to regularise the AGP: considering the matrix elements of A in the energy eigenbasis (eq. 5.11), one would intuitively expect counterdiabatic driving with the regularised AGP to suppress only those level transitions where $\eta^2 \ll \omega_{mn}^2$ [PCC+20]. In this sense, η can be viewed as a gap cutoff. Since we're only interested in suppressing the level transitions from the n-th eigenstate, a cutoff $\eta \sim \Delta_n^\nu$ for some $\nu > 0$, with Δ_n the minimum gap around the n-th eigenstate, should suffice. This energy cutoff η also turns out to be related to the cost of implementing approximate CD with the regularised AGP: as we will show later, the algorithm complexities grow with η^{-1} and a. Besides that, there is a connection (albeit qualitative) between η^{-1} and the thermodynamic cost associated with CD [CD17, ZCDCP16]. In these works, the norm of the AGP is introduced as a cost of implementing CD; at the same time, one can easily see that the spectral norm of the regularised truncated AGP satisfies $\|\mathrm{A}_{\eta,a}\| \leq \eta^{-1}(1 - \mathrm{e}^{-\eta a})\|\partial_\lambda \mathrm{H}\| \leq \eta^{-1}\|\partial_\lambda \mathrm{H}\|$.

For the remainder of the chapter, we will work in the limit $T \to 0$ (eq. 5.8). The most important reason for this is that, in the proofs of the error bounds that follow, the system hamiltonian always cancels out (see eq. 5.21 in the proof of lemma 5.3.1). As such, working in the limit $T \to 0$ and thereby leaving out the system hamiltonian merely simplifies the calculations. Nonetheless, one might argue that a better complexity can be achieved by choosing an optimal T which is not reflected in the presented bounds. We expect that such an optimal T^* would be at least on the order $T^* \in \Omega(\Delta^{-1})$, and possibly $T^* \in \Omega(\|\partial_\lambda \mathrm{H}\|_{\infty,\infty}/\Delta^2)$. (Larger times essentially revert the protocol to adiabatic evolution.) In such cases, we expect the added complexity of trotterising H to be asymptotically subdominant to that of CD in the limit $T \to 0$, since the complexity of AQC at a time scale T grows almost linearly in T (see proposition 5.6.2 in section 5.6). However, we do not expect such schemes to provide significant gains over either our CD algorithm or AQC, since CD in the $T \to 0$ limit can be essentially be viewed as AQC in a different coordinate system (i.e. a coordinate change from time to λ) and a $T > 0$ setting in a sense interpolates between the two.

We proceed to establish a more precise relation between η, a and Δ_n, which puts conditions on η and a for arbitrary approximation error.

5.3.1. Lemma. *Let* $\mathrm{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) = \mathcal{T}\exp[-\mathrm{i}\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\, \mathrm{A}_{\eta,a}(\lambda)]$, *where* $\mathrm{A}_{\eta,a}(\lambda)$ *is defined as in eq. 5.18, and let* $|n(\lambda_\mathrm{i})\rangle$ *be the* n*-th eigenstate of* $\mathrm{H}(\lambda_\mathrm{i})$. *Then*

$$\|(\mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \mathrm{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \leq \epsilon \tag{5.19}$$

if we take

$$\eta = \frac{1}{\sqrt{2}}\Delta_n^{3/2}\epsilon^{1/2}\|\partial_\lambda \mathrm{H}\|_{n,1}^{-1/2}; \qquad a = \frac{1}{\eta}\log\left(2\frac{\Delta_n+\eta}{\Delta_n\epsilon\eta}\|\partial_\lambda \mathrm{H}\|_{n,1}\right) \in \tilde{O}(\eta^{-1}). \tag{5.20}$$

Proof.
We first bound the unitary error by an error in the AGP:

$$\begin{aligned}
&\|(\mathrm{U}(\lambda_\mathrm{f},\lambda_\mathrm{i}) - \mathrm{U}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \\
&\quad = \|(\mathrm{U}^\dagger_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i})\mathrm{U}(\lambda_\mathrm{f},\lambda_\mathrm{i}) - \mathrm{I})|n(\lambda_\mathrm{i})\rangle\| \\
&\quad = \left\| \int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\, \frac{\mathrm{d}}{\mathrm{d}\lambda}(\mathrm{U}^\dagger_{\eta,a}(\lambda,\lambda_\mathrm{i})\mathrm{U}(\lambda,\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle \right\| \\
&\quad = \left\| \int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\, \mathrm{U}^\dagger_{\eta,a}(\lambda,\lambda_\mathrm{i})\left[\mathrm{A}(\lambda) - \mathrm{A}_{\eta,a}(\lambda)\right]\mathrm{U}(\lambda,\lambda_\mathrm{i})|n(\lambda_\mathrm{i})\rangle \right\| \\
&\quad \leq \int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\, \|(\mathrm{A}(\lambda) - \mathrm{A}_{\eta,a}(\lambda))|n(\lambda)\rangle\|
\end{aligned} \tag{5.21}$$

where the last inequality follows from a triangle inequality on the integral. To evaluate the AGP error, we make use of the energy eigenbasis expansion (omitting the λ argument for brevity and using that $\langle k|\partial_\lambda k\rangle = 0$)

$$\mathrm{e}^{-\mathrm{i}\mathrm{H}\tau}\partial_\lambda \mathrm{H}\,\mathrm{e}^{\mathrm{i}\mathrm{H}\tau} = \sum_k \partial_\lambda E_k |k\rangle\langle k| + \sum_{k;\, m\neq k} \mathrm{e}^{-\mathrm{i}(E_m - E_k)\tau}|m\rangle\langle m|\partial_\lambda \mathrm{H}|k\rangle\langle k|. \tag{5.22}$$

Since $\mathrm{e}^{-\eta|\tau|}$ is symmetric in τ, it follows that[2]

$$\sum_k \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{d}\tau\, \mathrm{sgn}(\tau)(1 - \mathrm{e}^{-\eta|\tau|}\mathbb{1}_{\tau\in[-a,a]})\partial_\lambda E_k |k\rangle\langle k| = 0. \tag{5.23}$$

Furthermore, for any eigenstate $|n\rangle$,

$$\begin{aligned}
&\sum_{\substack{k \\ m\neq k}} \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{d}\tau\, \mathrm{sgn}(\tau)(1 - \mathrm{e}^{-\eta|\tau|}\mathbb{1}_{\tau\in[-a,a]})\mathrm{e}^{-\mathrm{i}(E_m - E_k)\tau}|m\rangle\langle m|\partial_\lambda \mathrm{H}|k\rangle\langle k|n\rangle \\
&= -\mathrm{i}\sum_{m\neq n} c_{\eta,a}(\omega_{mn})\langle m|\partial_\lambda \mathrm{H}|n\rangle\,|m\rangle
\end{aligned} \tag{5.24}$$

where

$$c_{\eta,a}(\omega) := \frac{1}{\omega} - \frac{\omega}{\omega^2+\eta^2} + \mathrm{e}^{-a\,\eta}\frac{\cos(a\omega)\omega + \sin(a\omega)\eta}{\omega^2+\eta^2}. \tag{5.25}$$

[2]We abuse notation in eqs. 5.23 and 5.24 since $\int_{-\infty}^{\infty}\mathrm{d}\tau\,\mathrm{sgn}\,\tau$ does not converge. The implied meaning is $\lim_{\mu\to 0}\int_{-\infty}^{\infty}\mathrm{d}\tau\,\mathrm{e}^{-\mu|\tau|}\,\mathrm{sgn}\,\tau$.

Observe that

$$|c_{\eta,a}(\omega)| \leq g_{\eta,a}(\omega) := \frac{\eta^2}{|\omega|^3} + \mathrm{e}^{-\eta a}\left(\frac{1}{|\omega|} + \frac{1}{\eta}\right) \tag{5.26}$$

and that $g_{\eta,a}(\omega)$ is a decreasing function of $|\omega|$; therefore we may give the upper bound

$$\begin{aligned}
&\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda \, \|(\mathrm{A}(\lambda) - \mathrm{A}_\eta(\lambda))|n(\lambda)\rangle\| \\
&\qquad = \int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda \sqrt{\sum_{m\neq n} |c_{\eta,a}(\omega_{mn}(\lambda))|^2 |\langle n(\lambda)|\partial_\lambda \mathrm{H}(\lambda)|m(\lambda)\rangle|^2} \\
&\qquad \leq \int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda \, \max_m g_{\eta,a}(\omega_{mn}(\lambda)) \sqrt{\sum_{m\neq n} |\langle n(\lambda)|\partial_\lambda \mathrm{H}(\lambda)|m(\lambda)\rangle|^2} \\
&\qquad \leq \int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda \, \max_m g_{\eta,a}(\omega_{mn}(\lambda)) \|\partial_\lambda \mathrm{H}(\lambda)\| \\
&\qquad \leq \|\partial_\lambda \mathrm{H}\|_{n,1} g_{\eta,a}(\min_m \min_\lambda \omega_{mn}(\lambda)) \\
&\qquad = \|\partial_\lambda \mathrm{H}\|_{n,1} g_{\eta,a}(\Delta_n).
\end{aligned} \tag{5.27}$$

The requirement that this error be at most ϵ is then satisfied by setting

$$\eta = \Delta_n^{3/2}(\epsilon/2)^{1/2}\|\partial_\lambda \mathrm{H}\|_{n,1}^{-1/2} \quad \Rightarrow \quad \frac{\eta^2}{\Delta_n^3}\|\partial_\lambda \mathrm{H}\|_{n,1} \leq \frac{\epsilon}{2} \tag{5.28}$$

and

$$a = \frac{1}{\eta}\log\left(\frac{\Delta_n + \eta}{\Delta_n \eta \epsilon/2}\|\partial_\lambda \mathrm{H}\|_{n,1}\right) \quad \Rightarrow \quad \mathrm{e}^{-\eta a}\left(\frac{1}{\Delta_n} + \frac{1}{\eta}\right)\|\partial_\lambda \mathrm{H}\|_{n,1} \leq \frac{\epsilon}{2} \tag{5.29}$$

which proves the lemma. □

5.4 Gate-based counterdiabatic driving

With the regularised truncated AGP in hand, we can now discuss approaches to gate-based counterdiabatic driving. For this purpose, we will understand CD as the task of implementing the path-ordered exponential $\mathrm{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) = \mathcal{T}\exp[-\mathrm{i}\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda \mathrm{A}_{\eta,a}(\lambda)]$ with some unitary $\tilde{\mathrm{U}}(\lambda_\mathrm{f}, \lambda_\mathrm{i})$, up to a given precision ϵ; by a simple triangle equality, the total error $\|(\mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \tilde{\mathrm{U}}(\lambda_\mathrm{f}, \lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\|$ is then at most $O(\epsilon)$. In this section, we present a deterministic gate-based algorithm

that achieves exactly this goal, and provide an upper bound on the complexity of running this algorithm on a fault-tolerant gate-based quantum computer.

Evidently, time-dependent simulation of the ordered exponential $\mathrm{U}_{\eta,a}$ is an integral component of gate-based CD, and we will employ LTS formulae for this purpose, as described in section 5.2.2. However, since the LTS formalism assumes that the integrand in the exponent is a sum of finitely many operator terms, and $\mathrm{A}_{\eta,a}$ is defined as an operator integral, we first need to approximate $\mathrm{A}_{\eta,a}$ with a suitable sum. To this end, we give a construction that is similar to an operator-valued Riemann approximation to the integral, except we make a more clever choice of evaluation points and weights way to obtain a smaller error bound. As such, all terms in the sum are proportional to the integrand in eq. 5.18, i.e. an operator of the form $\exp[-\mathrm{i}\mathrm{H}(\lambda)\tau]\partial_\lambda \mathrm{H}(\lambda)\exp[\mathrm{i}\mathrm{H}(\lambda)\tau]$.

This leads to another issue: the mentioned time-dependent hamiltonian simulation routines assume that the operator terms in the sum are given, in the same sense that we assume $\mathrm{H}(\lambda)$ and $\partial_\lambda \mathrm{H}(\lambda)$ given as sums of simple (weighted unitary) terms. Clearly, this is not the case here, since one would first have to simulate all operators $\exp[\pm\mathrm{i}\mathrm{H}(\lambda)\tau]$ to construct the desired operator terms. To avoid nested simulation, we should integrate the simulation of the $\exp[\pm\mathrm{i}\mathrm{H}(\lambda)\tau]$ operators in the procedure for approximating $\mathrm{U}_{\eta,a}$. Fortunately, this can be done in a straightforward way using Trotter-based formulae, since when we exponentiate the integrand terms, the operators $\exp[\pm\mathrm{i}\mathrm{H}(\lambda)\tau]$ may be taken out of the exponent:

$$\begin{aligned} &\exp[-\mathrm{i}(\mathrm{e}^{-\mathrm{i}\mathrm{H}(\lambda)\tau}\partial_\lambda \mathrm{H}(\lambda)\mathrm{e}^{\mathrm{i}\mathrm{H}(\lambda)\tau})\delta\lambda] \\ =\ &\exp[-\mathrm{i}\mathrm{H}(\lambda)\tau]\exp[-\mathrm{i}\partial_\lambda \mathrm{H}(\lambda)\delta\lambda]\exp[\mathrm{i}\mathrm{H}(\lambda)\tau]. \end{aligned} \tag{5.30}$$

This follows immediately from the fact that $\exp[\mathrm{U}\mathrm{O}\mathrm{U}^\dagger] = \mathrm{U}\exp[\mathrm{O}]\mathrm{U}^\dagger$, for any O and unitary U, which may be shown using a Taylor expansion. As such, we avoid double exponentiation and are left with only a product of ordinarily simulatable operator exponentials.

In what follows, we first give a detailed description of our summation method, which gives rise to an AGP approximation $\mathrm{A}_{\eta,a}^{M,q} \approx \mathrm{A}_{\eta,a}$ and a corresponding evolution operator $\mathrm{U}_{\eta,a}^{M,q}(\lambda_\mathrm{f},\lambda_\mathrm{i}) = \mathcal{T}\exp[-\mathrm{i}\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}}\mathrm{d}\lambda\,\mathrm{A}_{\eta,a}^{M,q}(\lambda)]$. Subsequently, we give a LTS construction to simulate $\mathrm{U}_{\eta,a}^{M,q}$, producing the unitary $\tilde{\mathrm{U}}$. We establish conditions on the relevant parameters to upper bound the errors $\|(\mathrm{U}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}) - \mathrm{U}_{\eta,a}^{M,q}(\lambda_\mathrm{f},\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\|$ and $\|(\mathrm{U}_{\eta,a}^{M,q}(\lambda_\mathrm{f},\lambda_\mathrm{i}) - \tilde{\mathrm{U}}(\lambda_\mathrm{f},\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\|$; we then have that the error $\|(\mathrm{U}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}) - \tilde{\mathrm{U}}(\lambda_\mathrm{f},\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\|$ is upper bounded by the sum of these two errors, by a triangle inequality. Finally, setting this error to at most $O(\epsilon)$ leads to a gate complexity expression for the algorithm.

5.4.1 Weighted sum approximation

The most elementary way of approximating an integral of a scalar-valued function is a Riemann sum: the integration range is partitioned into M subintervals and one approximates $\int_a^b \mathrm{d}x\, f(x) \approx \sum_{\kappa=1}^{M} f(x_\kappa)\delta x_\kappa$. A similar thing can be done with operator functions. Let us write

$$\mathrm{A}_{\eta,a} = \frac{1}{2}\int_0^a \mathrm{d}\tau\, \mathrm{e}^{-\eta\tau}\big(\partial_\lambda \mathrm{H}(\tau) - \partial_\lambda \mathrm{H}(-\tau)\big) \tag{5.31}$$

where we dropped the λ argument for legibility, and denote $\mathrm{O}(\tau) = \mathrm{e}^{-\mathrm{i}\mathrm{H}\tau}\mathrm{O}\mathrm{e}^{\mathrm{i}\mathrm{H}\tau}$ for any operator O. We partition the interval $[0, a]$ into M subintervals, picking $\tau_0 < \tau_1 < \cdots < \tau_M$ such that $\tau_0 = 0$ and $\tau_M = a$; define $\delta\tau_\kappa = \tau_\kappa - \tau_{\kappa-1}$. Note that the subintervals need not be uniform. The operator Riemann sum then takes the form

$$\mathrm{A}_{\eta,a} \approx \frac{1}{2}\sum_{\kappa=1}^{M} \delta\tau_\kappa \mathrm{e}^{-\eta\tau_\kappa}\big(\partial_\lambda \mathrm{H}(\tau_\kappa) - \partial_\lambda \mathrm{H}(-\tau_\kappa)\big). \tag{5.32}$$

However, this is a rather crude method in that the error in each subinterval is relatively large, so that many subintervals are needed to make the overall error small. We can address this in a more clever way by using quadrature methods based on Lagrange interpolation.

Lagrange interpolation and scalar quadrature

Scalar functions may be approximated as weighted sums through Lagrange interpolation. The idea of this interpolation method is to approximate a function $f(x)$ in some interval $[a, b]$ by some polynomial $p_q(\tau)$ of degree at most q such that the interpolation condition

$$f(x_\alpha) = p_q(x_\alpha), \quad x_0, \cdots, x_q \in [a, b] \tag{5.33}$$

is satisfied for a set of $q+1$ points in $[a, b]$. It can be shown that such a polynomial is unique, and is given by

$$p_q(x) = \sum_{\alpha=0}^{q} f(x_\alpha) l_\alpha(x) \tag{5.34}$$

where

$$l_\alpha(x) = \prod_{\substack{0\le\beta\le q \\ \beta\neq\alpha}} \frac{x - x_\beta}{x_\alpha - x_\beta} \tag{5.35}$$

are the Lagrange basis polynomials. It is straightforward to check that $l_\alpha(x_\beta) = \delta_{\alpha\beta}$, which also verifies the interpolation condition (eq. 5.33). Clearly, if f itself

is a polynomial of degree at most q, then $p_q = f$ since p_q is unique, and the interpolation is exact. For other functions, it may be shown [SBG+13] that the remainder $r_q(x) = f(x) - p_q(x)$ at any point in $[a, b]$ is given by

$$r_q(x) = \tilde{l}_q(x) \frac{f^{(q+1)}(\xi)}{(q+1)!} \tag{5.36}$$

for some $\xi \in [a, b]$, where $\tilde{l}_q(x) = \prod_{\alpha=0}^{q}(x - x_\alpha)$.

