Strategies for braiding and ground state preparation in digital quantum hardware

Herasymenko, Y.

Strategies for braiding and ground state preparation in digital quantum hardware

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.dr.ir. H. Bijl,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 20 april 2022
klokke 15.00 uur

door

Yaroslav Romanovych Herasymenko

geboren te Kyiv (Oekraïne)
in 1995

Promotor:	prof.dr. C.W.J. Beenakker
Co-promotor:	dr. T.E. O'Brien (Google Quantum AI, Munich, Germany)

Promotiecommissie:	prof.dr. Y. Gefen (Weizmann Institute, Israel)
	prof.dr. B.M. Terhal (Technische Universiteit Delft)
	prof.dr. J. Aarts
	dr. W. Löffler
	prof.dr. J. Zaanen

Casimir PhD series Delft-Leiden 2022-10
ISBN 978-90-8593-521-6

This work originates as part of the research programme of the Foundation for Fundamental Research on Matter (FOM), and falls as of April 1, 2017 under the responsibility of Foundation for Nederlandse Wetenschappelijk Onderzoek Instituten (NWO-I), which is part of the Dutch Research Council (NWO).

Cover: An interlacing of colored threads which illustrates braiding and the notion of a quantum ground state. Regions with regular patterns represent the onset of order at low energies. The orange-to-blue gradient on the threads shows the transition from hot to cold as one moves closer to the ground state.

To my family and to Marina.

Contents

1 Introduction

1.1 Preface

A quantum computer is a special kind of computer that perform its tasks by leveraging the laws of quantum mechanics. For particular classes of problems, such devices are expected to dramatically outperform conventional (classical) computers [1], such as modern-day laptops or even supercomputers. The greatest possibilities are offered by *digital* quantum computers, which possess the flexibility to perform arbitrary quantum computation. The word 'digital' implies that such a computer operates on standardized memory registers and digitizes the desired computation into elementary operations called gates*. The ability of a digital quantum computer to perform an arbitrary computation is known as universality. Although highly promising due to their universality, application-ready digital quantum computers are hard to implement. In particular, most such realizations need an exceptional degree of precision and noise isolation [2]. To this date, significant questions remain open: for instance, how to reliably implement a digital quantum computer? Furthermore, once such a computer is available, how to utilize it effectively?

One possible way to realize digital quantum computation is by employing a technique called braiding [3–5]. The key elements in such an implementation are anyons — particle-like energy excitations supported by some quantum materials. In a quantum computation, information is to be encoded and modified via mutual interchanges, or braiding, of anyons. The term 'braiding' refers to a similar exchange of strands in the familiar process of making a braid. Anyons and their braiding could be promising building blocks for a digital quantum computer. However, existing theoretical blueprints for anyons so far were elusive to implement [6–10]. This thesis addresses the issue by putting forward alternative proposals for anyon braiding and detection.

For a functioning quantum computer, a promising early application is

*This method can be contrasted with analog quantum computation, which does not require digitization but instead is run on a special-purpose device built to perform a given type of computation.

simulating physical systems [11–14]. One type of potential simulation is preparing a ground state — the lowest energy state of a quantum system. Information obtained from simulated ground states can be used in research or engineering since many chemicals and materials occupy low-energy configurations. On a classical computer, representing a quantum ground state is generally unfeasible [15–17], due to the many-body quantum correlations present in the state. On the other hand, a quantum computer can reproduce these correlations efficiently due to the quantum nature of its hardware. The very procedure of preparing the ground state, however, is less trivial and is currently a subject of active research [18]. In this thesis, we layout several new strategies towards achieving this goal.

1.2 Topological matter and braiding

Solid-state quantum materials have the ability to enter special *topological* phases [19, 20], characterized by elementary excitations with unusual properties. In 2D materials, the topological nature of the phase can manifest itself in so-called *anyonic* statistics of its excitations [3, 21]. In contrast to bosons or fermions, an interchange (braiding) of two *anyons* can modify the system wavefunction by a nontrivial complex phase or even a unitary operation. Perhaps more strikingly, this modification depends only on the topology of the braided worldlines of the anyons. If a species of anyons realizes non-commuting unitary operations via braiding, such anyon statistics is called non-abelian.

Non-abelian statistics is at the foundation of promised quantum computing applications of anyons. Because braiding is discrete, the computation realized by such interchange procedures is a digital one. For some species of anyons, the set of these quantum operations is even universal – they allow to access the Jones polynomials of the knot theory, which are proven to efficiently encode an arbitrary quantum computation [22]. Finally, due to its topological nature, anyon braiding is expected to be robust to external noise and perturbations.

The above properties make topological matter a viable platform for quantum computation. However, the existing proposals for such an application have not yet been realized in a lab. It is therefore of special interest to propose alternative realization platforms for non-abelian statistics. For the same reason, there is an active interest in the novel experimental signatures of such anyonic excitations. We focus on these questions as posed for two types of topological systems: topological superconductors and Fractional Quantum Hall materials.

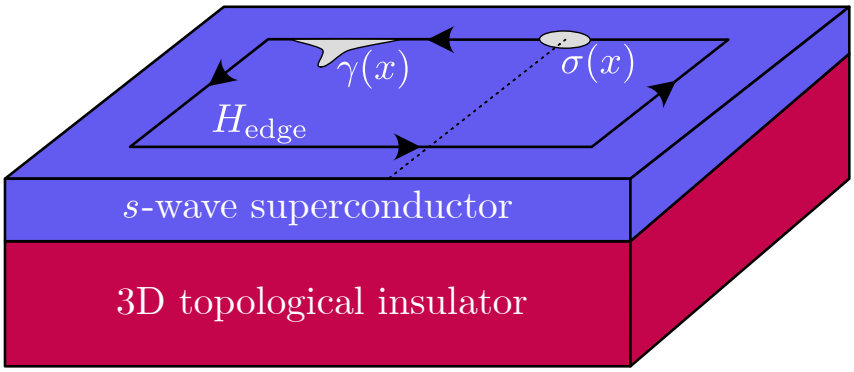

Figure 1.1: Topological superconductivity. A topological superconductor can be effectively realized in a superconductor-topological insulator heterostructure [25]. It is characterized by the presence of a gapless edge mode whose physics is governed by the Hamiltonian H_{edge} (1.2.2). The elementary excitations of the edge mode are Majorana fermions $\gamma(x)$ and edge vortices $\sigma(x)$. (the dotted line shows the associated branch cut)

1.2.1 Topological superconductivity

A good example of a topological superconductor in 2 dimensions is a p-wave superconductor [3, 21, 23, 24]. As a model, one may consider the following ('$p+ip$') Bogoliubov-de Gennes Hamiltonian:

$$H_{p+ip} = \sum_p \left[(\frac{p^2}{2m} - \mu)(c_p^\dagger c_p - b_p^\dagger b_p) + (p_x + ip_y)c_p^\dagger b_p + (p_x - ip_y)b_p^\dagger c_p \right], \tag{1.2.1}$$

where c (b) operators describe the electron (hole) degrees of freedom*. For $\mu > 0$, the Hamiltonian (1.2.1) enters a topological phase, while the phase at $\mu < 0$ is referred to as trivial. Both of these phases are characterized by an energy gap for the bulk excitations. But unlike in the trivial phase, in the topological phase a finite sample of superconductor (1.2.1) would host an additional gapless mode at its edge (Fig. 1.1). This edge mode is chiral, being effectively described by the Hamiltonian H_{edge}:

$$H_{\text{edge}} = \int_x i\gamma(x)\partial_x\gamma(x)dx, \tag{1.2.2}$$

*The absence of a spin degree of freedom in this toy model is justified for a p-wave superconductor, which allows for a superconducting pairing in the spin-polarized channel.

where $\gamma(x)$ is a Majorana fermion: an operator satisfying $\gamma^\dagger(x) = \gamma(x)$ that describes a neutral quasiparticle i.e. one that does not carry charge.