The integral $\int_a^b f(x)\mathrm{d}x$ is then approximated as

$$\int_a^b f(x)\mathrm{d}x \approx \int_a^b p_q(x)\mathrm{d}x = \sum_{a=0}^{q} f(x_\alpha) \underbrace{\int_a^b l_\alpha(x)\mathrm{d}x}_{=w_\alpha}. \tag{5.37}$$

Note that $\sum_{\alpha=0}^{q} w_\alpha = b-a$; this can be seen by taking $f(x) = 1$, in which case the approximation is exact (because f is a polynomial of degree zero). For general functions, the error is

$$R = \int_a^b r_q(x)\mathrm{d}x = \frac{f^{(q+1)}(\xi)}{(q+1)!} \int_a^b \tilde{l}_q(x)\mathrm{d}x. \tag{5.38}$$

Clearly, $|R|$ is bounded above by

$$|R| \leq \frac{(b-a)^{q+2} \max_{\xi\in[a,b]} |f^{(q+1)}(\xi)|}{(q+1)!}. \tag{5.39}$$

Applying quadrature to the AGP

We can apply these ideas by evaluating the integrand of eq. 5.31 at multiple points in each of the M subintervals, and weighing each term accordingly:

$$\mathrm{A}_{\eta,a} \approx \mathrm{A}_{\eta,a}^{M,q} = \frac{1}{2} \sum_{\kappa=1}^{M} \delta\tau_\kappa \sum_{\alpha=0}^{q} w_{\kappa,\alpha} \mathrm{e}^{-\eta\tau_{\kappa,\alpha}} \big(\partial_\lambda \mathrm{H}(\tau_{\kappa,\alpha}) - \partial_\lambda \mathrm{H}(-\tau_{\kappa,\alpha})\big). \tag{5.40}$$

Here, $q \leq 0$ is introduced as a free parameter, so that $q+1$ evaluations are made within each subinterval. The weights $w_{\kappa,\alpha}$ are determined through Lagrange interpolation and are such that $\sum_{\alpha=0}^{q} w_{\kappa,\alpha} = 1$ for all κ. Furthermore, they depend only on our choice of the interpolation points $\tau_{\kappa,\alpha}$ and not on the value of the integrand at those points. We will not specify the choice of the interpolation points, and instead rely on general Lagrange interpolation bounds for estimating the error made in the approximation. This gives rise to the following lemma.

5.4.1. Lemma. *Let* $\mathrm{U}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i})$ *be given as in lemma 5.3.1, let* $\mathrm{A}^{M,q}_{\eta,a}(\lambda)$ *be given as in eq. 5.40 and let* $\mathrm{U}^{M,q}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}) = \mathcal{T}\exp[-\mathrm{i}\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}}\mathrm{d}\lambda\,\mathrm{A}^{M,q}_{\eta,a}(\lambda)]$. *Assume* $\eta \leq \min_{\lambda\in[\lambda_\mathrm{i},\lambda_\mathrm{f}]}\|\mathrm{H}(\lambda)\|$. *If we choose*

$$\tau_\kappa = -\frac{q+2}{\eta}\log\Big(1-\frac{\kappa}{M}(1-\mathrm{e}^{-\eta a/(q+2)})\Big), \qquad \kappa\in\{0,\dots,M\} \tag{5.41}$$

and if we set, for any choice of interpolation points $\tau_{\kappa,0}<\tau_{\kappa,1}<\cdots<\tau_{\kappa,q}$ *within each subinterval* $[\tau_{\kappa-1},\tau_\kappa]$,

$$w_{\kappa,\alpha} = \frac{1}{\delta\tau_\kappa}\int_{\tau_{\kappa-1}}^{\tau_\kappa}\mathrm{d}\tau\prod_{\substack{0\leq\beta\leq q\\ \beta\neq\alpha}}\frac{\tau-\tau_{\kappa,\beta}}{\tau_{\kappa,\alpha}-\tau_{\kappa,\beta}}, \qquad \alpha\in\{0,\dots,q\} \tag{5.42}$$

then

$$\|(\mathrm{U}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i})-\mathrm{U}^{M,q}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \leq \epsilon \tag{5.43}$$

provided that

$$M \geq \max\left\{\frac{3\mathrm{e}(2a)^{1+1/(q+1)}}{\epsilon^{1/(q+1)}(q+1)}\|\mathrm{H}\|_{\infty,\infty}\|\partial_\lambda\mathrm{H}\|_{n,1}^{1/(q+1)}, \mathrm{e}^{\eta a/(q+2)}-1\right\}. \tag{5.44}$$

Proof sketch.
As before, we bound $\|(\mathrm{U}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i})-\mathrm{U}^{M,q}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \leq \int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}}\mathrm{d}\lambda\|\mathrm{A}_{\eta,a}-\mathrm{A}^{M,q}_{\eta,a}\|_n$, and then apply another triangle inequality to bound this quantity by a sum of contributions from each of the M subintervals:

$$\|\mathrm{A}_{\eta,a}-\mathrm{A}^{M,q}_{\eta,a}\|_n \leq \frac{1}{2}\sum_{\kappa=1}^{M}\Bigg(\Bigg\|\int_{\tau_{\kappa-1}}^{\tau_\kappa}\mathrm{d}\tau\,\mathrm{e}^{-\eta\tau}\partial_\lambda\mathrm{H}(\tau)-\sum_{\alpha=0}^{q}w_{\kappa,\alpha}\mathrm{e}^{-\eta\tau_{\kappa,\alpha}}\partial_\lambda\mathrm{H}(\tau_{\kappa,\alpha})\Bigg\|_n + (\tau\leftrightarrow-\tau)\Bigg). \tag{5.45}$$

Given that the weights $w_{\kappa,\alpha}$ in eq. 5.42 are chosen according to the Lagrange polynomial basis (cf. eqs. 5.35 and 5.37), it turns out that the general quadrature error bound in eq. 5.39 may be used in an element-wise fashion to bound the expression in eq. 5.45. We obtain

$$\|\mathrm{A}_{\eta,a}-\mathrm{A}^{M,q}_{\eta,a}\|_n \leq Q_n^q\sum_{\kappa=1}^{M}\delta\tau_\kappa^{q+2}\mathrm{e}^{-\eta\tau_{\kappa-1}}; \qquad Q_n^q = \frac{(3\|\mathrm{H}\|)^{q+1}\|\partial_\lambda\mathrm{H}\|_n}{(q+1)!}. \tag{5.46}$$

The next step is to choose the intervals $\delta\tau_\kappa$. The motivation behind the choice in eq. 5.41 is to make the intervals small when $\mathrm{e}^{-\eta\tau_\kappa}$ is large, and large when $\mathrm{e}^{-\eta\tau_\kappa}$ is small. In this way, every term in the sum in eq. 5.46 is almost constant, and can be bounded precisely by a constant at most $Q_n^q(2a/M)^{q+2}$ assuming

$M \geq \mathrm{e}^{\eta a/(q+2)} - 1$. This gives a better error bound than, for example, intervals of constant size.

In the end, with a bit of rewriting we find

$$\|(\mathrm{U}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}) - \mathrm{U}^{M,q}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \leq \frac{3^{q+1}(2a)^{q+2}}{M^{q+1}(q+1)!}\|\mathrm{H}\|^{q+1}_{\infty,\infty}\|\partial_\lambda \mathrm{H}(\lambda)\|_{n,1} \tag{5.47}$$

and this error can be made at most ϵ by taking M as in eq. 5.44. □

Full proof.
As before, we start from the observation that

$$\|(\mathrm{U}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}) - \mathrm{U}^{M,q}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \leq \int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\, \|\mathrm{A}_{\eta,a}(\lambda) - \mathrm{A}^{M,q}_{\eta,a}(\lambda)\|_{n(\lambda)}. \tag{5.48}$$

By a triangle inequality, we have

$$\begin{aligned}
&\|\mathrm{A}_{\eta,a} - \mathrm{A}^{M,q}_{\eta,a}\|_n \\
&\quad\leq \frac{1}{2}\sum_{\kappa=1}^{M}\Bigg(\left\|\int_{\tau_{\kappa-1}}^{\tau_\kappa}\mathrm{d}\tau\,\mathrm{e}^{-\eta\tau}\,\partial_\lambda \mathrm{H}(\tau) - \sum_{\alpha=0}^{q} w_{\kappa,\alpha}\,\mathrm{e}^{-\eta\tau_{\kappa,\alpha}}\partial_\lambda \mathrm{H}(\tau_{\kappa,\alpha})\right\|_n + \\
&\qquad\qquad \left\|\int_{\tau_{\kappa-1}}^{\tau_\kappa}\mathrm{d}\tau\,\mathrm{e}^{-\eta\tau}\,\partial_\lambda \mathrm{H}(-\tau) - \sum_{\alpha=0}^{q} w_{\kappa,\alpha}\,\mathrm{e}^{-\eta\tau_{\kappa,\alpha}}\partial_\lambda \mathrm{H}(-\tau_{\kappa,\alpha})\right\|_n\Bigg).
\end{aligned} \tag{5.49}$$

The first term on the right-hand side of eq. 5.49 can be upper bounded by expanding in the eigenbasis of H and applying element-wise interpolation. For each vector element $\mathrm{e}^{-\eta\tau}\langle m|\partial_\lambda \mathrm{H}(\tau)|n\rangle$, which is a scalar function of τ, we can then use the error bound in eq. 5.39. This works for our vector-valued case because $\mathrm{argmax}_{\sigma_m\in[\tau_{\kappa-1},\tau_\kappa]}\left|\langle m|\frac{\mathrm{d}^{q+1}}{\mathrm{d}\tau^{q+1}}\partial_\lambda(\mathrm{e}^{-\eta\tau}\mathrm{H}(\tau))|_{\tau=\sigma_m}|n\rangle\right|$ is identical for every m. We obtain

$$\begin{aligned}
&\left\|\int_{\tau_{\kappa-1}}^{\tau_\kappa}\mathrm{d}\tau\,\mathrm{e}^{-\eta\tau}\,\partial_\lambda \mathrm{H}(\tau) - \sum_{\alpha=0}^{q} w_{\kappa,\alpha}\,\mathrm{e}^{-\eta\tau_{\kappa,\alpha}}\partial_\lambda \mathrm{H}(\tau_{\kappa,\alpha})\right\|_n \\
&= \left(\sum_m\left|\int_{\tau_{\kappa-1}}^{\tau_\kappa}\mathrm{d}\tau\,\mathrm{e}^{-\eta\tau}\,\langle m|\partial_\lambda \mathrm{H}(\tau)|n\rangle - \sum_{\alpha=0}^{q} w_{\kappa,\alpha}\,\mathrm{e}^{-\eta\tau_{\kappa,\alpha}}\langle m|\partial_\lambda \mathrm{H}(\tau_{\kappa,\alpha})|n\rangle\right|^2\right)^{1/2} \\
&\leq \frac{\delta\tau_\kappa^{q+2}}{(q+1)!}\left(\sum_m \max_{\sigma_m\in[\tau_{\kappa-1},\tau_\kappa]}\left|\langle m|\frac{\mathrm{d}^{q+1}}{\mathrm{d}\tau^{q+1}}(\mathrm{e}^{-\eta\tau}\partial_\lambda \mathrm{H}(\tau))\Big|_{\tau=\sigma_m}|n\rangle\right|^2\right)^{1/2} \\
&= \frac{\delta\tau_\kappa^{q+2}}{(q+1)!}\mathrm{e}^{-\eta\tau_{\kappa-1}}\left\|\sum_{s=0}^{q+1}\binom{q+1}{s}(-\eta)^s[\underbrace{-\mathrm{iH},[-\mathrm{iH},\cdots,[-\mathrm{iH}}_{q+1-s},\partial_\lambda \mathrm{H}]\cdots]]\right\|_n
\end{aligned}$$

$$\leq \frac{\delta\tau_\kappa^{q+2}}{(q+1)!}\mathrm{e}^{-\eta\tau_{\kappa-1}}\sum_{s=0}^{q+1}\binom{q+1}{s}\eta^s(2\|\mathrm{H}\|)^{q+1-s}\|\partial_\lambda \mathrm{H}\|_n$$
$$= \frac{\delta\tau_\kappa^{q+2}}{(q+1)!}\mathrm{e}^{-\eta\tau_{\kappa-1}}(\eta+2\|\mathrm{H}\|)^{q+1}\|\partial_\lambda \mathrm{H}\|_n. \tag{5.50}$$

The calculation for the $\tau \leftrightarrow -\tau$ term in eq. 5.49 is identical; therefore

$$\|\mathrm{A}_{\eta,a}(\lambda) - \mathrm{A}_{\eta,a}^{M,q}(\lambda)\|_{n(\lambda)} \leq Q_n^q(\lambda)\sum_{\kappa=1}^{M}\delta\tau_\kappa^{q+2}\mathrm{e}^{-\eta\tau_{\kappa-1}} \tag{5.51}$$

where we have defined

$$Q_n^q = \frac{(\eta+2\|\mathrm{H}(\lambda)\|)^{q+1}\|\partial_\lambda \mathrm{H}(\lambda)\|_{n(\lambda)}}{(q+1)!}. \tag{5.52}$$

We will now choose the τ_κ, so as to ultimately derive the required value of M. Instead of using constant intervals $\delta\tau_\kappa$, it should be intuitively beneficial to pick $\delta\tau_\kappa$ small for those κ where $\mathrm{e}^{-\eta\tau}$ is large, and pick $\delta\tau_\kappa$ large when $\mathrm{e}^{-\eta\tau}$ is small. An ansatz that makes every term in the sum $\sum_{\kappa=1}^{M}\delta\tau_\kappa^{q+2}\mathrm{e}^{-\eta\tau_{\kappa-1}}$ approximately constant can be found by regarding $\tau(\kappa)$ as a continuous function of κ and solving the differential equation[3]

$$\frac{d\tau}{d\kappa}\mathrm{e}^{-\eta\tau/(q+2)} = \zeta \tag{5.53}$$

for some positive constant ζ, subject to the boundary condition $\tau(0) = 0$. This differential equation is solved by $\tau(\kappa) = -\frac{q+2}{\eta}\log(1-\frac{\eta\kappa\zeta}{q+2})$; the condition $\tau(M) = a$ then implies

$$\zeta = \frac{q+2}{\eta M}(1-\mathrm{e}^{-\eta a/(q+2)}) \leq \frac{a}{M} \tag{5.54}$$

so we obtain the solution

$$\tau(\kappa) = -\frac{q+2}{\eta}\log\Big(1-\frac{\kappa}{M}(1-\mathrm{e}^{-\eta a/(q+2)})\Big). \tag{5.55}$$

Each term in the sum then contributes approximately ζ^{q+2}. To make this exact, we Taylor expand $\tau(\kappa)$ about $\kappa-1$ to obtain[4]

$$\delta\tau_\kappa\mathrm{e}^{-\eta\tau_{\kappa-1}/(q+2)} - \underbrace{\mathrm{e}^{-\eta\tau_{\kappa-1}/(q+2)}\frac{\mathrm{d}\tau}{\mathrm{d}\kappa}(\kappa-1)}_{=\zeta}$$
$$= \mathrm{e}^{-\eta\tau_{\kappa-1}/(q+2)}\frac{1}{2}\frac{\mathrm{d}^2\tau}{\mathrm{d}\kappa^2}(\hat{\kappa}) \qquad (\text{for some } \hat{\kappa}\in[\kappa-1,\kappa])$$

[3] I thank Kareljan Schoutens for suggesting this ansatz.

[4] To be precise, we Taylor expand $\tau(x)$ about $x = \kappa-1$ with quadratic Lagrange remainder, and this expansion is then evaluated at κ.

$$
\begin{aligned}
&= \frac{q+2}{2\eta}\mathrm{e}^{-\eta\tau_{\kappa-1}/(q+2)}\left(\frac{1-\mathrm{e}^{-\eta a/(q+2)}}{M-(1-\mathrm{e}^{-\eta a/(q+2)})\hat{\kappa}}\right)^2 \\
&\leq \frac{q+2}{2\eta}\Big(1-\frac{\kappa-1}{M}(1-\mathrm{e}^{-\eta a/(q+2)})\Big)\left(\frac{(1-\mathrm{e}^{-\eta a/(q+2)})}{M-(1-\mathrm{e}^{-\eta a/(q+2)})\kappa}\right)^2 \\
&= \frac{q+2}{2M\eta}\left(\frac{[1-\mathrm{e}^{-\eta a/(q+2)}]^2}{[M-(1-\mathrm{e}^{-\eta a/(q+2)})\kappa]}+\frac{[1-\mathrm{e}^{-\eta a/(q+2)}]^3}{[M-(1-\mathrm{e}^{-\eta a/(q+2)})\kappa]^2}\right) \\
&\leq \frac{(q+2)[1-\mathrm{e}^{-\eta a/(q+2)}]}{2M\eta}\left(\frac{\mathrm{e}^{\eta a/(q+2)}-1}{M}+\Big(\frac{\mathrm{e}^{\eta a/(q+2)}-1}{M}\Big)^2\right) \\
&\leq \frac{(q+2)[1-\mathrm{e}^{-\eta a/(q+2)}]}{M\eta} = \zeta \qquad \text{if } M \geq \mathrm{e}^{\eta a/(q+2)}-1. \qquad (5.56)
\end{aligned}
$$

In the end, we find that the interpolation error obeys

$$
\|\mathrm{A}_{\eta,a}-\mathrm{A}_{\eta,a}^{M,q}\|_n \leq M\Big(\frac{2a}{M}\Big)^{q+2}Q_n^q = \frac{(2a)^{q+2}}{M^{q+1}}\frac{(\eta+2\|\mathrm{H}\|)^{q+1}\|\partial_\lambda \mathrm{H}\|_n}{(q+1)!} \qquad (5.57)
$$

implying that

$$
\begin{aligned}
&\|(\mathrm{U}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i})-\mathrm{U}_{\eta,a}^{M,q}(\lambda_\mathrm{f},\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \\
&\qquad \leq \frac{(2a)^{q+2}}{M^{q+1}(q+1)!}\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\,(\eta+2\|\mathrm{H}(\lambda)\|_\infty)^{q+1}\|\partial_\lambda \mathrm{H}(\lambda)\|_{n(\lambda)}. \qquad (5.58)
\end{aligned}
$$

We simplify the expression above by using the assumption that $\eta \leq \min_\lambda \|\mathrm{H}(\lambda)\|$, and by using Hölder's inequality[5]:

$$
\begin{aligned}
&\|(\mathrm{U}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i})-\mathrm{U}_{\eta,a}^{M,q}(\lambda_\mathrm{f},\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \\
&\qquad \leq \frac{3^{q+1}(2a)^{q+2}}{M^{q+1}(q+1)!}\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\,\|\mathrm{H}(\lambda)\|_\infty^{q+1}\|\partial_\lambda \mathrm{H}(\lambda)\|_{n(\lambda)} \\
&\qquad \leq \frac{3^{q+1}(2a)^{q+2}}{M^{q+1}(q+1)!}\|\mathrm{H}\|_{\infty,\infty}^{q+1}\|\partial_\lambda \mathrm{H}(\lambda)\|_{n,1}. \qquad (5.59)
\end{aligned}
$$

To make this error at most ϵ, it is sufficient to take

$$
M \geq \max\left\{\frac{3\mathrm{e}(2a)^{1+1/(q+1)}}{\epsilon^{1/(q+1)}(q+1)}\|\mathrm{H}\|_{\infty,\infty}\|\partial_\lambda \mathrm{H}\|_{n,1}^{1/(q+1)}, \mathrm{e}^{\eta a/(q+2)}-1\right\} \qquad (5.60)
$$

where we used Stirling's approximation to rewrite $((q+1)!)^{-(q+1)} \leq \mathrm{e}^{1-1/(q+1)}(q+1)^{-1} \leq \mathrm{e}(q+1)^{-1}$. □

[5]Define the integral norm $\|f\|_p = \big(\int_a^b \mathrm{d}x\,|f(x)|^p\big)^{1/p}$ with $\|f\|_\infty = \sup_{x\in[a,b]}|f(x)|$ and let f, g be such that $\|f\|_p$, $\|g\|_p$ and $\|fg\|_p$ are bounded for $1 \leq p \leq \infty$; then Hölder's inequality states that $\|fg\|_1 \leq \|f\|_p\|g\|_{p'}$ if $1/p+1/p'=1$. We choose $p=\infty$ and $p'=1$.

5.4.2 Lie-Trotter-Suzuki expansion of the weighted sum

The path-ordered exponential $\mathrm{U}_{\eta,a}^{M,q}(\lambda_\mathrm{f},\lambda_\mathrm{i}) = \mathcal{P}\exp[-\mathrm{i}\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}}\mathrm{d}\lambda\,\mathrm{A}_{\eta,a}^{M,q}(\lambda)]$ may be implemented by a k^{th}-order r-segment LTS product formula. These formulas follow the same description as given in section 5.2.2, except we will be working in λ space instead of time space.

The first-order single-segment formula $\tilde{\mathrm{U}}_{1,1}(\lambda_\mathrm{f},\lambda_\mathrm{i})$, for the operator

$$\mathrm{A}_{\eta,a}^{M,q}(\lambda) = \sum_{\kappa=-M}^{M}\sum_{\alpha=0}^{q}\mathrm{C}_{\kappa,\alpha}(\lambda) \tag{5.61}$$

where

$$\begin{aligned}\mathrm{C}_{\kappa,\alpha}(\lambda) &= \mathrm{e}^{-\mathrm{i}\mathrm{H}(\lambda)\tau_{\kappa,\alpha}}\,\mathrm{B}_{\kappa,\alpha}(\lambda)\,\mathrm{e}^{\mathrm{i}\mathrm{H}(\lambda)\tau_{\kappa,\alpha}};\\ \mathrm{B}_{\kappa,\alpha}(\lambda) &= \frac{1}{2}\delta\tau_\kappa w_{\kappa,\alpha}\,\mathrm{e}^{-\eta|\tau_{\kappa,\alpha}|}\operatorname{sgn}\tau_{\kappa,\alpha}\partial_\lambda\mathrm{H}(\lambda)\end{aligned} \tag{5.62}$$

(with $\mathrm{C}_{0,\alpha}=0$), is given by

$$\begin{aligned}\tilde{\mathrm{U}}_{1,1}(\lambda_\mathrm{f},\lambda_\mathrm{i}) = &\left(\prod_{\kappa=-M}^{M}\prod_{\alpha=0}^{q}\exp[-\mathrm{i}\mathrm{C}_{\kappa,\alpha}(\lambda_\mathrm{i}+\delta\lambda/2)\delta\lambda/2]\right)\\ &\times\left(\prod_{\kappa=M}^{-M}\prod_{\alpha=q}^{0}\exp[-\mathrm{i}\mathrm{C}_{\kappa,\alpha}(\lambda_\mathrm{i}+\delta\lambda/2)\delta\lambda/2]\right)\end{aligned} \tag{5.63}$$

where $\delta\lambda = \lambda_\mathrm{f}-\lambda_\mathrm{i}$. Here, the operator exponentials in the product take the form

$$\begin{aligned}\exp[-\mathrm{i}\mathrm{C}_{\kappa,\alpha}(\lambda)\delta\lambda] = &\quad \exp[-\mathrm{i}\mathrm{H}(\lambda)\tau_{\kappa,\alpha}]\\ &\times\exp\left[-\frac{\mathrm{i}}{2}\mathrm{e}^{-\eta|\tau_{\kappa,\alpha}|}\operatorname{sgn}\tau_{\kappa,\alpha}w_{\kappa,\alpha}\partial_\lambda\mathrm{H}(\lambda)\,\delta\tau_\kappa\,\delta\lambda\right]\\ &\times\exp[\mathrm{i}\mathrm{H}(\lambda)\tau_{\kappa,\alpha}].\end{aligned} \tag{5.64}$$

Higher-order, multi-segment formulas are constructed as in eqs. 5.13 and 5.15. In the same way, we can use lemma 5.2.1 to put a bound on the number of path-independent operator exponentials required to achieve a certain precision. We apply this result to our case by giving an explicit expression for Λ corresponding to the ordered exponential $\mathrm{U}_{\eta,a}^{M,q}$. If the spectral distance between $\mathrm{U}_{\eta,a}^{M,q}$ and its LTS decomposition $\tilde{\mathrm{U}}_{k,r}$ is at most ϵ, then the euclidean state distance $\|(\mathrm{U}_{\eta,a}^{M,q}(\lambda_\mathrm{f},\lambda_\mathrm{i})-\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f},\lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\|$ is also at most ϵ.

To use lemma 5.2.1, we require bounds on the operator norms of the higher-order derivatives of the $\mathrm{C}_{\kappa,\alpha}$ operators. For the case of interpolating hamiltonians $\mathrm{H}(\lambda) = \mathrm{H}_\mathrm{i} + f(\lambda)\mathrm{H}_\mathrm{p}$, these norms take on a simple form, since

$$\left\|\frac{\partial^p}{\partial\lambda^p}\left(\mathrm{e}^{-\mathrm{i}\mathrm{H}(\lambda)\tau}\partial_\lambda\mathrm{H}\mathrm{e}^{-\mathrm{i}\mathrm{H}(\lambda)\tau}\right)\right\| = \left\|\frac{\partial^{p+1}f(\lambda)}{\partial\lambda^{p+1}}\mathrm{H}_\mathrm{p}\right\| = \|\partial_\lambda^{p+1}\mathrm{H}(\lambda)\|; \tag{5.65}$$

however, for general hamiltonians, the higher-order derivatives of $\mathrm{C}_{\kappa,\alpha}(\lambda)$ are less straightforward and their norms can grow rapidly with p, as a result of compounding product rules. Since there can be a significant discrepancy in the the scaling of the derivatives between interpolating case $\mathrm{H}(\lambda) = \mathrm{H_i} + f(\lambda)\mathrm{H_p}$ and the general case, and better bounds on the norm of $\|(\partial^p/\partial\lambda^p)\mathrm{C}_{\kappa,\alpha}(\lambda)\|$ require more specifics about the definition of $\mathrm{H}(\lambda)$, we will stick to the interpolating case for the rest of the section. A condition on the number r, and thereby the number of exponentials required in our gate-based counterdiabatic driving algorithm, is then given by the following lemma.

5.4.2. Lemma. *Let* $\mathrm{H}(\lambda) = \mathrm{H_i} + f(\lambda)\mathrm{H_p}$ *for hermitian* $\mathrm{H_i}$ *and* $\mathrm{H_p}$ *and a scalar function* $f(\lambda)$ *defined on the interval* $[\lambda_\mathrm{i}, \lambda_\mathrm{f}]$ *that is* $2k+1$ *times differentiable. Let* $\mathrm{U}^{M,q}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i})$, $\{\tau_\kappa\}$, $\{\tau_{\kappa,\alpha}\}$ *and* $\{w_{\kappa,\alpha}\}$ *be as in lemma 5.4.1, and let* $\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f}, \lambda_\mathrm{i})$ *be the* k^{th}*-order* r*-segment LTS decomposition of* $\mathrm{U}^{M,q}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i})$ *as in eq. 5.15. Let the conditions of lemma 5.4.1 be satisfied. Suppose*

$$\max_{1\le p\le 2k+1}\left[\left(\frac{2(1-\mathrm{e}^{-\eta a})}{\eta}\|\partial_\lambda^p \mathrm{H}\|_{\infty,\infty}\right)^{1/p}\right] \le \tilde{\Lambda}. \tag{5.66}$$

If $\|\partial_\lambda \mathrm{H}\|_{n,1} \ge 3^{-(q+1)}\frac{\eta}{1-\mathrm{e}^{-\eta a}}$ *and* $\epsilon \le \min\{(9/10)(5/3)^k\tilde{\Lambda}\delta\lambda, 1\}$, *then*

$$\|\mathrm{U}^{M,q}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f}, \lambda_\mathrm{i})\| \le \epsilon \tag{5.67}$$

provided that

$$r \ge 5k\tilde{\Lambda}\delta\lambda\left(\frac{5}{3}\right)^k\left(\frac{\tilde{\Lambda}\delta\lambda}{\epsilon}\right)^{1/2k}. \tag{5.68}$$

Proof.
We merely need to bound the sum $\sum_{\kappa=-M}^{M}\sum_{\alpha=0}^{q}\left\|\frac{\partial^p}{\partial\lambda^p}\mathrm{C}_{\kappa,\alpha}(\lambda)\right\|$ to determine the $\tilde{\Lambda}$ quantity from lemma 5.2.1. Clearly,

$$\sum_{\kappa=-M}^{M}\sum_{\alpha=0}^{q}\left\|\frac{\partial^p}{\partial\lambda^p}\mathrm{C}_{\kappa}(\lambda)\right\| = \frac{1}{2}\|\partial_\lambda^{p+1}\mathrm{H}(\lambda)\|\sum_{\kappa=-M}^{M}\sum_{\alpha=0}^{q}\delta\tau_\kappa w_{\kappa,\alpha}\,\mathrm{e}^{-\eta|\tau_{\kappa,\alpha}|}. \tag{5.69}$$

We see that the sum on the right-hand side of this equation approximates the integral $\int_{-a}^{a}\mathrm{d}\tau\,\mathrm{e}^{-\eta|\tau|}$. Considering only the positive half of the integration range, i.e. $\kappa > 0$, we have

$$\sum_{\kappa=1}^{M}\sum_{\alpha=0}^{q}\delta\tau_\kappa w_{\kappa,\alpha}\mathrm{e}^{-\eta\tau_{\kappa,\alpha}} = \sum_{\kappa=1}^{M}\left[\int_{\tau_{\kappa-1}}^{\tau_\kappa}\mathrm{d}\tau\,\mathrm{e}^{-\eta\tau} + R_\kappa\right] = \int_0^a \mathrm{d}\tau\,\mathrm{e}^{-\eta\tau} + R \tag{5.70}$$

where $\mathbb{R}_\kappa \in \mathbb{R}$, $R = \sum_{\kappa=1}^{M} R_\kappa$ and, much like in the proof of lemma 5.4.1,

$$|R_\kappa| \le \frac{\eta^{q+1}\delta\tau_\kappa^{q+2}}{(q+1)!}\mathrm{e}^{-\eta\tau_\kappa} \le \frac{\eta^{q+1}}{(q+1)!}\Big(\frac{2a}{M}\Big)^{q+2}. \tag{5.71}$$

but since the conditions of lemma 5.4.1 are assumed satisfied, we have

$$\frac{3^{q+1}(2a)^{q+2}}{M^{q+1}(q+1)!}\|\mathrm{H}\|_{\infty,\infty}^{q+1}\|\partial_\lambda \mathrm{H}\|_{n,1} \leq \epsilon \leq 1 \tag{5.72}$$

as well as $\eta \leq \min_\lambda \|\mathrm{H}(\lambda)\|$, so that

$$|R| \leq \frac{\eta^{q+1}(2a)^{q+2}}{M^{q+1}(q+1)!} \leq \frac{\eta^{q+1}\epsilon}{3^{q+1}\|\mathrm{H}\|_{\infty,\infty}^{q+1}\|\partial_\lambda \mathrm{H}\|_{n,1}} \leq \frac{1}{3^{q+1}\|\partial_\lambda \mathrm{H}\|_{n,1}}. \tag{5.73}$$

The premise that $\|\partial_\lambda \mathrm{H}\|_{n,1} \geq 3^{-(q+1)}\frac{\eta}{1-\mathrm{e}^{-\eta a}}$ then directly implies that $|R| \leq \frac{1-\mathrm{e}^{-\eta a}}{\eta}$. The calculation for $\kappa < 0$ is identical, and we finally have

$$\sum_{\kappa=-M}^{M}\sum_{\alpha=0}^{q}\left\|\frac{\partial^p}{\partial\lambda^p}\mathrm{C}_\kappa(\lambda)\right\| \leq \frac{2(1-\mathrm{e}^{-\eta a})}{\eta}\|\partial_\lambda^{p+1}\mathrm{H}(\lambda)\|. \tag{5.74}$$

Therefore the condition on $\tilde{\Lambda}$ in eq. 5.66 is sufficient to fulfill the requirements of lemma 5.2.1, and the result follows immediately. □

The assumption on $\|\partial_\lambda \mathrm{H}\|_{n,1}$ is justified since $3^{-(q+1)}$ can be made arbitrarily small and we work in the regime where η is small and a is large (note $\frac{\eta}{1-\mathrm{e}^{-\eta a}}$ approaches zero in the limit $a \to \infty$, $\eta \to 0$).

5.4.3 Gate complexity

Finally, we turn to the quantum gate complxity of the gate-based CD algorithm. In order to bound this complexity, we need to associate with each path-independent operator exponential in the product formula $\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f}, \lambda_\mathrm{i})$ a gate complexity and sum over all operator exponentials. In the case of AQC, all these operator exponentials were assumed to be simulatable with a single multi-qubit rotation gate, and unit cost was associated with each simulation. This is not the case for CD, since the constituents of the LTS formulae are exponentials of $\mathrm{H}(\lambda)$ and $\partial_\lambda \mathrm{H}(\lambda)$. Simulation costs for these exponentials in terms of quantum gates are provided through established time-independent hamiltonian simulation methods. We use a technique known as qubitisation [LC19b] for its optimality in the relevant parameters. In short, if an operator H is given as a linear combination of unitaries[6] $\mathrm{H} = \sum_{i=1}^{\ell}\beta_i \mathrm{V}_i$, qubitisation provides a routine that simulates $\mathrm{e}^{\pm \mathrm{i}\mathrm{H}\theta}$ to error ϵ in spectral norm using $O(\|\boldsymbol{\beta}\|_1\theta + \log 1/\epsilon)$ queries to an oracle that

[6]The qubitisation subroutine requires that the hamiltonian be supplied as a LCU or as a sparse matrix with oracle entry access. Since the former is more natural and more common form for most physics settings (in particular many-body problems), we go with the LCU formulation here.

provides access to the hamiltonian, where $\|\boldsymbol{\beta}\|_1 = \sum_j |\beta_j|$. Since this oracle can be implemented with $O(\ell)$ elementary gates [SBM06, CW12, LC19b], the simulation can be done with $O([\|\boldsymbol{\beta}\|_1\theta + \log 1/\epsilon]\ell)$ gates.[7]

With a clear definition of gate complexity, we can now state the main theorem of this chapter, establishing a gate complexity bound for our gate-based counterdiabatic driving algorithm.

5.4.3. Theorem (gate-based CD). *Suppose* $\mathrm{H}(\lambda) = \mathrm{H_i} + f(\lambda)\mathrm{H_p} = \sum_{i=1}^{\ell} \beta_i(\lambda)\mathrm{V}_i$ *with hermitian* $\mathrm{H_i}$ *and* $\mathrm{H_p}$ *and a scalar function* $f(\lambda)$ *that is* $2k+1$ *times differentiable on the interval* $[\lambda_\mathrm{i}, \lambda_\mathrm{f}]$, $\beta_i(\lambda) \in \mathbb{R}$ *and* V_i *unitary. Let* $\{|n(\lambda)\rangle\}$ *be the set of its instantaneous eigenstates. Suppose the spectrum of* $\mathrm{H}(\lambda)$ *around an eigenstate* $|n(\lambda)\rangle$ *has a gap of at least* Δ_n *on the entire interval. Let* $\mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) = \sum_n |n(\lambda_\mathrm{f})\rangle\langle n(\lambda_\mathrm{i})|$. *Define the quantities* $\gamma_p = (\sqrt{2}\Delta_n^{-3/2}\|\partial_\lambda \mathrm{H}\|_{n,1}^{1/2}\|\partial_\lambda^p \mathrm{H}\|_{\infty,\infty})^{1/p}$, $p^* = \arg\max_{1\le p\le 2k+1}(\epsilon^{-1/2p}\gamma_p)$ *and* $\tilde{\Lambda} = \epsilon^{-1/2p^*}\gamma_{p^*}$. *Furthermore, let* $\epsilon > 0$, $q \in \mathbb{N}_{>0}$ *and suppose that*

(a) $\epsilon \le \min\{[(9/10)(5/3)^k \gamma_{p^*}(\lambda_\mathrm{f} - \lambda_\mathrm{i})]^{1-\frac{1}{2p^*+1}}, 1\}$;

(b) $\|\partial_\lambda \mathrm{H}\|_{n,1} \ge 3^{-2(q+1)/3}\Delta_n \epsilon^{1/3}$;

(c) $\min_{\lambda\in[\lambda_\mathrm{i},\lambda_\mathrm{f}]} \|\mathrm{H}(\lambda)\| \ge \Delta_n^{3/2}\epsilon^{1/2}\|\partial_\lambda \mathrm{H}\|_{n,1}^{-1/2}$.

Then there exists a quantum algorithm $\tilde{\mathrm{U}}(\lambda_\mathrm{f}, \lambda_\mathrm{i})$ *such that*

$$\|(\mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \tilde{\mathrm{U}}(\lambda_\mathrm{f}, \lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \le O(\epsilon) \tag{5.75}$$

which can be implemented with

$$\tilde{O}\left(\left(\frac{25}{3}\right)^k k(\Lambda\delta\lambda)^{1+\frac{1}{2k}} \frac{\|\partial_\lambda \mathrm{H}\|_{n,1}^{1+\frac{1}{4k}+\frac{3}{2(q+1)}}}{\epsilon^{1+\frac{3}{4k}+\frac{3}{2(q+1)}}\Delta_n^{3+\frac{3}{4k}+\frac{3}{2(q+1)}}}\|\boldsymbol{\beta}\|_{1,\infty}\right) \tag{5.76}$$

quantum gates.
Here, $\|\boldsymbol{\beta}\|_{1,\infty} = \max_{\lambda\in[\lambda_\mathrm{i},\lambda_\mathrm{f}]} \sum_{i=1}^{\ell} |\beta_i(\lambda)|$ *and* $\Lambda = \max_{1\le p\le 2k+1} \|\partial_\lambda^p \mathrm{H}\|_{\infty,\infty}^{1/p}$.

Proof sketch.
Let $\tilde{\mathrm{U}}_{k,r}$ be as in lemma 5.4.2; assumptions (a)–(c) serve to fulfill the conditions of lemmas 5.4.1 and 5.4.2 so that, by a triangle inequality, $\|(\mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f}, \lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \in O(\epsilon)$ provided that η and a are set as in lemma 5.3.1. To finalise the proof, we must determine how $\tilde{\mathrm{U}}_{k,r}$ is decomposed into elementary quantum gates to precision ϵ (in spectral norm) and how many are needed. Since

[7] Since the qubitisation method is rather technical and its details are out of the scope of this chapter, we will use it as a black-box subroutine in the rest of the chapter. Nevertheless, the method is well known, and for readers interested in implementing our algorithm on quantum hardware or simulators, libraries exist that implement this subroutine for out-of-the-box usage. See for example ref. qt24.