Experimentally, fabricating a topological p-wave superconductor material (1.2.1) turns out to be extremely difficult [26]. Instead, a two-dimensional topological superconductor can effectively be realized [25] with a layer of an s-wave superconductor on a topological insulator substrate (Fig. 1.1). Such a heterostructure has a similar effective Hamiltonian to the $p + ip$ superconductor* (1.2.1) and features the gapless edge mode (1.2.2). Recent experimental works have attempted this realization of topological superconductivity [27, 28], and observed a signature consistent with the Majorana edge mode (1.2.2). However, alternative explanations of the observed data have also been proposed [9, 10, 29], and to date no consensus has been reached on whether the observed signal is from a Majorana mode. Optimistically, in the near-future one expects the body of such evidence to grow further [7], and two-dimensional topological superconductivity to be finally established and harnessed in a lab setting.

1.2.2 Majorana anyons: bulk and boundary

In addition to the Majorana edge mode, topological superconductors also provide a platform for anyonic braiding statistics. The anyonic excitation that has this statistics is an Abrikosov vortex of supercurrent [30], which in a topological superconductor hosts a zero-energy Majorana bound state (or zero-mode). While playing a crucial role in the vortex exchange properties, Majorana zero modes γ_α themselves obey statistics of a fermionic type:

$$\gamma_\alpha \gamma_\beta + \gamma_\beta \gamma_\alpha = 2\delta_{\alpha,\beta}. \tag{1.2.3}$$

The anyonic statistics of the vortices is rooted in the Aharonov-Bohm [31] effect imposed by the vortex. In particular, a 2π phase winding of the superconducting order parameter Δ can be translated into a π phase shift for the fermionic variables. This can be represented by a branch-cut boundary condition for the fermions (Fig. 1.2a), with implications for the vortex statistics [32]. Consider the example sketched in Fig. 1.2, where two vortices σ_1 and σ_2 are interchanged. Due to the inevitable crossing of a

*Although unlike H_{p+ip}, it satisfies the time-reversal symmetry.

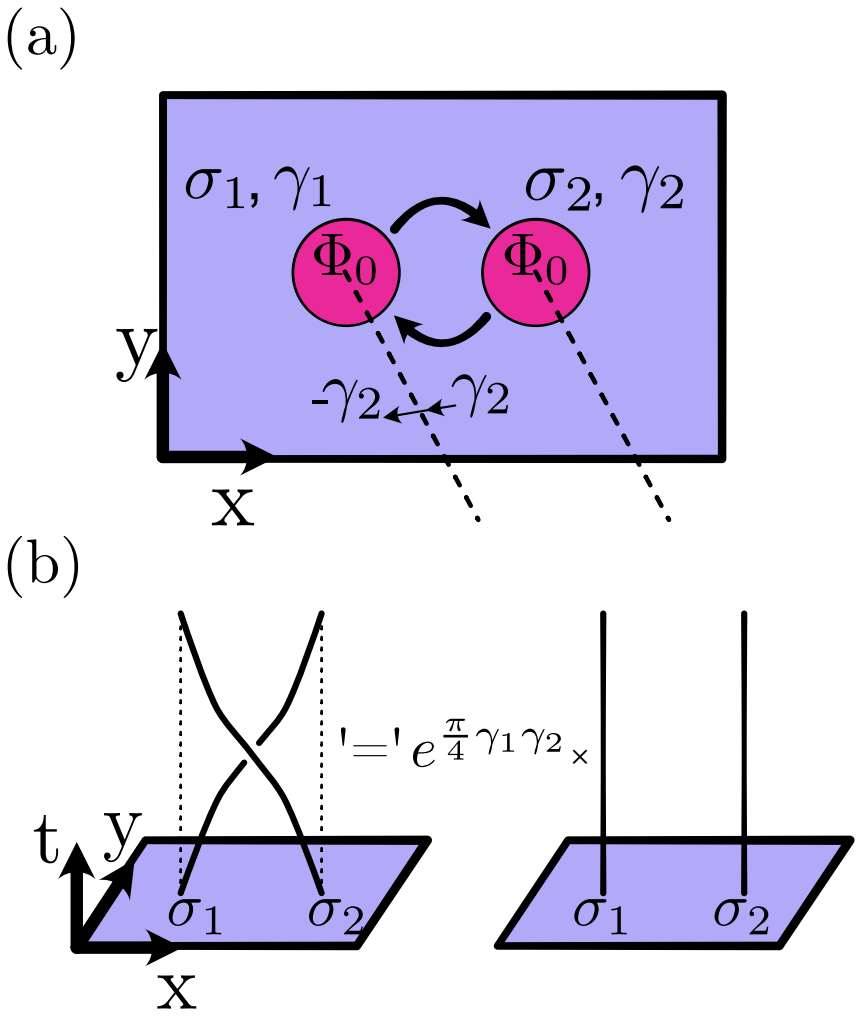

Figure 1.2: Braiding of vortices in a 2D topological superconductor. (a) In a topological superconductor, Abrikosov vortices $\sigma_{1,2}$ host Majorana fermion modes $\gamma_{1,2}$ at zero energy. The magnetic flux quantum $\Phi_0 = h/2e$ carried by each vortex induces a branch-cut boundary condition (dotted lines). Any fermion operator flips the sign once crossing the branch cut. Such crossing is guaranteed to happen whenever the two vortices are exchanged. (b) In space and time, the vortex exchange process implies braiding of their worldlines (on the left). Compared to no exchange (on the right), such a braiding operation transforms the Majorana degrees of freedom $\gamma_{1,2}$ with a unitary: $U_{12} = e^{\frac{\pi}{4}\gamma_1\gamma_2}$.

branch cut, Majorana operator γ_2 has to flip a sign. In algebraic form, this exchange operation produces a unitary U_{12} such that $U_{12}^{-1}\gamma_2 U_{12} = -\gamma_1$, $U_{12}^{-1}\gamma_1 U_{12} = \gamma_2$. By virtue of the Majorana operator algebra (1.2.3), these relations imply the following form of the unitary:

$$U_{12} = e^{\frac{\pi}{4}\gamma_1\gamma_2} = \frac{1}{\sqrt{2}}(\mathbf{I} + \gamma_1\gamma_2). \tag{1.2.4}$$

More generally, a similar exchange of vortices σ_α and σ_β would be represented by a unitary operation $U_{\alpha\beta} = e^{\frac{\pi}{4}\gamma_\alpha\gamma_\beta}$. As different $U_{\alpha\beta}$ do not generally commute, the vortices σ_α have the exchange statistics of non-abelian anyons.

An intense recent interest has been drawn to the relation between the edge modes and anyonic excitations. Since Majoranas play a role in the anyonic statistics of the bulk excitations, some have proposed to use edge Majoranas to produce the anyonic statistics [33]. Edge mode theory also permits vortex-like excitations, which can be characterized by branch cuts (illustrated in Fig. 1.1). Some of the work presented in this thesis is motivated by the question: can one reproduce the bulk anyonic statistics using edge vortices?

1.2.3 Fractional Quantum Hall Effect. Laughlin quasiparticles

In some two-dimensional materials subjected to strong magnetic fields at low temperatures, the Hall conductivity [34] is measured to be quantized:

$$\sigma_{xy} = q\frac{e^2}{h}, \tag{1.2.5}$$

which is known as the Quantum Hall Effect. In the Integer Quantum Hall Effect (IQHE) [35], the constant q in (1.2.5) is always an integer, while in the Fractional Quantum Hall Effect (FQHE) [36] it can also take fractional values. Similarly to topological superconductivity, the key to Quantum Hall physics is in the gapless edge modes [37], which in this case carry the quantized Hall currents. In the IQHE such gapless modes are populated by ordinary electrons, and the quantization (1.2.5) with an integer q follows from the conductance quantization theorem for 1D channels [38]. Meanwhile, the non-integer q of the FQHE materials seemingly contradicts this theorem. This striking effect is explained by the presence of many-body excitations with fractional charge, rather than electrons which carry unit charge, in the gapless edge mode. In the Laughlin model [39], which

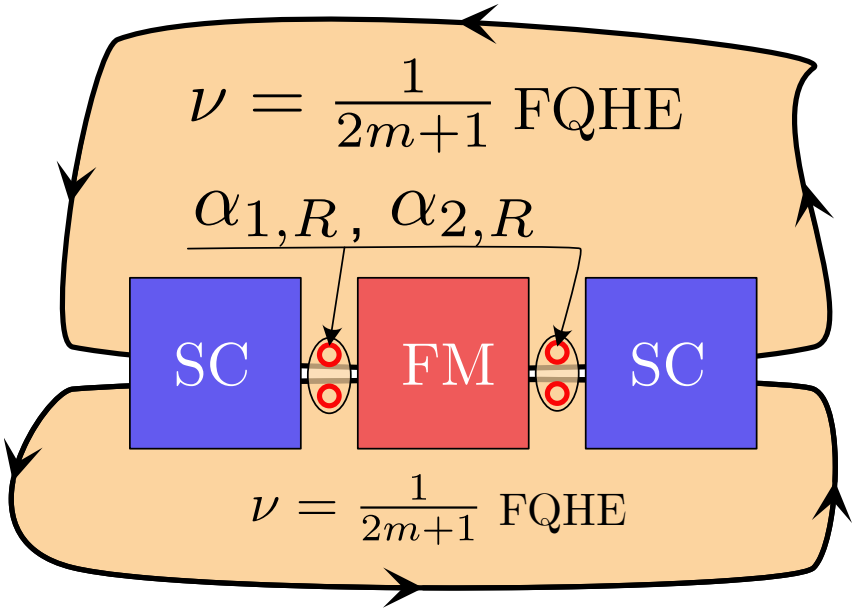

Figure 1.3: Combining two samples of Laughlin FQHE materials allows one to enact parafermionic zero-modes. For that, one has to deposit a sequence of ferromagnets (FM) and superconductors (SC) on top of the interfacing sample edges. Each interface between a ferromagnet and a superconductor hosts two zero-modes: one per edge mode. The respective operators are characterized by parafermionic statistics.

describes the subclass of FQHEs with inverse odd $q = \frac{1}{2n+1}$, there is a single species of such fractional excitations. These Laughlin quasiparticles $\psi(x)$ carry charge qe and have anyonic exchange statistics:

$$\psi(x)\psi(y) = \psi(y)\psi(x)e^{iq\pi}, \quad x > y. \tag{1.2.6}$$

Since the exchange unitary is represented only by an overall phase $e^{iq\pi}$, the Laughlin anyons are abelian.

With material in an FQHE phase of the Laughlin type, one can also achieve non-abelian braiding statistics. For this, one needs to orchestrate so-called parafermionic zero-modes [40]. This can be realized if a pair of counterpropagating FQHE edge modes are gapped out by a sequence of ferromagnets interlaced with superconductors (Fig. 1.3). On sites $j \in \mathbf{Z}$ between the superconducting and ferromagnetic domains, a parafermionic

zero-mode $\alpha_{j,L/R}$ emerges. L and R here stand for left-propagating and right-propagating, depending on the edge where $\alpha_{j,L/R}$ is localized. The parafermion operators $\alpha_{j,L/R}$ themselves have statistics similar to Laughlin quasiparticles, e.g.

$$\alpha_{j,R}\alpha_{k,R} = \alpha_{k,R}\alpha_{j,R}e^{iq\pi}, \quad j > k, \tag{1.2.7}$$

which is abelian. However, a gradual exchange of two parafermion-hosting sites can be used to produce a nonabelian operation [40]. Such a relation between abelian and non-abelian statistics of parafermions can be closely paralleled with Majorana braiding in a topological superconductor. Indeed, Majorana fermions in Abrikosov vortices do not carry nonabelian statistics, while the braiding of those vortices is nonabelian. In contrast to Majorana-hosting vortices, however, the non-abelian braiding of parafermionic sites allows for an even larger set of unitary operations [40]. In this thesis, we will use this fact as a motivation for our investigation of parafermions but not focus on their braiding per se. Instead, we are interested in constructing new and useful ways to characterize parafermionic zero mode statistics (1.2.7) in experiment.

1.3 Preparing ground states with a quantum computer

In this section, we transition from the basic physics of digital quantum hardware and move on to its potential utilization. One of the most promising applications of quantum computers is preparing a simulation of a many-body ground state. A version of this task is likely [18] to be among the first problems that are solved on a quantum computer with a speed-up relative to its classical counterparts. The ground state preparation problem is relevant to quantum physics as well as classical physics and computer science. As a natural example, it arises when studying the low-temperature properties of many-body systems. This includes [18] multi-electron systems in solid state physics, chemistry, and spin systems. In computer science, a cost function optimization is also a common task, with application to machine learning and logistics [41]. Such tasks can also be mapped onto a ground state finding problem for a particular Hamiltonian [16].

The hope for the success of quantum computers in ground state simulation lies in the exponentially greater expressiveness of a quantum computer compared to a classical one. However, one does not expect such an exponential speed-up in the context of every problem. In particular,

NP-complete ground state search problems, e.g. for classical spin models [42], are not expected to be polynomially solvable on a quantum computer. For quantum many-body Hamiltonians, accessing the ground state even in some restricted Hamiltonian families is already QMA-complete [43] (quantum analog of NP-complete). On the other hand, any *polynomially* complex quantum computation can be mapped [44] onto the problem of distilling the exact ground state of a particular Hamiltonian from a given good approximation thereof. These mathematical results highlight the relevance of ground state preparation to general quantum speed-ups and the significance of approximating target ground states efficiently.

To prepare a ground state on a quantum computer, one needs to build and utilize an appropriate quantum algorithm. On a broad level, there already exist multiple paradigms constructing such algorithms. These include adiabatic quantum computing [45], quantum phase estimation [46], and variational quantum algorithms [47]. An ongoing effort is devoted to optimizing such approaches for practical use. This optimization is especially critical if an algorithm is to be employed in the near-term [11], before existing quantum hardware is sufficiently protected from the errors. For such applications, the exact time complexity of an algorithm is as crucial as its asymptotic scaling. In addition to optimizing existing ideas, creating original methods for ground state preparation is also of interest. A novel approach to this problem may yield a sizeable improvement in time complexity, especially in the context of specific niche applications.