$\tilde{\mathrm{U}}_{k,r}$ is a sequence of alternating exponentials of the form $\mathrm{e}^{-\mathrm{i}\mathrm{H}\theta}$ and $\mathrm{e}^{-\mathrm{i}\partial_\lambda \mathrm{H}\theta}$, we only need to count the number of such exponentials and apply the standard qubitisation result mentioned earlier. Carrying out this counting and identifying the dominant terms yields

$$\text{gate complexity} \in \tilde{O}\big(\ell 5^{k-1} r M(q+1)\big) \tag{5.77}$$

where M and q are the number of subintervals and interpolation points in the integral approximation, as before. We note that the dependence on the precision ϵ is logarithmic and has been absorbed into the $\tilde{O}(\cdots)$ notation. Inserting the values for M from lemma 5.4.1 (eq. 5.44) and r from lemma 5.4.2 (eq. 5.68), recognising that $\tilde{\Lambda} \leq 2\eta^{-1}\Lambda$ and subsequently inserting η and a from lemma 5.3.1 (eq. 5.20) then lead to the final result. □

Full proof.
Consider $\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f}, \lambda_\mathrm{i})$ as in lemma 5.4.2, as well as $\mathrm{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i})$ for $\eta, a > 0$ as in lemma 5.3.1 and $\mathrm{U}^{M,q}_{\eta,a}$ for $M \in \mathbb{N}_{>0}$ as in lemma 5.4.1. Let $\tilde{\mathrm{U}}^{\mathrm{sim}}_{k,r}$ be the simulated version of $\tilde{\mathrm{U}}_{k,r}$, that is a version of $\tilde{\mathrm{U}}_{k,r}$ which replaces all (path-independent) operator exponentials with their (qubitisation) simulations. Clearly, if

(i) $\|(\mathrm{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \leq O(\epsilon)$,

(ii) $\|(\mathrm{U}^{M,q}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \mathrm{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \leq O(\epsilon)$,

(iii) $\|(\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \mathrm{U}^{M,q}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \leq O(\epsilon)$ and

(iv) $\|(\tilde{\mathrm{U}}^{\mathrm{sim}}_{k,r}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f}, \lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \leq O(\epsilon)$

then $\|(\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}))|n(\lambda_\mathrm{i})\rangle\| \leq O(\epsilon)$ by a triangle inequality. If we set $\eta = \frac{1}{\sqrt{2}}\Delta_n^{3/2}\epsilon^{1/2}\|\partial_\lambda \mathrm{H}\|_{n,1}^{-1/2}$ and $a = \frac{1}{\eta}\log\big(2\frac{\Delta_n+\eta}{\Delta_n \epsilon \eta}\|\partial_\lambda \mathrm{H}\|_{n,1}\big)$ as in eq. 5.20, then by lemma 5.3.1, condition (i) is satisfied. By assumption (c), $\min_{\lambda\in[\lambda_\mathrm{i},\lambda_\mathrm{f}]}\|\mathrm{H}(\lambda)\| \geq \eta$, so according to lemma 5.4.1, condition (ii) can be made true by setting parameters $\{\tau_\kappa\}$, $\{w_{\kappa,\alpha}\}$ and M as in eqs. 5.41, 5.42 and 5.44. Furthermore, assumption (b) implies that $\|\partial_\lambda \mathrm{H}\|_{n,1} \geq 3^{-(q+1)}\eta \geq 3^{-(q+1)}\frac{\eta}{1-\mathrm{e}^{-\eta a}}$; assumption (a) implies that $\epsilon \leq (9/10)(5/3)^k \tilde{\Lambda}(\lambda_\mathrm{f} - \lambda_\mathrm{i})$; and the choice of γ_p and $\tilde{\Lambda}$ ensures that eq. 5.66 holds. Therefore, by lemma 5.4.2, condition (iii) is fulfilled if we set r according to eq. 5.68.

What remains is to calculate the number of quantum gates required to simulate $\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f}, \lambda_\mathrm{i})$ to precision ϵ – that is, implement $\tilde{\mathrm{U}}^{\mathrm{sim}}_{k,r}(\lambda_\mathrm{f}, \lambda_\mathrm{i})$ such that condition (iv) is fulfilled – given the settings of η, a, M, q, k and r. For this purpose, we write $\mathcal{C}_\epsilon(\cdot)$ for the gate complexity of simulating a unitary to precision ϵ in spectral norm. We first compute the gate complexity for the simulation of the first-order, single-segment decomposition $\tilde{\mathrm{U}}_{1,1}(\lambda_\mathrm{f}, \lambda_\mathrm{i})$, as given in eq. 5.63, and

subsequently extend this to k^{th}-order, r-segment formulas $\tilde{\mathrm{U}}_{k,r}(\lambda_{\mathrm{f}}, \lambda_{\mathrm{i}})$. For this purpose, we conveniently relabel the κ, α indices with an additional subscript index $i = 1, \ldots, M(q{+}1)$, as follows: $(\kappa_1, \alpha_1) = (-M, 0)$, $(\kappa_{i+1}, \alpha_{i+1}) = (\kappa_i, \alpha_i{+}1)$ if $\alpha < q$ and $(\kappa_i + 1, 0)$ otherwise. Now, an important observation is that the $\exp[\pm \mathrm{i}\mathrm{H}\tau_{\kappa,\alpha}]$ exponentials surrounding each $\exp[-\mathrm{i}\mathrm{B}_{\kappa,\alpha}\delta\lambda]$ (see eqs. 5.62–5.64) in the product formula partially cancel out, since

$$\begin{aligned}
&\exp[-\mathrm{i}\mathrm{C}_{\kappa_i,\alpha_i}\delta\lambda]\exp[-\mathrm{i}\mathrm{C}_{\kappa_{i+1},\alpha_{i+1}}\delta\lambda] \\
&= \exp[-\mathrm{i}\mathrm{H}\tau_{\kappa_i,\alpha_i}]\exp[-\mathrm{i}\mathrm{B}_{\kappa_i,\alpha_i}\delta\lambda]\exp[\mathrm{i}\mathrm{H}\tau_{\kappa_i,\alpha_i}] \\
&\quad \times \exp[-\mathrm{i}\mathrm{H}\tau_{\kappa_{i+1},\alpha_{i+1}}]\exp[-\mathrm{i}\mathrm{B}_{\kappa_{i+1},\alpha_{i+1}}\delta\lambda]\exp[\mathrm{i}\mathrm{H}\tau_{\kappa_{i+1},\alpha_{i+1}}] \\
&= \exp[-\mathrm{i}\mathrm{H}\tau_{\kappa_i,\alpha_i}]\exp[-\mathrm{i}\mathrm{B}_{\kappa_i,\alpha_i}\delta\lambda]\exp[-\mathrm{i}\mathrm{H}(\tau_{\kappa_{i+1},\alpha_{i+1}} - \tau_{\kappa_i,\alpha_i})] \\
&\quad \times \exp[-\mathrm{i}\mathrm{B}_{\kappa_{i+1},\alpha_{i+1}}\delta\lambda]\exp[\mathrm{i}\mathrm{H}\tau_{\kappa_{i+1},\alpha_{i+1}}].
\end{aligned} \tag{5.78}$$

As such, we find that the task of simulating $\tilde{\mathrm{U}}_{1,1}(\lambda_{\mathrm{f}}, \lambda_{\mathrm{i}})$ to precision ϵ has gate complexity (taking $\delta\lambda = \lambda_{\mathrm{f}} - \lambda_{\mathrm{i}}$)

$$\begin{aligned}
\mathcal{C}_\epsilon(\tilde{\mathrm{U}}_{1,1}(\lambda_{\mathrm{f}}, \lambda_{\mathrm{i}})) &= 2\Bigg(\sum_{i=1}^{M(q+1)} \mathcal{C}_{\tilde{\epsilon}}(\exp[-\mathrm{i}\mathrm{B}_{\kappa_i,\alpha_i}(\lambda_{\mathrm{i}} + \delta\lambda/2)\delta\lambda/2]) \\
&\qquad + \sum_{i=1}^{M(q+1)-1} \mathcal{C}_{\tilde{\epsilon}}(\exp[-\mathrm{i}\mathrm{H}(\lambda_{\mathrm{i}} + \delta\lambda/2)(\tau_{\kappa_{i+1},\alpha_{i+1}} - \tau_{\kappa_i,\alpha_i})] \\
&\qquad + \mathcal{C}_{\tilde{\epsilon}}(\exp[-\mathrm{i}\mathrm{H}\tau_{-M,0}]) + \mathcal{C}(\exp[-\mathrm{i}\mathrm{H}\tau_{M,q}])\Bigg) \\
&\leq \sum_{\kappa=-M}^{M}\sum_{\alpha=0}^{q} O\left(\frac{1}{2}\mathrm{e}^{-\eta|\tau_{\kappa,\alpha}|}w_{\kappa,\alpha}\delta\tau_\kappa \|\partial_\lambda \boldsymbol{\beta}(\lambda_{\mathrm{i}} + \delta\lambda/2)\|_1 \delta\lambda + \log\frac{1}{\tilde{\epsilon}}\right) O(\ell) \\
&\qquad + \sum_{i=1}^{M(q+1)-1} O\left(\|\boldsymbol{\beta}(\lambda_{\mathrm{i}} + \delta\lambda/2)\|_1(\tau_{\kappa_{i+1},\alpha_{i+1}} - \tau_{\kappa_i,\alpha_i}) + \log\frac{1}{\tilde{\epsilon}}\right) O(\ell) \\
&\qquad + 2O\left(\|\boldsymbol{\beta}(\lambda_{\mathrm{i}} + \delta\lambda/2)\|_1 a + \frac{1}{\tilde{\epsilon}}\right) O(\ell) \\
&\leq O\left(\frac{1 - \mathrm{e}^{-\eta a}}{\eta}\|\partial_\lambda \boldsymbol{\beta}\|_{1,\infty}\delta\lambda + \|\boldsymbol{\beta}\|_{1,\infty} a + M(q+1)\log\frac{1}{\tilde{\epsilon}}\right) O(\ell)
\end{aligned} \tag{5.79}$$

where $\tilde{\epsilon}$ will be specified shortly, and we used the fact that $\frac{1}{2}\sum_{\kappa,\alpha}\mathrm{e}^{-\eta|\tau_{\kappa,\alpha}|}w_{\kappa,\alpha}\delta\tau_\kappa \leq 2\frac{1-\mathrm{e}^{-\eta a}}{\eta}$ from the proof of lemma 5.4.2. Now, to generalise this to $\tilde{\mathrm{U}}_{k,r}$, observe that a k^{th}-order LTS formula is a product of 5^{k-1} first-order formulas and that an r-segment formula is a product of r single-segment formulas. At the same time, the integration interval $\delta\lambda$ is divided into smaller subintervals such that the sum of the lengths of the subintervals over all first-order, single-segment factors

is exactly $\delta\lambda$. Therefore, only the second and the third term in the last line of eq. 5.79 grow with k and r, and we obtain

$$\begin{aligned}&\mathcal{C}_\epsilon(\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f},\lambda_\mathrm{i}))\\&\quad\leq O\bigg(\frac{1-\mathrm{e}^{-\eta a}}{\eta}\|\partial_\lambda\boldsymbol{\beta}\|_{1,\infty}\delta\lambda+5^{k-1}r\|\boldsymbol{\beta}\|_{1,\infty}a+5^{k-1}rM(q+1)\log\frac{1}{\tilde{\epsilon}}\bigg)O(\ell).\end{aligned}\tag{5.80}$$

We will now specify $\tilde{\epsilon}$. Evidently, $\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f},\lambda_\mathrm{i})$ is a product of $O(5^{k-1}rM(q+1))$ (time-independent) operator exponentials. For this product to be simulated to precision $O(\epsilon)$ in operator norm, we require all constituent operator exponentials to be simulated to precision $O(\epsilon/(5^{k-1}rM(q+1)))$, since errors (in the sense of spectral distance) add up at most linearly by the telescoping property of the spectral norm. This means that we must set $\tilde{\epsilon}=\epsilon/(5^{k-1}rM(q+1))$ to ensure that condition (iv) is satisfied.

We proceed to insert the values for M from lemma 5.4.1 (eq. 5.44) and r from lemma 5.4.2 (eq. 5.68). Since r grows superlinearly with $\tilde{\Lambda}$, $\tilde{\Lambda}$ grows with η^{-1} and M grows superlinearly with a, the third term in eq. 5.80 is dominant in the regime of small η and large a; therefore

$$\begin{aligned}\mathcal{C}_\epsilon(\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f},\lambda_\mathrm{i}))&\leq O\bigg(\ell 5^{k-1}rM(q+1)\log\frac{1}{\tilde{\epsilon}}\bigg)\\&\leq\tilde{O}\bigg(\ell\bigg(\frac{25}{3}\bigg)^k k(\tilde{\Lambda}\delta\lambda)^{1+1/2k}\frac{a^{1+1/(q+1)}}{\epsilon^{1/2k+1/(q+1)}}\|\boldsymbol{\beta}\|_{1,\infty}\|\partial_\lambda\mathrm{H}\|_{n,1}^{1/(q+1)}\bigg).\end{aligned}\tag{5.81}$$

Note that the $\log 1/\tilde{\epsilon}$ has been absorbed into the $\tilde{O}(\cdots)$ notation, since the gate complexity scales polynomially in $1/\epsilon$. Lastly, we insert $a\in\tilde{O}(\eta^{-1})$ with the previously specified expression $\eta=\frac{1}{\sqrt{2}}\Delta_n^{3/2}\epsilon^{1/2}\|\partial_\lambda\mathrm{H}\|_{n,1}^{-1/2}$ to find

$$\mathcal{C}_\epsilon(\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f},\lambda_\mathrm{i}))\leq\tilde{O}\bigg(\ell\bigg(\frac{25}{3}\bigg)^k k(\tilde{\Lambda}\delta\lambda)^{1+\frac{1}{2k}}\frac{\|\partial_\lambda\mathrm{H}\|_{n,1}^{\frac{1}{2}+\frac{3}{2(q+1)}}}{\epsilon^{\frac{1}{2}+\frac{1}{2k}+\frac{3}{2(q+1)}}\Delta_n^{\frac{3}{2}+\frac{3}{2(q+1)}}}\|\boldsymbol{\beta}\|_{1,\infty}\bigg).\tag{5.82}$$

Finally, we observe that $\tilde{\Lambda}$ grows at least linearly in η^{-1}, and that $\tilde{\Lambda}\leq 2\eta^{-1}\Lambda$. Inserting this inequality into eq. 5.82, we obtain the upper bound

$$\mathcal{C}_\epsilon(\tilde{\mathrm{U}}_{k,r}(\lambda_\mathrm{f},\lambda_\mathrm{i}))\leq\tilde{O}\bigg(\ell\bigg(\frac{25}{3}\bigg)^k k(\Lambda\delta\lambda)^{1+\frac{1}{2k}}\frac{\|\partial_\lambda\mathrm{H}\|_{n,1}^{1+\frac{1}{4k}+\frac{3}{2(q+1)}}}{\epsilon^{1+\frac{3}{4k}+\frac{3}{2(q+1)}}\Delta_n^{3+\frac{3}{4k}+\frac{3}{2(q+1)}}}\|\boldsymbol{\beta}\|_{1,\infty}\bigg)\tag{5.83}$$

which is exactly eq. 5.76. The theorem is thus proved. □

5.5 Sampling approach to gate-based CD

In this section, we describe an alternative, randomised method to gate-based counterdiabatic driving. It is based on the qDRIFT algorithm [Cam19, BCS+20], which is a sampling protocol for time-dependent hamiltonian simulation. As in trotterisation-based CD, λ space is discretised, in the sense that we approximate $\mathrm{U}_{\eta,a}$ as a product of unitaries

$$\mathrm{U}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}) \approx \tilde{\mathrm{U}}_{\eta,a}(\lambda_\mathrm{f},\lambda_\mathrm{i}) = \prod_{j=1}^{r} \tilde{\mathrm{U}}_{\eta,a}(\lambda_j,\lambda_{j-1}) \tag{5.84}$$

for a given partition of the interval $[\lambda_\mathrm{i},\lambda_\mathrm{f}]$ into r subintervals $[\lambda_{j-1},\lambda_j]$ such that $\lambda_0 = \lambda_\mathrm{i}$ and $\lambda_r = \lambda_\mathrm{f}$. However, the $\tilde{\mathrm{U}}(\lambda_j,\lambda_{j-1})$ are no longer deterministically constructed, but instead sampled according to an appropriate distribution. In other words, randomised evolution is provided by a channel $\prod_{j=1}^{r}\tilde{\mathcal{U}}(\lambda_j,\lambda_{j-1})$. The idea of randomised gate-based CD is rooted in a Monte Carlo evaluation of the time integral expression for $\mathrm{A}_{\eta,a}(\lambda_j)$. In this setting, one would sample

$$\begin{aligned} \mathrm{B}_{\eta,a}(\lambda,\tau) &= \mathrm{e}^{-\mathrm{i}\mathrm{H}(\lambda)\tau}\left(\frac{(1-\mathrm{e}^{-\eta a})\operatorname{sgn}\tau}{\eta}\partial_\lambda \mathrm{H}(\lambda)\right)\mathrm{e}^{\mathrm{i}\mathrm{H}(\lambda)\tau}, \\ P(\tau) &= \begin{cases} \frac{\eta}{2(1-\mathrm{e}^{-\eta a})}\mathrm{e}^{-\eta|\tau|} & |\tau|\le a \\ 0 & \text{else} \end{cases} \end{aligned} \tag{5.85}$$

so that the expected value of $\mathrm{B}_{\eta,a}(\lambda,\tau)$ over τ equals $\mathrm{A}_{\eta,a}(\lambda)$. The procedure for randomised evolution is then as follows: for each subinterval $[\lambda_{j-1},\lambda_j]$, we sample a $\lambda\in[\lambda_{j-1},\lambda_j]$ according to a distribution $p(\lambda)$ (which will be specified later) and a $\tau\in[-a,a]$ from the distribution $P(\tau)$ defined above, and apply a unitary $\tilde{\mathrm{U}}_{\eta,a}(\lambda,\tau)$ in each case. This gives rise to the channel

$$\tilde{\mathcal{U}}_{\eta,a}(\lambda_j,\lambda_{j-1})(\rho) = \int_{\lambda_{j-1}}^{\lambda_j}\mathrm{d}\lambda p(\lambda)\int_{-\infty}^{\infty}\mathrm{d}\tau P(\tau)\,\tilde{\mathrm{U}}_{\eta,a}(\lambda,\tau)\,\rho\,\tilde{\mathrm{U}}_{\eta,a}^{\dagger}(\lambda,\tau). \tag{5.86}$$

The intuitive choice of $\tilde{\mathrm{U}}_{\eta,a}(\lambda,\tau)$, which will be used in the algorithm, is then as follows:

$$\begin{aligned} \tilde{\mathrm{U}}_{\eta,a}(\lambda,\tau) &= \exp[-\mathrm{i}\mathrm{B}_{\eta,a}(\lambda,\tau)/p(\lambda)] \\ &= \exp[-\mathrm{i}\mathrm{H}(\lambda)\tau]\exp\left[-\mathrm{i}\,\frac{(1-\mathrm{e}^{-\eta a})\operatorname{sgn}\tau}{p(\lambda)\eta}\partial_\lambda\mathrm{H}(\lambda)\right]\exp[\mathrm{i}\mathrm{H}(\lambda)\tau]. \end{aligned} \tag{5.87}$$

We will now prove a bound on the number of quantum gates that is required to upper bound the resulting error in terms of the trace distance between the constructed channel and exact evolution. As in theorem 5.4.3, we assume that the hamiltonian H can be decomposed as a linear combination of finitely many unitaries, $\mathrm{H}(\lambda) = \sum_{j=1}^{\ell}\beta_j(\lambda)\mathrm{V}_j$ (this assumption is always fulfilled in finite-dimensional Hilbert spaces).