1.3.1 Variational quantum algorithms

A variational quantum algorithm (Fig. 1.4) is a quantum-classical hybrid algorithm which aims to approximate the ground state of a given Hamiltonian by utilizing the variational principle [47]. Since quantum circuits are in general exponentially hard to represent classically, a variational quantum algorithm employs a quantum circuit as a powerful variational ansatz. To enable the optimization procedure, the energy of the ansatz state is being measured at the end of the circuit by means of sampling from the target Hamiltonian. The quantum circuit is then tuned to ensure the minimization of said energy. Variational quantum algorithm is well-suited for use on near-term quantum devices [47], being not highly sensitive to noise and not requiring an implementation of deep quantum circuits. The success of this method depends on the efficiency of the energy measurement [48–50], choice of optimization procedure [51–53], and the expressiveness

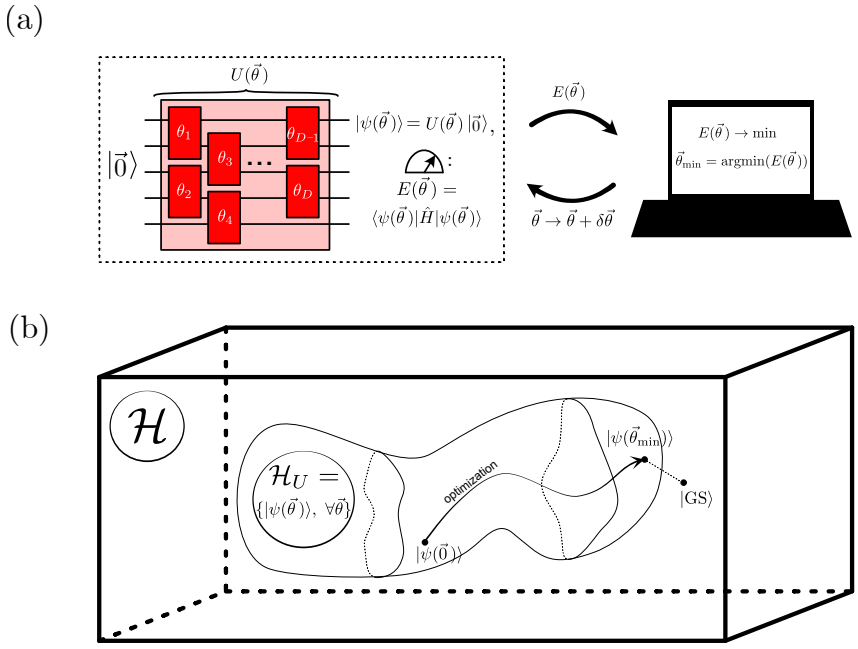

Figure 1.4: Variational quantum algorithms. (a) The hybrid quantum-classical scheme that was introduced in [47]. The algorithm employs a tunable ansatz circuit $U(\vec{\theta})$, and aims to find the value of parameters $\vec{\theta}_{\min}$ which minimizes the variational energy $E(\vec{\theta})$. Quantum hardware need not be continuously in a coherent state since the information is processed classically. This is the key advantage of such an algorithm for near-term implementation. (b) Schematic illustration of the ansatz expressivity. In the total Hilbert space of the system, $\mathcal{H}$, only the small subset $\mathcal{H}_U$ is spanned by the variational states. It is a manifold that is usually much lower in dimensionality (polynomial versus exponential), and generally, the ground state $|GS\rangle$ lies outside of it. The state $|\psi(\vec{\theta}_{\min})\rangle$ produced by the variational algorithm only approximates $|GS\rangle$. It is the task of ansatz construction that the expected separation between $|GS\rangle$ and $|\psi(\vec{\theta}_{\min})\rangle$ is ensured to be minimal. For that, $\mathcal{H}_U$ has to span the physically relevant part of the Hilbert space.

of the ansatz circuits [54–56]. Since the very potential for exponential speed-up is rooted in the capacity of the quantum circuit, such ansatz design is particularly crucial.

To be efficient, a variational ansatz circuit needs to be tailored to the problem at hand. In the quantum chemistry context, a common way to do

this [47, 54] is through the unitary coupled-cluster (UCC) method. The UCC state has the form:

$$|\psi_{\mathrm{UCC}}\rangle = U\,|\psi_{HF}\rangle = e^{T_1+T_2+T_3+\cdots}\,|\psi_{\mathrm{HF}}\rangle\,, \tag{1.3.1}$$

where $|\psi_{\mathrm{HF}}\rangle$ is a Hartree-Fock approximation to the fermionic ground state, and anti-Hermitian generators T_n consist of $2n$-fermion operators, each adding n excitations to the state $|\psi_{\mathrm{HF}}\rangle$. For example:

$$T_1 = \sum_{\alpha,\beta} T_1^{\alpha\beta} c_\alpha^\dagger c_\beta, \quad T_2 = \sum_{\mu,\nu,\lambda,\rho} T_2^{\mu\nu\lambda\rho} c_\mu^\dagger c_\nu^\dagger c_\lambda c_\rho \tag{1.3.2}$$

With the appropriate coefficients in T_n, UCC is able to lower the energy beyond that of $|\psi_{\mathrm{HF}}\rangle$ by introducing the right type of correlations. As using the generators with all possible orders n is computationally prohibitive, one often truncates the UCC to single and double excitation operators T_1, T_2 only. This is also referred to as UCCSD (SD stands for 'singles and doubles'). The efficiency of this approach is provable on the perturbative level, where it is ensured by the linked-cluster theorem [57].

To utilize UCCSD in a variational quantum algorithm, one needs to implement the unitary in Eq. (1.3.1) as a digital quantum circuit. In most cases this is not possible directly, and instead requires a Trotter-Suzuki approximation [58, 59]. For example, a UCCSD unitary can be approximated ('Trotterized') into K Trotter steps as follows:

$$e^{T_1+T_2} \simeq \prod_{k=1}^{K} \left[\prod_{\alpha,\beta} e^{\frac{1}{K}\left(T_1^{\alpha\beta} c_\alpha^\dagger c_\beta + \mathrm{h.c.}\right)} \prod_{\mu,\nu,\lambda,\rho} e^{\frac{1}{K}\left(T_2^{\mu\nu\lambda\rho} c_\mu^\dagger c_\nu^\dagger c_\lambda c_\rho + \mathrm{h.c.}\right)} \right]. \tag{1.3.3}$$

Given this form of the ansatz, realizing (1.3.3) as a quantum circuit is a routine procedure. Indeed, fermionic operators can be mapped onto Pauli matrices, and exponentials of Pauli strings can be represented with a quantum circuit using the standard procedures of [60]. By tuning the coefficients $T^{\alpha\beta}$, one can then employ the ansatz circuit (1.3.3) in a variational procedure.

Along with UCC, there exist alternative approaches to ansatz construction. One example is the Hamiltonian Variational Ansatz [56], which is based on the principle of adiabatic state preparation. Since adiabatic evolution is capable of producing complex ground states, its crudely Trotterized version is expected to have similar capabilities. The accuracy of such preparation is further enhanced by the variational procedure, making Hamiltonian Variational Ansatz a sound approach. Another method is

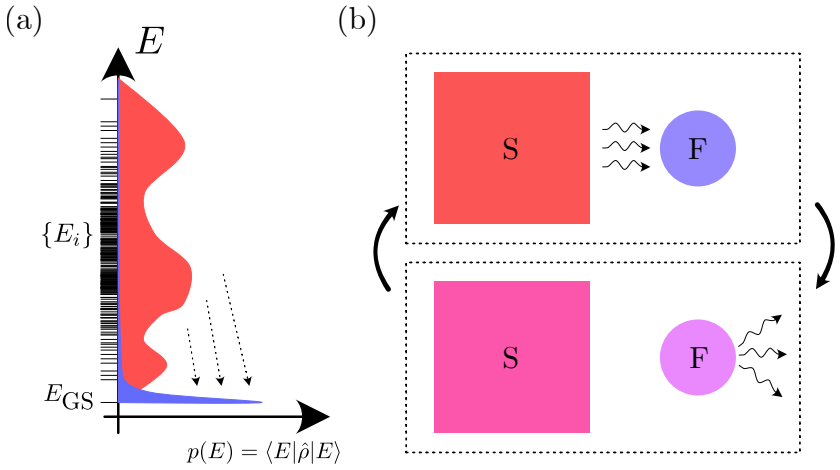

Figure 1.5: Quantum cooling. (a) In quantum mechanics, cooling implies a modification of the density matrix such that it shifts to lower energies (red to blue profile). With perfect cooling, the final state would be dominated by the ground state (blue profile). (b) To cool a physical system (S), it is standard to mediate the process by a controlled system (fridge, F). The process is two-fold. Firstly, the system and fridge are brought to interact. The energy from the system is then spontaneously transferred to the fridge (illustrated by a change of color). Secondly, the fridge has to emit surplus energy into the environment. The latter process is to be driven by external control, and would not happen under equilibrium thermodynamics. In algorithmic quantum cooling, one can achieve this by manually driving the fridge to its ground state.

ADAPT [55], which suggests to add tunable elements to the quantum circuit one at a time, while optimizing the parameters of such an ever-growing ansatz. The added unitary is to be picked adaptively, by estimating and optimizing the energy decrease that is expected from this addition. ADAPT is digital by construction and automatically adapts itself to the problem of interest, thus forming another viable approach to ansatz design. The methods presented above have both strengths and limitations — each is founded on specific analytical consideration, with some concessions made for compatibility with near-term quantum hardware. As such, there is an ongoing effort to design new methods of ansatz construction. The ultimate goal of these novel methods is to improve the stringent trade-off between the capacity of the circuit and its ease of implementation.

1.3.2 Dissipative quantum algorithms

Dissipative approach to quantum computing [61] implies performing a computation through an engineered evolution of an open quantum system, rather than an isolated one. This paradigm is less developed relative to the more established scheme of unitary-based computing, albeit it is theoretically sound. In particular, it was proven that any polynomially-complex algorithm performed with a standard quantum circuit can be mapped onto dissipative quantum hardware with only polynomial overhead [61]. Any evolution of an open quantum system, conversely, can be directly realized on digital quantum hardware by using ancillary registers. Intriguingly, with the latter procedure, one may apply the dissipative paradigm directly to digital quantum computing — potentially leading to new and powerful algorithms.

In the context of applying digital quantum hardware to ground state preparation, the dissipative paradigm is of direct relevance. Indeed, the low-energy states in nature are normally produced via a dissipative process of cooling (Fig. 1.5). Cooling is also employed in quantum engineering, for instance when initializing a null computational basis state of a quantum computer [62–64]. In the context of digital quantum computation, cooling was proposed early on [12] as a method of ground state simulation. In this case, even a single ancillary qubit can be used to emulate a fridge. Indeed, to drive the qubit to its ground state is straightforward — and being able to prepare the low energy state of the fridge is key to controlled cooling (Fig. 1.5b). Despite this being an interesting possibility, this idea remained largely undeveloped after the work of Lloyd [12]. Instead, most of research in algorithmic quantum cooling is focused on analog quantum computation [65, 66].

Another way to include dissipation into a state preparation protocol is to introduce weak measurements [67, 68]. Compared to conventional (projective) quantum measurements, weak measurements extract less information about the system and modify its state in a less drastic fashion. These unusual measurements require an auxilliary 'detector' system, which is coupled to the subject system only for a brief period of time. Application of such weak measurements to quantum state preparation is actively studied under the framework of quantum control theory [69–73] . In the traditional approaches, the system is controlled quantum-coherently, and the information obtained from the weak measurements can be fed back to modify the direction of this unitary evolution. A feedback of this type is classified as closed-loop quantum control. Compared to pre-defined unitary evolution, including such feedback allows to complete the state

preparation task with fewer resources or on a shorter timescale. Closed-loop quantum control can even be taken to the extreme, by removing the quantum-coherent part of the evolution altogether [69, 74, 75]. In this case, the weak measurements are used to inform the subsequent evolution, which is also driven by weak measurements. This approach, which goes under the name "control-free control", is now under active investigation both in experiment and theory [69].

1.4 This thesis

Chapter 2

Non-abelian statistics in topological superconductors can be realized by exchanging vortices that host Majorana zero-modes. However, using the bulk Abrikosov vortices for this purpose has proven to be impractical. Instead, many experimental groups turned to using Majorana zero-modes at the ends of superconducting nanowires, an effort that has also proven to be challenging. In this chapter, we propose to use a topological superconductor to realize the non-abelian braiding of an itinerant edge vortex with a bulk vortex. A voltage-driven Josephson junction can be employed to deterministically produce an edge vortex available for a braiding procedure. After a braiding operation, the vortices are to be fused back at another Josephson junction utilizing another voltage bias. We predict that the charge produced after the fusion is sensitive to the braiding operation. In particular, a single electron charge is produced after the braiding, and no charge if no braiding has occurred.

Chapter 3

We put forward a scheme to realize and detect another phenomenon of non-abelian statistics - anyon fusion. For this, we suggest employing the chiral edge vortex architecture introduced in the previous chapter. In a topological superconductor, by fusing Majorana-hosting vortices one produces a mixture of zero and one Dirac fermion. We design an experiment to capture this property, using four edge modes and four Josephson junction terminals. We propose to create two pairs of such vortices (1 and 2, 3 and 4) at the input terminals, and fuse them at the output terminals in a different configuration (1 with 3, 2 with 4). We predict, that the vortex fusion produces an equal weight superposition of two electrons and no electrons in the two output channels. This reproduces the fusion rule with

a direct effect on the output observables: charge transfer and fermion parity. In particular, we show that (a) the average fermion parity in each of the two leads is exactly vanishing upon fusion and (b) charge transfer is directly correlated, with the difference of output currents exhibiting zero noise while their sum remaining noiseful.