5.5.1. Theorem (randomised gate-based CD). *Suppose* $\mathrm{H}(\lambda) = \sum_{i=1}^{\ell} \beta_i(\lambda)\mathrm{V}_i$ *is hermitian and differentiable on the interval* $[\lambda_\mathrm{i}, \lambda_\mathrm{f}]$. *Define* $\tilde{\mathcal{U}}_{\eta,a}(\lambda_j, \lambda_{j-1})$ *and* $\tilde{\mathrm{U}}_{\eta,a}(\lambda, \tau)$ *as in eqs. 5.86 and 5.87 respectively, for any partition* $\lambda_\mathrm{i} = \lambda_0 < \cdots < \lambda_r = \lambda_\mathrm{f}$ *of* $[\lambda_\mathrm{i}, \lambda_\mathrm{f}]$. *Let* $\mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) = \mathcal{T}\exp[-\mathrm{i}\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\, \mathrm{A}(\lambda)]$ *and* $\mathcal{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i})(\rho) = \mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i})\rho\mathrm{U}^\dagger(\lambda_\mathrm{f}, \lambda_\mathrm{i})$, *and let* $\rho_n(\lambda) = |n(\lambda)\rangle\langle n(\lambda)|$. *Then there exists a partition and a distribution* $p(\lambda)$ *such that*

$$\left\| \mathcal{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i})(\rho_n(\lambda_\mathrm{i})) - \prod_{j=1}^{r} \tilde{\mathcal{U}}_{\eta,a}(\lambda_j, \lambda_{j-1})(\rho_n(\lambda_\mathrm{i})) \right\|_1 \leq \epsilon \tag{5.88}$$

and the channel $\prod_{j=1}^{r} \tilde{\mathcal{U}}_{\eta,a}(\lambda_j, \lambda_{j-1})$ *can be implemented with*

$$\tilde{O}\left(\Delta_n^{-9/2} \epsilon^{-5/2} \|\partial_\lambda \mathrm{H}\|_{\infty,1}^{7/2} \ell \left[\max_\lambda \|\boldsymbol{\beta}(\lambda)\|_1 + \|\partial_\lambda \mathrm{H}\|_{\infty,1} \max_\lambda \frac{\|\partial_\lambda \boldsymbol{\beta}(\lambda)\|_1}{\|\partial_\lambda \mathrm{H}(\lambda)\|} \right] \right) \tag{5.89}$$

quantum gates.

Proof.
Define $\mathcal{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i})(\rho) = \mathrm{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i})\rho\mathrm{U}_{\eta,a}^\dagger(\lambda_\mathrm{f}, \lambda_\mathrm{i})$ where as usual $\mathrm{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) = \mathcal{T}\exp[-\mathrm{i}\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\, \mathrm{A}_{\eta,a}(\lambda)]$ and $\mathrm{A}_{\eta,a}(\lambda)$ is the regularised truncated AGP as in 5.18. From lemma 5.3.1, it follows that $\|\mathcal{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i})(\rho_n(\lambda_\mathrm{i})) - \mathcal{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i})(\rho_n(\lambda_\mathrm{i}))\|_1 \leq \epsilon$ if we set η, a as in eq. 5.20. We are therefore interested in upper bounding the error

$$\left\| \mathcal{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i})(\rho_n(\lambda_\mathrm{i})) - \prod_{j=1}^{r} \tilde{\mathcal{U}}_{\eta,a}(\lambda_j, \lambda_{j-1})(\rho_n(\lambda_\mathrm{i})) \right\|_1 . \tag{5.90}$$

This error may be telescoped into multiple short-evolution errors:

$$(\text{eq. } 5.90) \leq \sum_{j=1}^{r} \|\mathcal{U}_{\eta,a}(\lambda_j, \lambda_{j-1})(\rho_n(\lambda_{j-1})) - \tilde{\mathcal{U}}_{\eta,a}(\lambda_j, \lambda_{j-1})(\rho_n(\lambda_{j-1}))\|_1. \tag{5.91}$$

We will now bound each term following a similar strategy to that of Berry et al. [BCS+20]. First we define scaled versions of $\mathcal{U}_{\eta,a}$ and $\tilde{\mathcal{U}}_{\eta,a}$:

$$\begin{aligned} \mathcal{U}_{\eta,a,s}(\lambda_j, \lambda_{j-1})(\rho) &= \mathrm{U}_{\eta,a,s}(\lambda_j, \lambda_{j-1})\rho\mathrm{U}_{\eta,a,s}^\dagger(\lambda_j, \lambda_{j-1}), \\ \mathrm{U}_{\eta,a,s}(\lambda_j, \lambda_{j-1}) &= \mathcal{P}\exp\left[-\mathrm{i}\int_{\lambda_{j-1}}^{\lambda_j} \mathrm{d}\lambda\, s\, \mathrm{A}_{\eta,a}(\lambda)\right]; \end{aligned} \tag{5.92}$$

$$\begin{aligned} \tilde{\mathcal{U}}_{\eta,a,s}(\lambda_j, \lambda_{j-1})(\rho) &= \int_{\lambda_{j-1}}^{\lambda_j} \mathrm{d}\lambda p(\lambda) \int_{-\infty}^{\infty} \mathrm{d}\tau P(\tau)\, \tilde{\mathrm{U}}_{\eta,a,s}(\lambda, \tau)\, \rho\, \tilde{\mathrm{U}}_{\eta,a,s}^\dagger(\lambda, \tau), \\ \tilde{\mathrm{U}}_{\eta,a,s}(\lambda, \tau) &= \exp[-\mathrm{i}\, s\, \mathrm{B}_\eta(\lambda, \tau)/p(\lambda)]. \end{aligned} \tag{5.93}$$

Since $\mathcal{U}_{\eta,0}(\rho) = \tilde{\mathcal{U}}_{\eta,0}(\rho) = \rho$, we may write (dropping the λ_j, λ_{j-1} arguments for legibility)

$$\begin{aligned}\|\mathcal{U}_{\eta,a}(\rho_n) - \tilde{\mathcal{U}}_{\eta,a}(\rho_n)\|_1 &= \big\| \{\mathcal{U}_{\eta,1}(\rho_n) - \mathcal{U}_{\eta,0}(\rho_n)\} - \{\tilde{\mathcal{U}}_{\eta,1}(\rho_n) - \tilde{\mathcal{U}}_{\eta,0}(\rho_n)\} \big\|_1 \\ &= \left\| \int_0^1 \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s} \big(\mathcal{U}_{\eta,a,s}(\rho_n) - \tilde{\mathcal{U}}_{\eta,a,s}(\rho_n)\big) \right\|. \end{aligned} \tag{5.94}$$

For the derivatives in the last line, we use lemma 6 from Berry et al. [BCS+20],

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathrm{U}_{\eta,a,s}(\lambda_j, \lambda_{j-1}) = \int_{\lambda_{j-1}}^{\lambda_j} \mathrm{d}\lambda\, \mathrm{U}_{\eta,a,s}(\lambda_j, \lambda)[-\mathrm{i}\mathrm{A}_{\eta,a}(\lambda)]\mathrm{U}_{\eta,a,s}(\lambda, \lambda_{j-1}), \tag{5.95}$$

and

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s} \tilde{\mathcal{U}}_{\eta,a,s}(\lambda_j, \lambda_{j-1})(\rho) = \int_{\lambda_{j-1}}^{\lambda_j} \mathrm{d}\lambda p(\lambda) \int_{-a}^{a} \mathrm{d}\tau\, \frac{\eta}{2(1-\mathrm{e}^{-\eta a})} \mathrm{e}^{-\eta|\tau|} \\ \times \mathrm{e}^{-\mathrm{i}\mathrm{H}(\lambda)\tau} \mathrm{e}^{-\mathrm{i}\,(1-\mathrm{e}^{-\eta a})\, s\, \mathrm{sgn}\,\tau\, \partial_\lambda \mathrm{H}(\lambda)/p(\lambda)\eta} \\ \times \Big[-\mathrm{i}\,\frac{(1-\mathrm{e}^{-\eta a})\,\mathrm{sgn}\,\tau}{p(\lambda)\eta}\, \partial_\lambda \mathrm{H}(\lambda),\, \mathrm{e}^{\mathrm{i}\mathrm{H}(\lambda)\tau} \rho\, \mathrm{e}^{-\mathrm{i}\mathrm{H}(\lambda)\tau}\Big] \\ \times \mathrm{e}^{\mathrm{i}\,(1-\mathrm{e}^{-\eta a})\, s\, \mathrm{sgn}\,\tau\, \partial_\lambda \mathrm{H}(\lambda)/p(\lambda)\eta} \mathrm{e}^{\mathrm{i}\mathrm{H}(\lambda)\tau}. \end{aligned} \tag{5.96}$$

We observe now that the derivatives of $\mathcal{U}_{\eta,a,s}$ and $\tilde{\mathcal{U}}_{\eta,a,s}$ agree at $s = 0$:

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathcal{U}_{\eta,a,s}(\lambda_j, \lambda_{j-1})(\rho)\Big|_{s=0} = \int_{\lambda_{j-1}}^{\lambda_j} \mathrm{d}\lambda\, [-\mathrm{i}\mathrm{A}_{\eta,a}(\lambda), \rho] = \frac{\mathrm{d}}{\mathrm{d}s} \tilde{\mathcal{U}}_{\eta,a,s}(\lambda_j, \lambda_{j-1})(\rho)\Big|_{s=0}; \tag{5.97}$$

hence, we may invoke the fundamental theorem of calculus a second time to express

$$\|\mathcal{U}_{\eta,a}(\rho_n) - \tilde{\mathcal{U}}_{\eta,a}(\rho_n)\|_1 = \left\| \int_0^1 \mathrm{d}s \int_0^s \mathrm{d}v \frac{\mathrm{d}^2}{\mathrm{d}v^2} \big(\mathcal{U}_{\eta,a,v}(\rho_n) - \tilde{\mathcal{U}}_{\eta,a,v}(\rho_n)\big) \right\|. \tag{5.98}$$

We proceed to bound this error following Berry et al. [BCS+20, proof of theorem 5]:

$$\begin{aligned}&\|\mathcal{U}_{\eta,a}(\rho) - \tilde{\mathcal{U}}_{\eta,a}(\rho)\|_1 \\ &\leq \int_0^1 \mathrm{d}s \int_0^s \mathrm{d}v \left\{ \left\| \frac{\mathrm{d}^2}{\mathrm{d}v^2} \mathrm{U}_{\eta,a,v}\, \rho\, \mathrm{U}_{\eta,a,v}^\dagger \right\|_1 + 2 \left\| \frac{\mathrm{d}}{\mathrm{d}v} \mathrm{U}_{\eta,a,v}\, \rho\, \frac{\mathrm{d}}{\mathrm{d}v} \mathrm{U}_{\eta,a,v}^\dagger \right\|_1 + \left\| \mathrm{U}_{\eta,a,v}\, \rho\, \frac{\mathrm{d}^2}{\mathrm{d}v^2} \mathrm{U}_{\eta,a,v}^\dagger \right\|_1 \right. \\ &\quad + \int_{\lambda_{j-1}}^{\lambda_j} \mathrm{d}\lambda\, p(\lambda) \int_{-\infty}^{\infty} \mathrm{d}\tau\, P(\tau) \\ &\quad \left. \times \left\| \Big[-\mathrm{i}\,\frac{(1-\mathrm{e}^{-\eta a})\,\mathrm{sgn}\,\tau}{p(\lambda)\eta}\, \partial_\lambda \mathrm{H}(\lambda),\, \Big[-\mathrm{i}\,\frac{(1-\mathrm{e}^{-\eta a})\,\mathrm{sgn}\,\tau}{p(\lambda)\eta}\, \partial_\lambda \mathrm{H}(\lambda),\, \mathrm{e}^{\mathrm{i}\mathrm{H}(\lambda)\tau} \rho\, \mathrm{e}^{-\mathrm{i}\mathrm{H}(\lambda)\tau} \Big] \Big] \right\|_1 \right\} \\ &\leq \int_0^1 \mathrm{d}s \int_0^s \mathrm{d}v \left\{ 4 \left(\int_{\lambda_{j-1}}^{\lambda_j} \mathrm{d}\lambda\, \|\mathrm{A}_{\eta,a}(\lambda)\| \right)^2 + 4 \int_{\lambda_{j-1}}^{\lambda_j} \mathrm{d}\lambda\, p(\lambda) \left(\frac{(1-\mathrm{e}^{-\eta a}) \|\partial_\lambda \mathrm{H}(\lambda)\|}{p(\lambda)\eta} \right)^2 \right\}. \end{aligned} \tag{5.99}$$

Now from the definition of $\mathrm{A}_{\eta,a}$, we find

$$\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda \, \|\mathrm{A}_{\eta,a}\| \leq \int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda \int_{-a}^{a} \mathrm{d}\tau \, \frac{1}{2} \mathrm{e}^{-\eta|\tau|} \|\partial_\lambda \mathrm{H}\| = \frac{1-\mathrm{e}^{-\eta a}}{\eta} \|\partial_\lambda \mathrm{H}\|_{\infty,1} \tag{5.100}$$

where $\|\partial_\lambda \mathrm{H}\|_{\infty,1} = \int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda \|\partial_\lambda \mathrm{H}(\lambda)\|$; if we now conveniently choose

$$p(\lambda) = \frac{\|\partial_\lambda \mathrm{H}(\lambda)\|}{\|\partial_\lambda \mathrm{H}\|_{\infty,1}} \tag{5.101}$$

then the last line of eq. 5.99 reduces to

$$\|\mathcal{U}_{\eta,a}(\lambda_j, \lambda_{j-1})(\rho) - \tilde{\mathcal{U}}_{\eta,a}(\lambda_j, \lambda_{j-1})(\rho)\|_1 \leq \frac{4(1-\mathrm{e}^{-\eta a})^2}{\eta^2} \|\partial_\lambda \mathrm{H}\|_{\infty,1}^2. \tag{5.102}$$

Lastly, we can directly follow the proof of theorem 7 from Berry et al. [BCS+20] to show that

$$r \geq \frac{4(1-\mathrm{e}^{-\eta a})^2}{\eta^2 \epsilon} \|\partial_\lambda \mathrm{H}\|_{\infty,1}^2 \tag{5.103}$$

is sufficient to guarantee that the error from eq. 5.90 satisfies

$$\left\| \mathcal{U}_{\eta,a}(\lambda_\mathrm{f}, \lambda_\mathrm{i})(\rho_n(\lambda_\mathrm{i})) - \prod_{j=1}^{r} \tilde{\mathcal{U}}_{\eta,a}(\lambda_j, \lambda_{j-1})(\rho_n(\lambda_\mathrm{i})) \right\|_1 \leq \epsilon. \tag{5.104}$$

We thus find that one should sample and apply a unitary $\tilde{\mathrm{U}}_{\eta,a}$ $\lceil r \rceil$ times, with r given by eq. 5.103, to achieve a total error of at most 2ϵ. But each of these unitaries will have to be simulated, which can be done through known hamiltonian simulation techniques. With qubitisation [LC19b], one can simulate an operator exponential $\exp[-\mathrm{i}\mathrm{H}\theta]$ of a hamiltonian $\mathrm{H} = \sum_{j=1}^{\ell} \beta_j \mathrm{V}_j$ to error ϵ with $O([\|\boldsymbol{\beta}\|_1 |\theta| + \log 1/\epsilon]\ell)$ elementary gates. For the simulation error of a product of r unitaries to be at most ϵ, each unitary should be simulated with error at most ϵ/r. Observing that $|\tau| \leq a$, $1 - \mathrm{e}^{\eta a} \leq 1$, $a \in \tilde{O}(\eta^{-1})$ and $\eta = \frac{1}{\sqrt{2}} \Delta_n^{3/2} \epsilon^{1/2} \|\partial_\lambda \mathrm{H}\|_{\infty,1}^{-1/2}$, we find the total gate complexity of simulating our channel to error ϵ in trace norm to be

$$\begin{aligned}
&\text{gate complexity} \\
&\leq O\left(r\ell \left[a \max_\lambda \|\boldsymbol{\beta}(\lambda)\|_1 + \log \frac{r}{\epsilon} \right] \right) + O\left(r\ell \left[\max_\lambda \frac{\|\partial_\lambda \boldsymbol{\beta}(\lambda)\|_1}{p(\lambda)\eta} + \log \frac{r}{\epsilon} \right] \right) \\
&\leq \tilde{O}\left(\frac{\|\partial_\lambda \mathrm{H}\|_{\infty,1}^2 \ell}{\eta^3 \epsilon} \left[\max_\lambda \|\boldsymbol{\beta}(\lambda)\|_1 + \max_\lambda \frac{\|\partial_\lambda \boldsymbol{\beta}(\lambda)\|_1 \|\partial_\lambda \mathrm{H}\|_{\infty,1}}{\|\partial_\lambda \mathrm{H}(\lambda)\|} \right] \right) \\
&\leq \tilde{O}\left(\Delta_n^{-9/2} \epsilon^{-5/2} \|\partial_\lambda \mathrm{H}\|_{\infty,1}^{7/2} \ell \left[\max_\lambda \|\boldsymbol{\beta}(\lambda)\|_1 + \|\partial_\lambda \mathrm{H}\|_{\infty,1} \max_\lambda \frac{\|\partial_\lambda \boldsymbol{\beta}(\lambda)\|_1}{\|\partial_\lambda \mathrm{H}(\lambda)\|} \right] \right)
\end{aligned} \tag{5.105}$$

which proves the theorem. □

5.6 Comparison to gate-based adiabatic computing

We now return to the matter of comparing counterdiabatic driving to adiabatic computing. The complexity of AQC is typically quantified by the physical evolution time T which sets the speed at which $\lambda(t)$ is varied. However, in the previous sections we have quantified the complexity of CD in terms of the number of digital quantum operations (quantum gates); as such, physical time, with which the performance of adiabatic computing is typically quantified, is insufficient here since it is not a meaningful quantity in the description of gate-based CD and therefore leads to an apples-to-oranges comparison. Instead, we will describe the complexity of AQC in the same gate-based fashion, counting the number of quantum gates with which gate-based AQC can be implemented. In this context, the standard adiabatic theorems which state bounds on the physical time are still important, since in the end, the number of quantum gates that is needed for gate-based AQC is mostly determined by the physical time.