Chapter 4

Parafermionic zero-modes allow to realize non-abelian braiding based on a Laughlin type Fractional Quantum Hall material. Unfortunately, they were not yet shown to be realized in an experiment. It is therefore of high interest to characterize parafermions with appropriate observables. To that end, we investigate the possibilities offered by the Clauser-Horne-Shimony-Holt (CHSH) inequality. It is a version of Bell inequality, which signifies the extent of quantum correlations between two separated physical systems. Firstly, we generalize this and other notions of Bell nonlocality to the context of non-hermitian and potentially non-commuting observables. These are characteristic for systems of parafermions, due to the anyonic nature of these quasiparticles. Secondly, for such generalized observables, we draw up several bounds and relations for the intra-system and inter-system correlations. We show how these can be probed with parafermions, and predict that our correlation bounds are saturated much tighter when the two subsystems host non-commuting, rather than commuting observables. Paradoxically, the non-commutation of observables in these separated anyonic systems could be interpreted as superluminal signaling. This is only a simulation of such signaling, however. We show that the relativistic causality is automatically restored when such paradoxical correlations are to be probed in a physical experiment.

Chapter 5

Successfully employing a variational quantum algorithm for ground state preparation requires the use of appropriate ansatz circuits. The main requirements for such an ansatz are high expressivity and ease of use with digital quantum hardware. One such prospective ansatz, inspired by computational quantum chemistry, is a Trotterized version of Unitary Coupled Cluster. The ansatz is analytically justified by the linked cluster theorem, which proves its efficiency in the perturbation theory. Unfortunately, this approach relies on Trotterization which is not exact. In this chapter, we put forward an approach to ansatz creation that follows the linked-cluster theorem while not relying on Trotterization. The basis of the construction

is Quantum Combinatorial Ansatz (QCA). QCA is given by a sequence of tunable Pauli string-generated rotations, spanning the entire N-qubit Hilbert space in a minimum number of such elementary rotations. We give a rule for a systematic reduction of QCA to practical size, based on many-body perturbation theory. This ansatz construction turns out to satisfy the linked-cluster theorem, therefore proving its efficiency in the perturbative limit. Finally, we numerically test a few variants of QCA-based ansatz constructions by applying them to Ising spin chains. We find that these allow for a good asymptotic convergence to the ground state in the paramagnetic and the ferromagnetic phases of this model. As expected from perturbative analytics, in the weakly coupled limit of the model, the variant of an ansatz construction that satisfies the linked-cluster theorem shows optimal performance.

Chapter 6

Using dissipative approaches in digital quantum hardware is an area of research that is currently under active development. One natural application is ground state preparation. In this chapter, we use the principle of cooling to design a ground state preparation algorithm. To that end, we propose to simulate the interaction between the target system and a single qubit, transferring the energy to the qubit. The qubit, therefore, plays the role of a fridge. To simulate the cooling process, this has to be combined with the internal system and fridge evolutions. The final ingredient of the protocol is the measurement and reset of the fridge qubit, which introduces the dissipative element necessary to reduce the system energy. We develop the protocol by first pinning down the case of a single-qubit system and then extending the protocol to a general system of N qubits. The role of energy conservation is established, given by an extension of the Fermi Golden Rule to the case of finite-time evolution. We study the issues of Trotter error and Heisenberg energy uncertainty and propose two scalable approaches to tackle these. The BangBang approach, characterized by high Trotter error and large Heisenberg uncertainty, yields low implementation complexity and is thus suitable for near-term implementation. The complementary LogSweep approach yields an asymptotically vanishing energy uncertainty and Trotter error. However, this asymptotic accuracy is at the cost of extensive circuit complexity. Numerically, we apply LogSweep to transverse-field Ising chains in their paramagnetic and critical states. The time complexity of such ground preparation turns out to scale algebraically with a desired precision.

Chapter 7

This chapter is focused on another dissipative approach to quantum state preparation: steering by generalized measurements. To reduce the time in which the target state is prepared, we propose to actively choose the applied measurements on-the-go. This active choice is to be based on the information obtained from the previous measurements. The possible policies for active-decision steering are influenced by the presence of entanglement in the target state. For strongly correlated target states, creating a policy for active-decision steering is most challenging, due to the vastness of the many-body Hilbert space. Therefore, for efficient decision-making special Hilbert space "navigation techniques" are needed. Two approximate ways of representing such navigation are developed: via (i) a cost function landscape and (ii) a semiclassical Quantum State Machine. From such simplified Hilbert space representations, one derives the respective active-decision policies. We numerically apply these policies to two paradigmatic targets: AKLT state and W-state. In each case, the introduction of active decision achieves up to factor 10 speed-up of the target state preparation.

2 Deterministic creation and braiding of chiral edge vortices

2.1 Introduction

Non-Abelian anyons have the property that a pairwise exchange operation may produce a different state, not simply related to the initial state by a phase factor [76]. Because such "braiding" operations are protected from local sources of decoherence they are in demand for the purpose of quantum computations [3]. Charge $e/4$ quasiparticles in the $\nu = 5/2$ quantum Hall effect were the first candidates for non-Abelian statistics [77], followed by vortices in topological superconductors [32, 78].

Because experimental evidence for non-Abelian anyons in the quantum Hall effect [79, 80] has remained inconclusive, the experimental effort now focuses on the superconducting realizations [81]. While the mathematical description of the braiding operation (the Clifford algebra) is the same in both realizations, the way in which braiding is implemented is altogether different: In the quantum Hall effect one uses the chiral motion along the edge to exchange pairs of non-Abelian anyons and demonstrate non-Abelian statistics [82–84]. In contrast, in a superconductor the non-Abelian anyons are midgap states ("zero-modes") bound to a defect (a vortex [25, 85] or the end-point of a nanowire [86–88]). Because they are immobile, existing proposals to demonstrate non-Abelian statistics do not actually exchange the zero-modes in real space [89–93].

Topological superconductors do have chiral edge modes [78], and recent experimental progress [27] has motivated the search for ways to use the chiral motion for a braiding operation [33]. The obstruction one needs to overcome is that the Majorana fermions which propagate along the edge of a superconductor have conventional *fermionic* exchange statistics. In the quantum Hall effect each charge $e/4$ quasiparticle contains a zero-mode and the exchange of two quasiparticles is a non-Abelian operation on a topological qubit encoded in the zero-modes. However, Majorana fermions

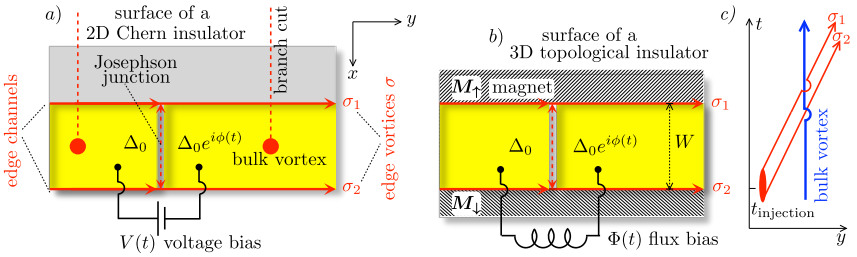

Figure 2.1: Panels *a)* and *b)*: Josephson junction geometries to deterministically inject a pair of edge vortices σ_1, σ_2 in chiral edge channels at opposite boundaries of a superconductor (yellow). The injection happens in response to a 2π increment in the superconducting phase difference $\phi(t)$, driven by a time-dependent voltage $V(t)$ or flux $\Phi(t)$. In panel *a)* edge vortex σ_1 crosses the 2π branch cut of a bulk vortex, resulting in a fermion parity switch. Panel *c)* shows the corresponding braiding of world lines in space-time: an overpass indicates that the vortex crosses a branch cut.

contain no zero-mode which might encode a topological qubit, one needs vortices for that.