Several versions of the adiabatic theorem exist. Nonrigorous, "folk" versions typically estimate $T \in \Omega(\Delta_n^{-2})$; but since we conducted a rigorous complexity analysis of CD, such statements do not suffice for a comparison. Rigorous theorems typically yield a cubic or near-quadratic worst-case scaling in Δ_n^{-1}, though a quadratic dependence may be obtained [EH12] if it is assumed that $\mathrm{H}(\lambda)$ belongs to a Gevrey class[8]. A general result, and, to the best of our knowledge, the tightest bound in terms of the minimum gap Δ_n is due to Reichardt [Rei04]. We quote this result here.

5.6.1. Lemma (adapted from Reichardt [Rei04]). *Let $T > 0$, and let $\lambda(\cdot) : \mathbb{R} \to \mathbb{R}$ be a function such that $\lambda(0) = \lambda_\mathrm{i}$ and $\lambda(T) = \lambda_\mathrm{f}$. Suppose $\mathrm{H}(\lambda)$ is hermitian, gapped and $K+1$ times differentiable in λ on the interval $[\lambda_\mathrm{i}, \lambda_\mathrm{f}]$, with $p \geq 1$. Let $\mathrm{U}(T,0) = \mathcal{T}\exp[-\mathrm{i}\int_0^T \mathrm{d}t\,\mathrm{H}(\lambda(t))]$. Let $|n(\lambda_\mathrm{i})\rangle$ be the the n-th instantaneous eigenstate of $\mathrm{H}(\lambda_\mathrm{i})$, and $|n(\lambda_\mathrm{f})\rangle = \lim_{T\to\infty}\mathrm{U}(T,0)|n(\lambda_\mathrm{i})\rangle$ (i.e. the n-th eigenstate of $\mathrm{H}(\lambda_\mathrm{f})$), and assume $\|\partial_\lambda^p \mathrm{H}(\lambda)\| \leq \Gamma$ for all $s \leq K+1$. Then*

$$|\langle n(\lambda_\mathrm{f})|\mathrm{U}(T,0)|n(\lambda_\mathrm{i})\rangle|^2 \geq 1-\epsilon^2 \tag{5.106}$$

provided that

$$T \geq \Omega(\epsilon^{-1/K}\Delta_n^{-(2+1/K)}\Gamma^{1+1/K}) \tag{5.107}$$

where Δ_n is the minimum spectral gap around $|n(\lambda)\rangle$ on the interval $[\lambda_\mathrm{i}, \lambda_\mathrm{f}]$.

We proceed to put a bound on the gate complexity of gate-based AQC by combining lemmas 5.6.1 and 5.2.1. Lemma 5.2.1 gives a bound on the number

[8] A hamiltonian $\mathrm{H}(\lambda)$ belongs to a Gevrey class G^α if $\partial_\lambda \mathrm{H} \neq 0\ \forall\lambda \in [\lambda_\mathrm{i}, \lambda_\mathrm{f}]$ and there exist constants $C, D > 0$ such that for all $p > 1$, $\max_{\lambda\in[\lambda_\mathrm{i},\lambda_\mathrm{f}]} \|\partial_\lambda^p \mathrm{H}(\lambda)\| \leq CD^p k^{\alpha k}$.

of operator exponentials with which a time evolution in a time interval of given length can be simulated by a LTS decomposition (cf. section 5.2.2). If we assume the hamiltonian is given as a linear combination of unitaries $\mathrm{H}(t) = \sum_{i=1}^{\ell} \beta_i(t)\mathrm{V}_i$, every exponential is a multi-qubit rotation gate; if required, such gates can be decomposed into a number of elementary gates (single-qubit rotations, CNOT gates and the like) that depends only on the locality of the respective unitary V_i (i.e. the number of qubit it acts on nontrivially). We will consider the maximum locality of these terms as a constant, and as such, in big-O notation, we equate the number of elementary gates to the number of (time-independent) operator exponentials. The following then gives a bound on the number of operator exponentials to implement gate-based AQC.

5.6.2. Proposition (gate-based AQC). *Let $T > 0$ as in lemma 5.6.1, and let $\lambda(t)$ be such that $\lambda(0) = \lambda_\mathrm{i}$ and $\lambda(T) = \lambda_\mathrm{f}$. Suppose $\mathrm{H}(\lambda(t)) = \sum_{i=1}^{\ell} \beta_i(\lambda(t))\mathrm{V}_i$ is hermitian and $2k+1$ times differentiable in t on the interval $[0, T]$, with $\beta_i(\lambda(t)) \in \mathbb{R}$ and V_i unitary. Let $\{|n(\lambda(t))\rangle\}$ be the set of its instantaneous eigenstates. Suppose the spectrum of $\mathrm{H}(\lambda(t))$ around an eigenstate $|n(\lambda(t))\rangle$ has a gap of at least Δ_n on the entire interval. Let $\mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) = \sum_n |n(\lambda_\mathrm{f})\rangle\langle n(\lambda_\mathrm{i})|$. Assume $\|\partial_\lambda^p \mathrm{H}(\lambda)\| \leq \Gamma$ for all $p \leq 2k+1$ and $T \in \Theta(\epsilon^{-1/2k} \Delta_n^{-(2+1/2k)} \Gamma^{1+1/2k})$. Then there exists a gate-based quantum algorithm $\tilde{\mathrm{U}}(T, 0)$ such that*

$$\|(\mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \tilde{\mathrm{U}}(T, 0))|n(\lambda_\mathrm{i})\rangle\| \leq O(\epsilon) \tag{5.108}$$

which can be implemented with

$$O\left(\ell \left(\frac{25}{3}\right)^k k \Lambda^{1+\frac{1}{2k}} \frac{\Gamma^{(1+\frac{1}{2k})^2}}{\epsilon^{\frac{1}{2k}(2+\frac{1}{2k})} \Delta_n^{(2+\frac{1}{2k})(1+\frac{1}{2k})}}\right) \tag{5.109}$$

quantum gates.
Here, $\Lambda = \max_{0 \leq p \leq 2k} \|\partial_t^p \boldsymbol{\beta}\|_{1,\infty}^{1/(p+1)}$ and $\|\partial_t^p \boldsymbol{\beta}\|_{1,\infty} = \max_{t \in [0,T]} \sum_{i=1}^{\ell} |\partial_t^p \beta_i|$.

Proof.
Let $\mathrm{U}(T, 0) = \mathcal{T} \exp[-\mathrm{i} \int_0^T \mathrm{d}t\, \mathrm{H}(t)]$ and identify $\tilde{\mathrm{U}}(T, 0)$ as a k^{th}-order r-segment LTS decomposition of $\mathrm{U}(T, 0)$. Observe that a first-order single-segment LTS decomposition $\tilde{\mathrm{U}}_{1,1}$ of $\mathrm{U}(T, 0)$ (eq. 5.12) contains 2ℓ operator exponentials, and that $\tilde{\mathrm{U}}_{k,r}$ is a product of $2\ell 5^{k-1} r$ operator exponentials. Lemma 5.6.1 gives the conditions under which $\|(\mathrm{U}(\lambda_\mathrm{f}, \lambda_\mathrm{i}) - \mathrm{U}(T, 0))|n(\lambda_\mathrm{i})\rangle\| \leq O(\epsilon)$; we use the bound on r from lemma 5.2.1, which assures that $\|(\mathrm{U}(T, 0) - \tilde{\mathrm{U}}(T, 0))|n(\lambda_\mathrm{i})\rangle\| \leq O(\epsilon)$ so that, by a triangle inequality, eq. 5.108 holds. To use this bound, we first notice that the choice $\Lambda = \max_{0 \leq p \leq 2k} \|\partial_t^p \boldsymbol{\beta}\|_{1,\infty}^{1/(p+1)}$ is sufficient to fulfill the requirement of eq. 5.16 in lemma 5.2.1. From lemma 5.6.1 we then insert T to obtain that

$$O\left(\ell \left(\frac{25}{3}\right)^k k \Lambda^{1+\frac{1}{2k}} T^{1+\frac{1}{2k}} \epsilon^{-\frac{1}{2k}}\right) = O\left(\ell \left(\frac{25}{3}\right)^k k \Lambda^{1+\frac{1}{2k}} \frac{\Gamma^{(1+\frac{1}{2k})^2}}{\epsilon^{\frac{1}{2k}(2+\frac{1}{2k})} \Delta_n^{(2+\frac{1}{2k})(1+\frac{1}{2k})}}\right) \tag{5.110}$$

quantum gates are sufficient to guarantee an overall error of at most $O(\epsilon)$. □

In short, we find that the physical time $T \in \Theta(\epsilon^{-o(1)}\Delta_n^{-(2+o(1))})$ leads to a similar gate complexity upper bound of $O(\epsilon^{-o(1)}\Delta_n^{-(2+o(1))})$.

Besides the bound by Reichardt and the subsequent gate complexity, another result that deserves a mention is that by Jansen et al. [JRS07]. Their bound is $T \geq \Omega\big(\epsilon^{-1}\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\, \|\partial_\lambda \mathrm{H}(\lambda)\|^2/\omega_n(\lambda)^3\big)$ for hamiltonians that are twice differentiable, where $\omega_n(\lambda)$ is the instantaneous gap around the eigenstate $|n(\lambda)\rangle$ of $\mathrm{H}(\lambda)$. At first glance, this seems equivalent to Reichardt's result for $K = 1$. However, the authors point out that, at least for unstructured (Grover) search, the result should not be viewed as an inverse cubic dependence on the mimimum gap. They show that instead, $\int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\, \omega_n(\lambda)^{-3} \in \Theta(\Delta_n^{-2})$, and that further improvement is achieved by tuning the interpolation function $f(\lambda)$ to the instantaneous gap, which leads to a linear scaling in Δ_n^{-1}. Similar results were obtained for the quantum linear systems problem (QLSP) by Cunningham et al. [CR24]. Looking at proposition 5.6.2, we would therefore also find a quadratic or linear inverse gap scaling in the number of quantum gates.

Comparing the $\tilde{O}(\epsilon^{-(1+o(1))}\Delta_n^{-(3+o(1))})$ bound from theorem 5.4.3 for CD to the $\tilde{O}(\epsilon^{-o(1)}\Delta_n^{-(2+o(1))})$ from proposition 5.6.2 for AQC, it appears that there is room for improvement in the analysis of gate-based CD. One such improvement can be easily spotted: in the proof of lemma 5.3.1, we determined η from the condition $2\eta^2 \int_{\lambda_\mathrm{i}}^{\lambda_\mathrm{f}} \mathrm{d}\lambda\, \frac{\|\partial_\lambda \mathrm{H}(\lambda)\|}{\omega_n^3(\lambda)} \leq 2\eta^2 \|\partial_\lambda \mathrm{H}\|_{\infty,1}\Delta_n^{-3} \leq \epsilon$. However, in the same vein as the argument put forward by Jansen et al. [JRS07], the integral may in certain cases be reduced to $\Theta(\Delta_n^{-2})$, which leads to a linear gap dependence $\eta \in \Theta(\epsilon^{1/2}\Delta_n)$. Since the gate complexity of CD depends at most quadratically on η^{-1} (up to logarithmic factors), we obtain an (almost) inverse quadratic overall gap scaling instead of inverse cubic. Improving this to linear for problems like unstructured search is however less straightforward, and likely requires a variable $\eta(\lambda)$ throughout the driving, which may be tuned to the instantaneous gap, instead of a constant η. In principle, one could adopt optimisation schemes to find good cutoff functions $\eta(\lambda)$, in a way similar to optimisation of scheduling functions [CSL+23]. However, given the difficulty of analysing optimisation methods, this would add a layer of uncertainty regarding the computational complexity of the complete procedure.

5.7 Conclusion

In this chapter, we have presented what is, to the best of our knowledge, the first fully gate-based quantum algorithm for counterdiabatic driving. This algorithm is constructed from the regularised truncated adiabatic gauge potential (eq. 5.18). By discretising the integral form of this AGP approximation, it is fed into a

Lie-Trotter-Suzuki formula to produce a gate-based algorithm that may be run on a quantum computer. We have shown that this algorithm, starting from an initial eigenstate $|n(\lambda_\mathrm{i})\rangle$, can be run with $\tilde{O}(\epsilon^{-(1+o(1))}\Delta_n^{-(3+o(1))})$ quantum gates in the worst case in order to achieve a fidelity at least $1-\epsilon^2$ with the target eigenstate $|n(\lambda_\mathrm{f})\rangle$ (theorem 5.4.3). Here, Δ_n is the minimum energy gap around the instantaneous eigenstate $|n(\lambda)\rangle$. The $o(1)$ scaling is an inverse linear dependence on the order k of the LTS formula (see eqs. 5.63 and 5.13) and the degree q of the Lagrange interpolation polynomial used in the discretisation of the integral form. We remark that q can be made large cheaply using purely classical precomputation of the quadrature weights.

This scaling is almost equivalent to the $O(\Delta^{-(3+o(1))}\epsilon^{-(1+o(1))})$ gate complexity bound that follows from applying LTS simulation to the worst-case result of Jansen et al. [JRS07] (up to logarithmic factors), but is polynomially worse than the $O(\Delta^{-(2+o(1))}\epsilon^{-o(1)})$ bound that results in a similar way from the scaling given by Reichardt [Rei04]. However, an inverse quadratic gap scaling of the form $\tilde{O}(\Delta^{-(2+o(1))}\epsilon^{-1+o(1)})$ may be achieved with our gate-based CD method in certain cases such as unstructured (Grover) search and the QLSP problem.

All in all, CD remains an interesting approach for quantum state preparation. After all, there may exist more efficient gate-based CD algorithms or better complexity bounds. And while CD cannot be generally faster than AQC, at least in terms of gap scaling – this would violate Grover optimality – our work opens up a new formalism for finding gate complexity upper bounds based on an approximation of the AGP.

Furthermore, CD can still be valuable in settings where quantum resources are scarce. For example, in noisy setups where the ability to work on small timescales is paramount to countering noise, adiabaticity is not available; achieving satisfactory fidelities then necessitates the use of some kind of classically precomputed shortcut field. We also remark that the algorithm presented in this chapter is intended as a fault-tolerant routine, whose implementation may remain out of reach for the near future, especially for settings with small minimum gaps which lead to deep circuits.

Lastly, we remark that, in order to obtain the presented gate complexity of CD, it turned out crucial to tailor the AGP approximation to the eigenstate of interest, through an appropriately chosen nonzero gap cutoff. This insight that it is unnecessary to suppress all transitions in the spectrum allowed us to reduce the complexity down from the dimensionality of the operator space to a scaling in the minimum gap around that specific eigenstate. While this idea has been recognised [DR08, MP24], it was not taken into account in the currently most prevalent algorithmic CD approaches [SP17, CPSP19, Bha23, TdC24]. As such, we would like to stress its importance once again, and hope to see it replicated in future work on counterdiabatic driving.

Bibliography

[ADP16] Fabien Alet, Kedar Damle, and Sumiran Pujari. Sign-problem-free monte carlo simulation of certain frustrated quantum magnets. *Phys. Rev. Lett.*, 117:197203, Nov 2016.

[AL18] Tameem Albash and Daniel A. Lidar. Adiabatic quantum computation. *Reviews of Modern Physics*, 90(1), January 2018.

[Ami09] M. H. S. Amin. Consistency of the adiabatic theorem. *Physical Review Letters*, 102(22):220401, jun 2009.

[ATS03] Dorit Aharonov and Amnon Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. *Conference Proceedings of the Annual ACM Symposium on Theory of Computing*, pages 20–29, 1 2003.

[AvDK+07] Dorit Aharonov, Wim van Dam, Julia Kempe, Zeph Landau, Seth Lloyd, and Oded Regev. Adiabatic Quantum Computation is Equivalent to Standard Quantum Computation. *SIAM Journal on Computing*, 37(1):166–194, January 2007.

[BBGP16] Johannes A. Buchmann, Denis Butin, Florian Göpfert, and Albrecht Petzoldt. *Post-Quantum Cryptography: State of the Art*, pages 88–108. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016.

[BCG+24] Sergey Bravyi, Andrew W. Cross, Jay M. Gambetta, Dmitri Maslov, Patrick Rall, and Theodore J. Yoder. High-threshold and low-overhead fault-tolerant quantum memory. *Nature*, 627(8005):778–782, March 2024.

[BCNP22] Budhaditya Bhattacharjee, Xiangyu Cao, Pratik Nandy, and Tanay Pathak. Krylov complexity in saddle-dominated scrambling. *Journal of High Energy Physics*, 2022(5), 5 2022.

[BCS+20] Dominic W. Berry, Andrew M. Childs, Yuan Su, Xin Wang, and Nathan Wiebe. Time-dependent Hamiltonian simulation with L^1-norm scaling. *Quantum*, 4:254, 4 2020.

[Bel64] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics Physique Fizika*, 1:195–200, Nov 1964.

[BF28] M. Born and V. Fock. Beweis des Adiabatensatzes. *Zeitschrift für Physik*, 51(3):165–180, March 1928.

[BGB+18] Ryan Babbush, Craig Gidney, Dominic W. Berry, Nathan Wiebe, Jarrod McClean, Alexandru Paler, Austin Fowler, and Hartmut Neven. Encoding electronic spectra in quantum circuits with linear T complexity. *Physical Review X*, 8:041015, 10 2018.

[BGG+20] F. Boudot, P. Gaudry, A. Guillevic, N. Heninger, E. Thomé, and P. Zimmermann. Comparing the difficulty of factorization and discrete logarithm: a 240-digit experiment. LISTSERV – NMBRTHRY archives, `https://listserv.nodak.edu/cgi-bin/wa.exe?A2=NMBRTHRY;dc42ccd1.2002`, 2020.

[Bha13] Rajendra Bhatia. *Matrix analysis*, volume 169. Springer Science & Business Media, 2013.

[Bha23] Budhaditya Bhattacharjee. A Lanczos approach to the adiabatic gauge potential. arXiv:2302.07228, 2023.

[BK02] Sergey B. Bravyi and Alexei Yu. Kitaev. Fermionic quantum computation. *Annals of Physics*, 298(1):210–226, May 2002.

[BKS09] Sergio Boixo, Emanuel Knill, and Rolando Somma. Eigenpath traversal by phase randomization. *Quantum Info. Comput.*, 9(9):833–855, sep 2009.

[BKS10] S. Boixo, E. Knill, and R. D. Somma. Fast quantum algorithms for traversing paths of eigenstates. *arXiv: Quantum Physics*, 5 2010.

[BLS+22] Dolev Bluvstein, Harry Levine, Giulia Semeghini, Tout T. Wang, Sepehr Ebadi, Marcin Kalinowski, Alexander Keesling, Nishad Maskara, Hannes Pichler, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin. A quantum processor based on coherent transport of entangled atom arrays. *Nature*, 604(7906):451–456, April 2022.

[BRSS19] J.L.F. Barbón, E. Rabinovici, R. Shir, and R. Sinha. On the evolution of operator complexity beyond scrambling. *Journal of High Energy Physics*, 2019(10), 10 2019.

[BSL+16] R. Barends, A. Shabani, L. Lamata, J. Kelly, A. Mezzacapo, U. Las Heras, R. Babbush, A. G. Fowler, B. Campbell, Yu Chen, Z. Chen, B. Chiaro, A. Dunsworth, E. Jeffrey, E. Lucero, et al. Digitized adiabatic quantum computing with a superconducting circuit. *Nature*, 534(7606):222–226, June 2016.

[BWM+18] Ryan Babbush, Nathan Wiebe, Jarrod McClean, James McClain, Hartmut Neven, and Garnet Kin Lic Chan. Low-depth quantum simulation of materials. *Physical Review X*, 8:011044, 3 2018.

[BZC+23] Pascal Baßler, Matthias Zipper, Christopher Cedzich, Markus Heinrich, Patrick H. Huber, Michael Johanning, and Martin Kliesch. Synthesis of and compilation with time-optimal multi-qubit gates. *Quantum*, 7:984, April 2023.

[CAB+21] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, September 2021.

[Cai21] Zhenyu Cai. Resource-efficient purification-based quantum error mitigation. arXiv:2107.07279, 2021.

[Cam19] Earl Campbell. Random compiler for fast hamiltonian simulation. *Phys. Rev. Lett.*, 123:070503, Aug 2019.

[CBB+23] Zhenyu Cai, Ryan Babbush, Simon C. Benjamin, Suguru Endo, William J. Huggins, Ying Li, Jarrod R. McClean, and Thomas E. O'Brien. Quantum error mitigation. *Reviews of Modern Physics*, 95(4):045005, December 2023.

[CD17] Steve Campbell and Sebastian Deffner. Trade-off between speed and cost in shortcuts to adiabaticity. *Phys. Rev. Lett.*, 118:100601, Mar 2017.

[Cep10] David M. Ceperley. *6. An Overview of Quantum Monte Carlo Methods*, pages 129–136. De Gruyter, Berlin, Boston, 2010.

[CGdW24] Yanlin Chen, András Gilyén, and Ronald de Wolf. A quantum speed-up for approximating the top eigenvectors of a matrix. arXiv:2405.14765, 2024.

[CKBG23] Chi-Fang Chen, Michael J. Kastoryano, Fernando G. S. L. Brandão, and András Gilyén. Quantum thermal state preparation, 2023.

[CMMS20] Chris Cade, Lana Mineh, Ashley Montanaro, and Stasja Stanisic. Strategies for solving the Fermi-Hubbard model on near-term quantum computers. *Phys. Rev. B*, 102:235122, Dec 2020.

[CMN+18] Andrew M. Childs, Dmitri Maslov, Yunseong Nam, Neil J. Ross, and Yuan Su. Toward the first quantum simulation with quantum speedup. *Proceedings of the National Academy of Sciences*, 115(38):9456–9461, 7 2018.

[CPDD23] Ieva Cepaite, Anatoli Polkovnikov, Andrew J. Daley, and Callum W. Duncan. Counterdiabatic optimized local driving. *PRX Quantum*, 4:010312, 1 2023.

[CPGW22] Toby Cubitt, David Perez-Garcia, and Michael M. Wolf. Undecidability of the spectral gap. *Forum of Mathematics, Pi*, 10, 2022.

[CPSP19] Pieter W. Claeys, Mohit Pandey, Dries Sels, and Anatoli Polkovnikov. Floquet-engineering counterdiabatic protocols in quantum many-body systems. *Phys. Rev. Lett.*, 123:090602, 8 2019.

[CR24] Joseph Cunningham and Jérémie Roland. Eigenpath traversal by poisson-distributed phase randomisation. arXiv:2406.03972, 2024.

[CSL+23] Jeremy Côté, Frédéric Sauvage, Martín Larocca, Matías Jonsson, Lukasz Cincio, and Tameem Albash. Diabatic quantum annealing for the frustrated ring model. *Quantum Science and Technology*, 8(4):045033, oct 2023.

[CST+21a] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe, and Shuchen Zhu. Theory of trotter error with commutator scaling. *Phys. Rev. X*, 11:011020, Feb 2021.

[CST+21b] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe, and Shuchen Zhu. Theory of Trotter error with commutator scaling. *Phys. Rev. X*, 11:011020, Feb 2021.

[CTV17] Earl T. Campbell, Barbara M. Terhal, and Christophe Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549(7671):172–179, September 2017.

[Cub23] Toby S. Cubitt. Dissipative ground state preparation and the dissipative quantum eigensolver. arXiv:2303.11962, 2023.

[CW12] Andrew M. Childs and Nathan Wiebe. Hamiltonian simulation using linear combinations of unitary operations. *Quantum Info. Comput.*, 12(11–12):901–924, 11 2012.

[Dam14] Bogdan Damski. Counterdiabatic driving of the quantum Ising model. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(12):P12019, 12 2014.

[dCRZ12] Adolfo del Campo, Marek M. Rams, and Wojciech H. Zurek. Assisted finite-rate adiabatic passage across a quantum critical point: Exact solution for the quantum Ising model. *Phys. Rev. Lett.*, 109:115703, 9 2012.

[DK07] Yu. N. Demkov and P. B. Kurasov. Von Neumann-Wigner theorem: Level repulsion and degenerate eigenvalues. *Theoretical and Mathematical Physics*, 153(1):1407–1422, Oct 2007.

[DN05] Christopher M. Dawson and Michael A. Nielsen. The Solovay-Kitaev algorithm, 2005.

[DR03] Mustafa Demirplak and Stuart A. Rice. Adiabatic population transfer with control fields. *The Journal of Physical Chemistry A*, 107(46):9937–9945, 2003.

[DR05] Mustafa Demirplak and Stuart A. Rice. Assisted adiabatic passage revisited. *The Journal of Physical Chemistry B*, 109(14):6838–6844, 2005. PMID: 16851769.

[DR08] Mustafa Demirplak and Stuart A. Rice. On the consistency, extremal, and global properties of counterdiabatic fields. *The Journal of Chemical Physics*, 129(15):154111, 10 2008.

[dW23] Ronald de Wolf. Quantum computing: lecture notes. arXiv:1907.09415, 2023.

[EBK+23] Simon J. Evered, Dolev Bluvstein, Marcin Kalinowski, Sepehr Ebadi, Tom Manovitz, Hengyun Zhou, Sophie H. Li, Alexandra A. Geim, Tout T. Wang, Nishad Maskara, Harry Levine, Giulia Semeghini, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin. High-fidelity parallel entangling gates on a neutral-atom quantum computer. *Nature*, 622(7982):268–272, October 2023.

[ECP10] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. *Reviews of Modern Physics*, 82(1):277–306, February 2010.

[EH12] Alexander Elgart and George A. Hagedorn. A note on the switching adiabatic theorem. *Journal of Mathematical Physics*, 53(10):102202, 9 2012.

[EKC+22] S. Ebadi, A. Keesling, M. Cain, T. T. Wang, H. Levine, D. Bluvstein, G. Semeghini, A. Omran, J.-G. Liu, R. Samajdar, X.-Z. Luo, B. Nash, X. Gao, B. Barak, E. Farhi, et al. Quantum optimization of maximum independent set using Rydberg atom arrays. *Science*, 376(6598):1209–1215, jun 2022.

[Fey82] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21:467–488, 6 1982.

[FGGS00] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. Quantum computation by adiabatic evolution. *arXiv: Quantum Physics*, 1 2000.

[FZZ+22] Akhil Francis, Ephrata Zelleke, Ziyue Zhang, Alexander F. Kemper, and James K. Freericks. Determining ground-state phase diagrams on quantum computers via a generalized application of adiabatic state preparation. *Symmetry*, 14:809, 4 2022.

[Gau28] J. A. Gaunt. A theory of Hartree's atomic fields. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(2):328–342, 1928.

[GB17] Christian Gross and Immanuel Bloch. Quantum simulations with ultracold atoms in optical lattices. *Science*, 357(6355):995–1001, 2017.

[GMC16] Yimin Ge, András Molnár, and J. Ignacio Cirac. Rapid adiabatic preparation of injective projected entangled pair states and gibbs states. *Physical Review Letters*, 116(8), February 2016.

[GORK+19] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martinez-Garaot, and J. G. Muga. Shortcuts to adiabaticity: Concepts, methods, and applications. *Reviews of Modern Physics*, 91(4), 10 2019.

[GS18] David J. Griffiths and Darrell F. Schroeter. *Introduction to Quantum Mechanics*. Cambridge University Press, 3 edition, 2018.

[GSLW19] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of*

Computing, STOC 2019, page 193–204, New York, NY, USA, 2019. Association for Computing Machinery.

[Har28] D. R. Hartree. The wave mechanics of an atom with a non-coulomb central field. part ii. some results and discussion. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(1):111–132, 1928.

[Has10] M. B. Hastings. Quasi-adiabatic continuation for disordered systems: Applications to correlations, lieb-schultz-mattis, and hall conductance. arXiv:1001.5280, 2010.

[HK64] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.

[HL22] Mingxia Huo and Ying Li. Dual-state purification for practical quantum error mitigation. *Physical Review A*, 105(2):022427, February 2022.

[HML22] Andreas Hartmann, Glen Bigan Mbeng, and Wolfgang Lechner. Polynomial scaling enhancement in the ground-state preparation of ising spin models via counterdiabatic driving. *Phys. Rev. A*, 105:022614, 2 2022.

[HMO+21] William J. Huggins, Sam McArdle, Thomas E. O'Brien, Joonho Lee, Nicholas C. Rubin, Sergio Boixo, K. Birgitta Whaley, Ryan Babbush, and Jarrod R. McClean. Virtual distillation for quantum error mitigation. *Physical Review X*, 11(4):041036, nov 2021.

[HPD+21] Narendra N. Hegade, Koushik Paul, Yongcheng Ding, Mikel Sanz, F. Albarrán-Arriagada, Enrique Solano, and Xi Chen. Shortcuts to adiabaticity in digitized adiabatic quantum computing. *Phys. Rev. Appl.*, 15:024038, Feb 2021.

[HW05] M. B. Hastings and Xiao-Gang Wen. Quasiadiabatic continuation of quantum states: The stability of topological ground-state degeneracy and emergent gauge invariance. *Phys. Rev. B*, 72:045141, Jul 2005.

[IBM24] IBM. Quantum roadmap. `https://www.ibm.com/roadmaps/quantum/`, 2024. Accessed 01 Nov 2024.

[Ion24] IonQ. How we achieved our 2024 performance target of #AQ 35. `https://ionq.com/posts/how-we-achieved-our-2024-performance-target-of-aq-35`, 2024. Accessed 01 Nov 2024.

[Jar95] C. Jarzynski. Geometric phases and anholonomy for a class of chaotic classical systems. *Phys. Rev. Lett.*, 74:1732–1735, 3 1995.

[JMM+22] David Joseph, Rafael Misoczki, Marcos Manzano, Joe Tricot, Fernando Dominguez Pinuaga, Olivier Lacombe, Stefan Leichenauer, Jack D. Hidary, Phil Venables, and Royal Hansen. Transitioning organizations to post-quantum cryptography. *Nature*, 605:237 – 243, 2022.

[Joh15] Steven G. Johnson. Saddle-point integration of C_∞ "bump" functions, 2015.

[Jor24] Stephen Jordan. The quantum algorithm zoo. `https://quantumalgorithmzoo.org/`, 2024. Accessed 01 Nov 2024.

[JRS07] Sabine Jansen, Mary-Beth Ruskai, and Ruedi Seiler. Bounds for the adiabatic approximation with applications to quantum computation. *Journal of Mathematical Physics*, 48(10), 10 2007.

[Kat50] Tosio Kato. On the adiabatic theorem of quantum mechanics. *Journal of the Physical Society of Japan*, 5(6):435–439, 1950.

[KEA+23] Youngseok Kim, Andrew Eddins, Sajant Anand, Ken Xuan Wei, Ewout van den Berg, Sami Rosenblatt, Hasan Nayfeh, Yantao Wu, Michael Zaletel, Kristan Temme, and Abhinav Kandala. Evidence for the utility of quantum computing before fault tolerance. *Nature*, 618(7965):500–505, June 2023.

[KFS24] Hyeongjin Kim, Matthew Fishman, and Dries Sels. Variational adiabatic transport of tensor networks. *PRX Quantum*, 5:020361, Jun 2024.

[KGB+20] Ian D. Kivlichan, Craig Gidney, Dominic W. Berry, Nathan Wiebe, Jarrod McClean, Wei Sun, Zhang Jiang, Nicholas Rubin, Austin Fowler, Alán Aspuru-Guzik, Hartmut Neven, and Ryan Babbush. Improved fault-tolerant quantum simulation of condensed-phase correlated electrons via trotterization. *Quantum*, 4:296, 7 2020.

[Kit96] Alexei Y. Kitaev. Quantum measurements and the abelian stabilizer problem. *Electron. Colloquium Comput. Complex.*, TR96-003, 1996.

[Kit97] A Yu Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191, dec 1997.

[KKR06] Julia Kempe, Alexei Kitaev, and Oded Regev. The complexity of the local hamiltonian problem. *SIAM Journal on Computing*, 35(5):1070–1097, 2006.

[KMP03] Henrik Koch, Alfredo Sánchez De Merás, and Thomas Bondo Pedersen. Reduced scaling in electronic structure calculations using Cholesky decompositions. *The Journal of Chemical Physics*, 118:9481, 5 2003.

[Koc21] Bálint Koczor. Exponential Error Suppression for Near-Term Quantum Devices. *Physical Review X*, 11(3):031057, September 2021.

[KP20] Manoj Kumar and Pratap Pattnaik. Post quantum cryptography(pqc) - an overview: (invited paper). In *2020 IEEE High Performance Extreme Computing Conference (HPEC)*, pages 1–9, 2020.

[KRL+23] Andrew D. King, Jack Raymond, Trevor Lanting, Richard Harris, Alex Zucca, Fabio Altomare, Andrew J. Berkley, Kelly Boothby, Sara Ejtemaee, Colin Enderud, Emile Hoskinson, Shuiyuan Huang, Eric Ladizinsky, Allison J. R. MacDonald, Gaelen Marsden, et al. Quantum critical dynamics in a 5,000-qubit programmable spin glass. *Nature*, 617(7959):61–66, May 2023.

[KRR+23] Alexander Kramida, Yuri Ralchenko, Joseph Reader, et al. Atomic spectra database – NIST standard reference database 78. `https://www.nist.gov/pml/atomic-spectra-database`, 2023. Accessed 01 Nov 2024.

[KS65] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.

[KSB19a] Mária Kieferová, Artur Scherer, and Dominic W. Berry. Simulating the dynamics of time-dependent hamiltonians with a truncated Dyson series. *Phys. Rev. A*, 99:042314, 4 2019.

[KSB19b] Mária Kieferová, Artur Scherer, and Dominic W. Berry. Simulating the dynamics of time-dependent hamiltonians with a truncated dyson series. *Phys. Rev. A*, 99:042314, Apr 2019.

[KSMP17] Michael Kolodrubetz, Dries Sels, Pankaj Mehta, and Anatoli Polkovnikov. Geometry and non-adiabatic response in quantum and classical systems. *Physics Reports*, 697:1–87, 6 2017.

[LBG+21] Joonho Lee, Dominic W. Berry, Craig Gidney, William J. Huggins, Jarrod R. McClean, Nathan Wiebe, and Ryan Babbush. Even more efficient quantum computations of chemistry through tensor hypercontraction. *PRX Quantum*, 2:030305, 9 2021.

[LC17] Guang Hao Low and Isaac L. Chuang. Optimal hamiltonian simulation by quantum signal processing. *Phys. Rev. Lett.*, 118:010501, Jan 2017.

[LC19a] Guang Hao Low and Isaac L. Chuang. Hamiltonian Simulation by Qubitization. *Quantum*, 3:163, July 2019.

[LC19b] Guang Hao Low and Isaac L. Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, 7 2019.

[LC19c] Guang Hao Low and Isaac L. Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, July 2019.

[LJY15] Zi-Xiang Li, Yi-Fan Jiang, and Hong Yao. Solving the fermion sign problem in quantum monte carlo simulations by majorana representation. *Phys. Rev. B*, 91:241117, Jun 2015.

[LL81] L.D. Landau and E.M. Lifshitz. *Quantum Mechanics: Non-Relativistic Theory*. Course of Theoretical Physics. Elsevier Science, 1981.

[Llo96] Seth Lloyd. Universal quantum simulators. *Science*, 273:1073–1078, 8 1996.

[LSR21] Jessica Lemieux, Artur Scherer, and Pooya Ronagh. Reflection-based adiabatic state preparation. *arXiv: Quantum Physics*, 11 2021.

[LTW+24] Martin Larocca, Supanut Thanasilp, Samson Wang, Kunal Sharma, Jacob Biamonte, Patrick J. Coles, Lukasz Cincio, Jarrod R. McClean, Zoë Holmes, and M. Cerezo. A review of barren plateaus in variational quantum computing, 2024.

[LW19] Guang Hao Low and Nathan Wiebe. Hamiltonian simulation in the interaction picture. *arXiv: Quantum Physics*, 5 2019.

[MAAV14] Michaël Mariën, Koenraad M. R. Audenaert, Karel Van Acoleyen, and Frank Verstraete. Entanglement rates and the stability of the area law for the entanglement entropy. arXiv:1411.0680, 2014.

[Mes62] Albert Messiah. *Quantum mechanics: Volume II*. North-Holland Publishing Company Amsterdam, 1962.

[MKL24] Conor Mc Keever and Michael Lubasch. Towards adiabatic quantum computing using compressed quantum circuits. *PRX Quantum*, 5:020362, Jun 2024.

[ML22] Glen Bigan Mbeng and Wolfgang Lechner. Rotated ansatz for approximate counterdiabatic driving. arXiv:2207.03553, 2022.

[MP34] Chr. Møller and M. S. Plesset. Note on an approximation treatment for many-electron systems. *Phys. Rev.*, 46:618–622, Oct 1934.

[MP24] Stewart Morawetz and Anatoli Polkovnikov. Efficient Paths for Local Counterdiabatic Driving. arXiv:2401.12287, 2024.

[MRQ+21] Xiao Mi, Pedram Roushan, Chris Quintana, Salvatore Mandrà, Jeffrey Marshall, Charles Neill, Frank Arute, Kunal Arya, Juan Atalaya, Ryan Babbush, Joseph C. Bardin, Rami Barends, Joao Basso, Andreas Bengtsson, Sergio Boixo, et al. Information scrambling in quantum circuits. *Science*, 374(6574):1479–1483, 2021.

[MS20] Ashley Montanaro and Stasja Stanisic. Compressed variational quantum eigensolver for the Fermi-Hubbard model. *arXiv: Quantum Physics*, 6 2020.

[MST+19] Mario Motta, Chong Sun, Adrian T. K. Tan, Matthew J. O'Rourke, Erika Ye, Austin J. Minnich, Fernando G. S. L. Brandão, and Garnet Kin-Lic Chan. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nature Physics*, 16(2):205–210, November 2019.