In this chapter we show how one can exploit the chiral motion along the edge of a topological superconductor to exchange zero-modes in real space. The key innovative element of our design, which distinguishes it from Ref. 33, is the use of a biased Josephson junction to *on demand* inject a pair of isolated vortices into chiral edge channels. Previous studies of such "edge vortices" relied on quantum fluctuations of the phase to create a vortex pair in the superconducting condensate [94–97], but here the injection is entirely *deterministic*. When the two mobile edge vortices encircle a localized bulk vortex their fermion parity switches from even to odd, as a demonstration of non-Abelian braiding statistics. The entire operation, injection–braiding–detection, can be carried out fully electrically, without requiring time-dependent control over Coulomb interactions or tunnel probabilities.

2.2 Edge vortex injection

Fig. 2.1 shows different ways in which the edge vortex can be injected: driven by a flux bias or by a voltage bias over a Josephson junction. We show two possible physical systems that support chiral edge channels moving in the *same* direction on opposite boundaries of the superconductor.

Both are hybrid systems, where a topologically trivial superconductor (spin-singlet s-wave pair potential Δ_0) is combined with a topologically nontrivial material: a 2D Chern insulator (quantum anomalous Hall insulator) [27, 98] (panel a) or a 3D topological insulator gapped on the surface by ferromagnets with opposite magnetisation $M_{\uparrow,\downarrow}$ [94, 99] (panel b).

The superconducting phase difference $\phi(t)$ across the Josephson junction is incremented with 2π by application of a voltage pulse $V(t)$ (with $\int V(t)dt = h/2e$), or by an $h/2e$ increase of the flux $\Phi(t)$ through an external superconducting loop. If the width W of the superconductor is large compared to the coherence length $\xi_0 = \hbar v/\Delta_0$, the edge channels at $x = \pm W/2$ are not coupled by the Josephson junction — except when ϕ is near π, as follows from the junction Hamiltonian [25, 99]

$$H_{\rm J} = v p_x \sigma_z + \Delta_0 \sigma_y \cos(\phi/2). \tag{2.2.1}$$

The Pauli matrices act on excitations moving in the $\pm x$ direction with velocity v, in a single mode for ξ_0 large compared to the thickness of the junction in the y-direction.

At $\phi = \pi$ a Josephson vortex passes through the superconductor [100, 101]. A Josephson vortex is a 2π phase winding for the pair potential, so a π phase shift for an unpaired fermion. As explained in Ref. 102, the passage of the Josephson vortex leaves behind a pair of edge vortices: a phase boundary $\sigma(y)$ on each edge, at which the phase of the Majorana fermion wave function $\psi(y)$ jumps by π. Because of the reality constraint on ψ, a π phase jump (a minus sign) is stable: it can only be removed by merging with another π phase jump. And because the phase boundary is tied to the fermion wave function, it shares the same chiral motion, $\sigma(y,t) = \sigma(y - vt)$.

2.3 Braiding of an edge vortex with a bulk vortex

Two vortices may be in a state of odd or even fermion parity, meaning that when they fuse they may or may not leave behind an unpaired electron. The fermion parity of vortices σ_1 and σ_2 is encoded in the ± 1 eigenvalue of the parity operator $P_{12} = i\gamma_1\gamma_2$, where γ_n is the Majorana operator associated with the zero-mode in vortex n^*. The two edge vortices

*Abrikosov vortices in the bulk have a normal core, which edge vortices lack. Both are non-Abelian anyons because a zero-mode does not need a normal core, see the explicit calculation for a coreless Josephson vortex in Ref. 103.

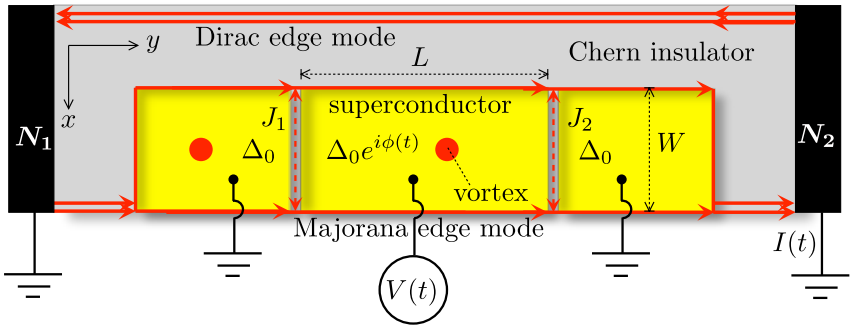

Figure 2.2: Starting from the layout of Fig. 2.1*a*, we have inserted a second Josephson junction (J_2) and we have added normal metal contacts (N_1, N_2) to measure the current $I(t)$ carried by the edge modes in response to the voltage $V(t)$ applied to the superconductor. A unit charge per 2π increment of ϕ is transferred from the superconductor into the normal metal contact. The counterpropagating Dirac edge mode along the upper edge of the Chern insulator is decoupled from the superconductor and plays no role in the analysis.

are created at the Josephson junction in a state of even fermion parity, $P_{12} = +1$, but as illustrated in Fig. 2.1*a* that may change as they move away from the junction: If one of the edge vortices, say σ_1, crosses the branch cut of the phase winding around a bulk vortex, γ_1 picks up a minus sign and the fermion parity $P_{12} \mapsto -1$ switches from even to odd [32]. This is the essence of the non-Abelian braiding statistics of vortices.

2.4 Detection of the fermion-parity switch

Fig. 2.2 shows the voltage-biased layout for a fully electrical measurement. The fermion parity of the edge vortices cannot be detected if they remain separated on opposite edges, so we first fuse them at a second Josephson junction. The characteristic time scale of the injection process is the time $t_{\rm inj} \simeq (\xi_0/W)(d\phi/dt)^{-1}$ when $\phi(t)$ is within ξ_0/W from π, and if the distance L between the two Josephson junctions is less than $vt_{\rm inj}$ we can neglect the time delay between the injection at the first junction J_1 and the fusion at the second junction J_2. This is convenient, because then the whole process can be driven by a single voltage pulse $V(t)$ applied to the region $|y| < L/2$ between the two junctions, relative to the grounded regions $y < -L/2$ and $y > L/2$ outside.

Both these grounded regions are connected to normal metal electrodes

N_1 and N_2 and the electrical current $I(t)$ between them is measured. As we will now show, the transferred charge $Q = \int I(t)dt$ is quantized at unit electron charge if the region between the Josephson junctions contains a bulk vortex, while $Q = 0$ if it does not.

2.5 Mapping onto a scattering problem

Tunneling of edge vortices driven by quantum fluctuations of the phase is a many-body problem of some complexity [102]. We avoid this because we rely on an external bias to inject the edge vortices, hence the phase $\phi(t)$ can be treated as a classical variable with a given time dependence.

The dynamics of the Majorana fermions remains fully quantum mechanical, governed by the Hamiltonian

$$H = i \begin{pmatrix} -v\partial/\partial y & -\mu[y,\phi(t)] \\ \mu[y,\phi(t)] & -v\partial/\partial y \end{pmatrix} \equiv v p_y \sigma_0 + \mu \sigma_y. \tag{2.5.1}$$

(We set $\hbar = 1$.) The 2×2 Hermitian matrix H acts on the Majorana fermion wave functions $\Psi = (\psi_1, \psi_2)$ at opposite edges of the superconductor, both propagating in the $+y$ direction (hence the unit matrix σ_0) The interedge coupling μ multiplies the σ_y Pauli matrix to ensure that H is purely imaginary and the wave equation $\partial\Psi/\partial t = -iH\Psi$ is purely real (as it should be for a Majorana fermion).