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[NGL21] Tommaso Nottoli, Jürgen Gauss, and Filippo Lipparini. Second-order CASSCF algorithm with the Cholesky decomposition of the two-electron integrals. *Journal of Chemical Theory and Computation*, 17:6819–6831, 11 2021.

[NMRMA+24] Pratik Nandy, Apollonas S. Matsoukas-Roubeas, Pablo Martinez-Azcona, Anatoly Dymarsky, and Adolfo del Campo. Quantum dynamics in Krylov space: methods and applications. arXiv:2405.09628, 2024.

[OAG+23] T. E. O'Brien, G. Anselmetti, F. Gkritsis, V. E. Elfving, S. Polla, W. J. Huggins, O. Oumarou, K. Kechedzhi, D. Abanin,

R. Acharya, I. Aleiner, R. Allen, T. I. Andersen, K. Anderson, M. Ansmann, et al. Purification-based quantum error mitigation of pair-correlated electron simulations. *Nature Physics*, 19(12):1787–1792, October 2023.

[OIWF22] Bryan O'Gorman, Sandy Irani, James Whitfield, and Bill Fefferman. Intractability of electronic structure in a fixed basis. *PRX Quantum*, 3:020322, May 2022.

[OPR+21] Thomas E. O'Brien, Stefano Polla, Nicholas C. Rubin, William J. Huggins, Sam McArdle, Sergio Boixo, Jarrod R. McClean, and Ryan Babbush. Error Mitigation via Verified Phase Estimation. *PRX Quantum*, 2(2):020317, May 2021.

[PAO23] Stefano Polla, Gian-Luca R. Anselmetti, and Thomas E. O'Brien. Optimizing the information extracted by a single qubit measurement. *Physical Review A*, 108(1):012403, July 2023.

[PCA+19] Daniel E. Parker, Xiangyu Cao, Alexander Avdoshkin, Thomas Scaffidi, and Ehud Altman. A universal operator growth hypothesis. *Phys. Rev. X*, 9:041017, 10 2019.

[PCC+20] Mohit Pandey, Pieter W. Claeys, David K. Campbell, Anatoli Polkovnikov, and Dries Sels. Adiabatic eigenstate deformations as a sensitive probe for quantum chaos. *Phys. Rev. X*, 10:041017, 10 2020.

[PCFL20] G. Passarelli, V. Cataudella, R. Fazio, and P. Lucignano. Counterdiabatic driving in the quantum annealing of the p-spin model: A variational approach. *Phys. Rev. Res.*, 2:013283, 3 2020.

[PM24] Gaopei Pan and Zi Yang Meng. *The sign problem in quantum Monte Carlo simulations*, page 879–893. Elsevier, 2024.

[Pog13] Richard Pogge. Representative visible-light emission-line spectra of various chemical elements. `https://www.astronomy.ohio-state.edu/pogge.1/TeachRes/HandSpec/atoms.html`, 2013. Accessed 01 Nov 2024.

[Pol24] Stefano Polla. `StefanoPolla/Adiabatic-Mitigation: v1.0` repository. Zenodo, `https://zenodo.org/records/10581604`, January 2024.

[qt24] Xanadu quantum technologies. Pennylane `qml` documentation. `https://docs.pennylane.ai/en/stable/code/api/pennylane.Qubitization.html`, 2024.

[QuE24] QuEra. Error-corrected computers. `https://www.quera.com/qec`, 2024. Accessed 01 Nov 2024.

[RC02] Jeremie Roland and Nicolas J. Cerf. Quantum Search by Local Adiabatic Evolution. *Physical Review A*, 65(4):042308, March 2002.

[Rei04] Ben W. Reichardt. The quantum adiabatic optimization algorithm and local minima. In *Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '04, page 502–510, New York, NY, USA, 2004. Association for Computing Machinery.

[RJ08] I. Røeggen and Tor Johansen. Cholesky decomposition of the two-electron integral matrix in electronic structure calculations. *The Journal of Chemical Physics*, 128:194107, 5 2008.

[RLB22] Nicholas C. Rubin, Joonho Lee, and Ryan Babbush. Compressing many-body fermion operators under unitary constraints. *Journal of Chemical Theory and Computation*, 18:1480–1488, 3 2022.

[RPL10] A. T. Rezakhani, A. K. Pimachev, and D. A. Lidar. Accuracy versus run time in an adiabatic quantum search. *Physical Review A*, 82(5):052305, November 2010.

[RSGSS21] E. Rabinovici, A. Sánchez-Garrido, R. Shir, and J. Sonner. Operator complexity: a journey to the edge of Krylov space. *Journal of High Energy Physics*, 2021(6), 6 2021.

[RSGSS22] E. Rabinovici, A. Sánchez-Garrido, R. Shir, and J. Sonner. Krylov localization and suppression of complexity. *Journal of High Energy Physics*, 2022(3), 3 2022.

[Sac99] Subir Sachdev. Quantum phase transitions. *Physics World*, 12(4):33, April 1999.

[SB23] Paul Manuel Schindler and Marin Bukov. Counterdiabatic Driving for Periodically Driven Systems. arXiv:2310.02728, 2023.

[SBG+13] J. Stoer, R. Bartels, W. Gautschi, R. Bulirsch, and C. Witzgall. *Introduction to Numerical Analysis*. Texts in Applied Mathematics. Springer New York, 2013.

[SBM06] V.V. Shende, S.S. Bullock, and I.L. Markov. Synthesis of quantum-logic circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 25(6):1000–1010, 2006.

[Sla28] J. C. Slater. The self-consistent field and the structure of atoms. *Phys. Rev.*, 32:339–348, Sep 1928.

[SNKM24] Yuta Shingu, Tetsuro Nikuni, Shiro Kawabata, and Yuichiro Matsuzaki. Quantum annealing with error mitigation. *Physical Review A*, 109(4):042606, April 2024.

[SOMdC14] Hamed Saberi, Tomáš Opatrný, Klaus Mølmer, and Adolfo del Campo. Adiabatic tracking of quantum many-body dynamics. *Phys. Rev. A*, 90:060301, 12 2014.

[SP17] Dries Sels and Anatoli Polkovnikov. Minimizing irreversible losses in quantum systems by local counterdiabatic driving. *Proceedings of the National Academy of Sciences*, 114(20), 5 2017.

[SRL12] Jacob T. Seeley, Martin J. Richard, and Peter J. Love. The Bravyi-Kitaev transformation for quantum computation of electronic structure. *The Journal of Chemical Physics*, 137(22), December 2012.

[SSW+21] Pascal Scholl, Michael Schuler, Hannah J. Williams, Alexander A. Eberharter, Daniel Barredo, Kai-Niklas Schymik, Vincent Lienhard, Louis-Paul Henry, Thomas C. Lang, Thierry Lahaye, Andreas M. Läuchli, and Antoine Browaeys. Quantum simulation of 2D antiferromagnets with hundreds of Rydberg atoms. *Nature*, 595(7866):233–238, July 2021.

[STC22] Benjamin F. Schiffer, Jordi Tura, and J. Ignacio Cirac. Adiabatic spectroscopy and a variational quantum adiabatic algorithm. *PRX Quantum*, 3(2):020347, jun 2022.

[Suz92] Masuo Suzuki. General theory of higher-order decomposition of exponential operators and symplectic integrators. *Physics Letters A*, 165(5):387–395, 1992.

[SvVTP24] Benjamin F. Schiffer, Dyon van Vreumingen, Jordi Tura, and Stefano Polla. Virtual mitigation of coherent non-adiabatic transitions by echo verification. *Quantum*, 8:1346, May 2024.

[Tak13] Kazutaka Takahashi. Transitionless quantum driving for spin systems. *Phys. Rev. E*, 87:062117, 6 2013.

[TCC+22] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H. Booth, and Jonathan Tennyson. The variational quantum eigensolver: A review of methods and best practices. *Physics Reports*, 986:1–128, 11 2022.

[TdC24] Kazutaka Takahashi and Adolfo del Campo. Shortcuts to adiabaticity in Krylov space. *Phys. Rev. X*, 14:011032, 2 2024.

[TEM+89] A. Tonomura, J. Endo, T. Matsuda, T. Kawasaki, and H. Ezawa. Demonstration of single-electron buildup of an interference pattern. *American Journal of Physics*, 57(2):117–120, 02 1989.

[TSCT21] Minh C. Tran, Yuan Su, Daniel Carney, and Jacob M. Taylor. Faster Digital Quantum Simulation by Symmetry Protection. *PRX Quantum*, 2(1):010323, February 2021.

[TSRP16] Michael Tomka, Tiago Souza, Steven Rosenberg, and Anatoli Polkovnikov. Geodesic paths for quantum many-body systems. *arXiv: Quantum Gases*, 6 2016.

[TVS17] Felix Tennie, Vlatko Vedral, and Christian Schilling. Universal upper bounds on the Bose-Einstein condensate and the Hubbard star. *Physical Review B*, 96, 7 2017.

[vBLH+21] Vera von Burg, Guang Hao Low, Thomas Häner, Damian S. Steiger, Markus Reiher, Martin Roetteler, and Matthias Troyer. Quantum computing enhanced computational catalysis. *Physical Review Research*, 3:033055, 9 2021.

[vV24] Dyon van Vreumingen. Gate-based counterdiabatic driving with complexity guarantees. *Phys. Rev. A*, 110:052419, Nov 2024.

[vVS23] Dyon van Vreumingen and Kareljan Schoutens. Adiabatic ground-state preparation of fermionic many-body systems from a two-body perspective. *Phys. Rev. A*, 108:062603, Dec 2023.

[WB12] Nathan Wiebe and Nathan S. Babcock. Improved error-scaling for adiabatic quantum evolutions. *New Journal of Physics*, 14(1):013024, January 2012.

[WBAG11] James D. Whitfield, Jacob Biamonte, and Alan Aspuru-Guzik. Simulation of electronic structure hamiltonians using quantum computers. *Molecular Physics*, 109:735–750, 3 2011.

[WBHS10] Nathan Wiebe, Dominic Berry, Peter Høyer, and Barry C Sanders. Higher order decompositions of ordered operator exponentials. *Journal of Physics A: Mathematical and Theoretical*, 43(6):065203, 1 2010.

[WK22] Kianna Wan and Isaac H Kim. Fast digital methods for adiabatic state preparation. *arXiv: Quantum Physics*, 4 2022.

[WLL20] Shijie Wei, Hang Li, and GuiLu Long. A full quantum eigensolver for quantum chemistry simulations. *Research*, 2020, 1 2020.

[WWRL22] Jacob Watkins, Nathan Wiebe, Alessandro Roggero, and Dean Lee. Time Dependent Hamiltonian Simulation Using Discrete Clock Constructions. arXiv:2203.11353, 2022.

[WZ05] Congjun Wu and Shou-Cheng Zhang. Sufficient condition for absence of the sign problem in the fermionic quantum monte carlo algorithm. *Phys. Rev. B*, 71:155115, Apr 2005.

[XSY22] Qing Xie, Kazuhiro Seki, and Seiji Yunoki. Variational counterdiabatic driving of the hubbard model for ground-state preparation. *Phys. Rev. B*, 106:155153, 10 2022.

[YCCV+23] Yilun Yang, Arthur Christianen, Sandra Coll-Vinent, Vadim Smelyanskiy, Mari Carmen Bañuls, Thomas E. O'Brien, Dominik S. Wild, and J. Ignacio Cirac. Simulating prethermalization using near-term quantum computers. *PRX Quantum*, 4(3):030320, August 2023.

[YSBK13] Kevin C. Young, Mohan Sarovar, and Robin Blume-Kohout. Error Suppression and Error Correction in Adiabatic Quantum Computation: Techniques and Challenges. *Physical Review X*, 3(4):041013, November 2013.

[ZCDCP16] Yuanjian Zheng, Steve Campbell, Gabriele De Chiara, and Dario Poletti. Cost of counterdiabatic driving and work output. *Phys. Rev. A*, 94:042132, Oct 2016.

[ZJN+20] Hui Zhou, Yunlan Ji, Xinfang Nie, Xiaodong Yang, Xi Chen, Ji Bian, and Xinhua Peng. Experimental realization of shortcuts to adiabaticity in a nonintegrable spin chain by local counterdiabatic driving. *Phys. Rev. Appl.*, 13:044059, 4 2020.

Acknowledgments

First and foremost, I would like to thank my supervisors Kareljan and Luuk for providing me with the opportunity to pursue a doctoral degree here at the UvA, QuSoft and VU in Amsterdam. It has been a journey filled with challenges – which, as my father would say, I dearly needed – but also with moments of fulfillment and personal growth. I am grateful for the acquired skills in conducting scientific research, and I am certain they will be valuable throughout the rest of my life.

I am especially grateful to Kareljan for the active supervision throughout these four years. There are supervisors who make only little time for their students; you were the opposite. You ensured I stayed on track through our regular meetings, and you were always there when I needed something. Thank you for the thorough guidance and your contributions to my first paper.

I also want to extend my gratefulness to Luuk for the time spent on getting me started and guiding me through the earlier days of the PhD. I will remember the extensive discussions we had during our meetings, and the multitude of learning sources you always gave me. Suffice to say I learned a lot from working with you.

Furthermore, I thank Ronald who, despite not being my official supervisor, was always available for discussions and advice. You gave me useful pointers that improved the quality of the work in this thesis.

Additional thanks go to the members of my doctoral comittee, Amira, Anatoli, Jasper, Jordi and Ronald for their time as well as their willingness to subject my thesis to scrutiny and their commitment to the role of opposition.

Thanks to my coauthors Stefano, Ben and Jordi for the work on echo verification we did together. I enjoyed the whiteboard discussions, solving the problems and the collaboration overall.

Thank you also to Stefano, Jan and Emiel for the insightful discussions we had. Stefano, I was always impressed by your extensive knowledge and fascination with science, and I am certain they will be valued by many others in your future career. Jan, I had a lot of fun during our discussions in Aveiro, Innsbruck, Nijmegen,

Amsterdam and Glasgow. I wish you fruitful collaborations in the future and hope I got you interested in sabre fencing. Emiel, thank you for the ideas you put forward for my first project during the early days of my PhD.

To my paranymphs Fran and Mariia, thank you for your supportive role around and during the defence.

Of course, these acknowledgments would not be complete without mentioning the people at QuSoft, the VU quantum chemistry group and CWI. Thanks for all the nice conversations, jokes, foosball matches, trips, parties, game nights, padel sessions and fun moments in general. You all made the journey much more enjoyable and will certainly be missed.

Poojith, thank you for our friendship, your neverending hospitality and for introducing me to the fascinating culture of South India. I learnt a lot from you and you inspired me to explore the world beyond the Western borders. I hope I'll be able to maintain this fascination for as long as I walk around on this planet.

Léo, thank you as well for being friends, for joining me in the addiction to this silly foosball game and for the invitiation to your wonderful wedding. You managed to convince me that France consists of more than merely a bunch of chauvinists who refuse to speak English. (For one, the French are much better dancers than we are.) I was met with a warmth and hospitality that are sometimes not as easy to find here. Just know that our cheese is still the best in the world – nothing beats Gouda cheese. Also without mustard.

Adam, Llorenç, Rene and Philip, I'm grateful for the beautiful trip we made to Hawaiï after QIP 2022. Adam, thanks for the deep conversations about life and Herman Hesse, and please never stop doing handstands in random places. Llorenç, thank you for being there for me in my weakest moment, when... well, you know the story. What happened on Hawaiï, stays on Hawaiï, I suppose.

Likewise, I owe thanks to Fran, Adam and Davi for the wonderful winter escape to Tenerife. Fran, while we will probably never agree on political matters, you will forever be my champion. Davi, bastard, I really hope I didn't traumatise you for life with our primitive fencing session on the beach.

Quinten, it was nice to have someone around with a business-oriented mindset, and it was always refreshing to talk with you. Subha and Mani, thanks for the potlucks; I really enjoyed making and having food together with you.

The list of people at QuSoft to whom I'm indebted is too long to thank everyone explicitly, but these acknowledgments would be imcomplete without mentioning all of you. Thanks to Ailsa, Ake, Akshay, Álvaro, Amira, Anna, Arie, Arjan (did anyone say waffle machine?), Bob, Chris, Crownie, Daan, Dmitry, Farrokh, Filippo, Francesco, Freek, Galina, Garazi, Gina, Harold, Ido, Jana, Jelena, Joran, Jordi, Koen, Lorenzo, Luca, Lynn (having been paranymphs together implies we're married now, you say?), Marten, Maxim, Mert, Mehrdad, Niels, Nikhil, Randy (I didn't expect to ever play a board game with a 45-page rulebook; it was great), Salvatore, Sarah, Sebastian Zur, Sebastian Verschoor, Seenivasan, Yanlin, Yaroslav and Yfke as well as Christian, Florian, Harry, Jeroen,

John, Jonas, Jop, Ludovico, Maris, Peter and Stacey for making QuSoft the great environment it has been and still is today. And naturally, a big thank you to Suzanne and Doutzen for their hard work in keeping QuSoft running and organising the scientific retreats, which were a big highlight every year.

From the quantum chemistry group at the VU, I would like to thank Mariia, Emiel, Souloke, Sharath and Edison. It was always fun to banter with you after the QC^2 meetings and after work time. Mariia, thank you in particular for your help with language learning.

Thanks also to Simona and Max at CWI for all the foosball games and the fun moments together.

Of course, I am also grateful to all the people outside my PhD environment who brightened my days. First of all, thank you to all the people I met through the Jong UvA network for the fun activities we did together. Out of all those, I want to give special thanks to Maddalena for convincing me to join the Jong UvA board and your kindness in general, as well as my fellow board members Taco, Clemens and Lina for the fun year. Honestly, I never thought a bottle of tomato juice could ever trigger such strong emotions... Anyway, it was great to throw money out of the window with you and we should totally twerk off the wall some time!

Another big thank you goes to *Ludendi* for the weekly game sessions, weekend getaways and the endless stream of stupid jokes that are probably best kept within the group. I'm really grateful to have you as a long-lasting group of friends and am looking forward to many more amazing years. To all of you, je moeder! Tom, thank you in addition for the beautiful trips to Peru (masaje? masaje?), Vienna (Baustelle?), Istanbul (eylül, kürk?) and southern France (jij agressieve badeend!).

To the friends from *Benidorm*, thanks for all the fun dinners, drinks and holidays and the friendship in general. Thanks in particular to Manon for making the effort to organise the wonderful nights at your home.

Thanks to my flat mates Henrik and Danish for the fun conversations about language, politics and ballroom dancing.

Dear Imke, thank you for your inspiration. When working on proving theorems all day, it can be easy to forget the power and importance of fascination for the world and the presence of all those "infinitely interesting people" close by.

Lastly, to my brother Rocco: thanks for the unwavering camaraderie and all the silly jokes; and to my parents Bert and Edit: thanks for being there for me at any time, for your unconditional support and for all the things you've done for me. This thesis would not have been possible without you.

شاد زی با سیاه چشمان شاد
که جهان نیست جز فسانه و باد
ز آمده تنگدل نباید بود
و ز گزشته نکرد باید یاد