For low-energy, long-wavelength wave packets the y-dependence of the interedge coupling may be replaced by a delta function, $\mu[y,\phi(t)] = v\delta(y)\eta(t)$. This "instaneous scattering approximation" [104] is valid if the transit time $t_{\rm transit} \simeq L/v$ of the wave packet through the system is short compared to the characteristic time scale $t_{\rm inj}$ of the vortex injection, hence if $d\phi/dt \ll v\xi_0/A_{\rm junction}$, where $A_{\rm junction} = WL$ is the area of the region between J_1 and J_2. In this regime there is no need to explicitly consider the vortex dynamics in between the Josephson junctions, instead we can treat this as a scattering problem "from the outside".

Incoming and outgoing states are related by

$$\Psi_{\rm out}(E) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, S(\omega)\Psi_{\rm in}(E-\omega), \tag{2.5.2}$$

where $S(\omega)$ is the adiabatic (or "frozen") scattering matrix,

$$S(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\omega t} S(t), \quad S(t) = \exp\bigl(-i\eta(t)\sigma_y\bigr), \tag{2.5.3}$$

describing the scattering at $E = 0$ for a fixed $\phi(t)$.

As we shall see in a moment, the transferred charge is independent of how $\eta(t) = \eta[\phi(t)]$ is varied as a function of time, only the net increment $\delta\eta = \eta(t \to \infty) - \eta(t \to -\infty)$ matters. When there is no vortex in the region between the two Josephson junctions J_1 and J_2 there is no difference between $\phi = 0$ and $\phi = 2\pi$ hence $\delta\eta = 0$. On the contrary, when there is a bulk vortex in this region we find*

$$\eta = 2 \arccos\left(\frac{\cos(\phi/2) + \tanh\beta}{1 + \cos(\phi/2)\tanh\beta}\right), \quad \beta = \frac{W}{\xi_0}\cos\frac{\phi}{2}, \tag{2.5.4}$$

hence $\delta\eta = 2\pi$. More generally, when there are N_{vortex} vortices between J_1 and J_2 the phase increment is

$$\delta\eta = \pi(1 - (-1)^{N_{\text{vortex}}}). \tag{2.5.5}$$

In Fig. 2.3 we show that the analytical result (2.5.4) agrees well with a computer simulation (using Kwant [105]) of a lattice model of a quantum anomalous Hall insulator with induced s-wave superconductivity [98].

2.6 Transferred charge

The expectation value of the transferred charge†,

$$Q = e\int_0^\infty \frac{dE}{2\pi}\,\langle \Psi^\dagger_{\text{out}}(E)\sigma_y\Psi_{\text{out}}(E)\rangle, \tag{2.6.1}$$

is given at zero temperature, when

$$\langle \Psi_{\text{in},n}(E)\Psi_{\text{in},m}(E')\rangle = \delta_{nm}\delta(E - E')\theta(-E), \tag{2.6.2}$$

by an integral over positive excitation energies,

$$Q = \frac{e}{4\pi^2}\int_{0^+}^\infty d\omega\,\omega\,\mathrm{Tr}\,S^\dagger(\omega)\sigma_y S(\omega). \tag{2.6.3}$$

(The factor $\omega = \int_0^\infty dE\,\theta(\omega - E)$ appears from the integration over the step function.) Because $S(-\omega) = S^*(\omega)$ the integrand in Eq. (2.6.3) is an

*The calculation of the scattering phase shift $\eta(\phi)$ is given in App. 2.A. Eq. (2.5.4) for $0 \le \phi \le 2\pi$ repeats periodically modulo 2π.

†The charge operator $Q = e\sigma_z$ in the electron-hole basis transforms into $Q = e\sigma_y$ in the basis of Majorana fermions.

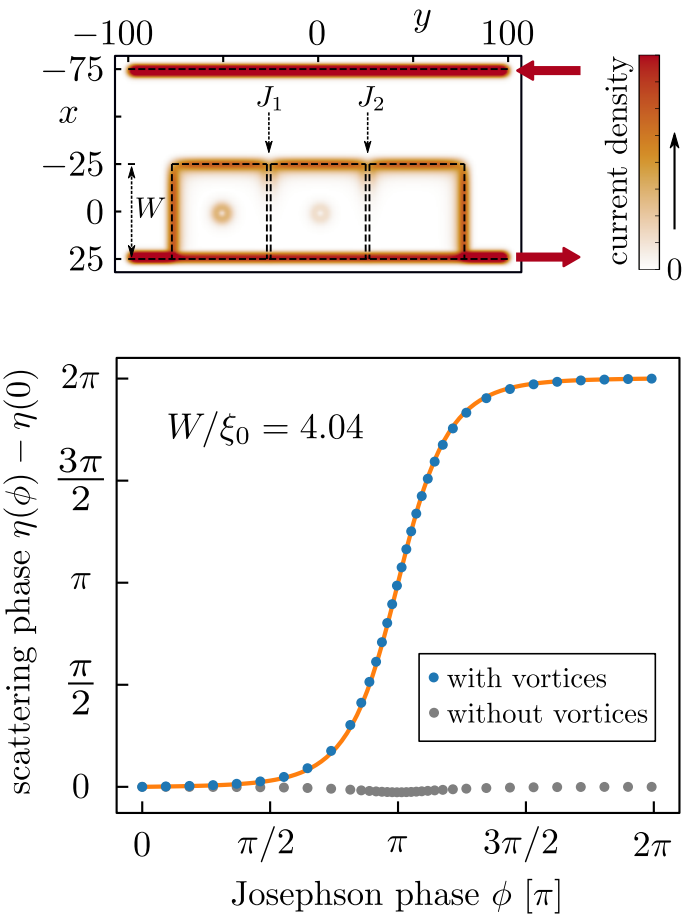

Figure 2.3: *Bottom panel:* Scattering phase $\eta(\phi) - \eta(0)$ according to Eq. (2.5.4) (solid curve) and as obtained numerically (blue data points) from a lattice model [98] of the system shown in Fig. 2.2. There are no fit parameters in the comparison, the ratio $W/\xi_0 = 4.04$ was obtained directly from the simulation The grey data points show the result without vortices, when there is no net increment as ϕ advances from 0 to 2π.

even function of ω and the integral can be extended to negative ω,

$$\begin{aligned} Q &= \frac{e}{8\pi^2} \int_{-\infty}^{\infty} d\omega\, \omega \,\mathrm{Tr}\, S^\dagger(\omega)\sigma_y S(\omega) \\ &= \frac{ie}{4\pi} \int_{-\infty}^{\infty} dt\, \mathrm{Tr}\, S^\dagger(t)\sigma_y \frac{\partial}{\partial t} S(t). \end{aligned} \tag{2.6.4}$$

This is the superconducting analogue of Brouwer's charge-pumping formula [106] (see Ref. 107 for an alternative derivation).

Substitution of $S(t) = \exp\bigl(-i\eta(t)\sigma_y\bigr)$ results in

$$Q = (e/2\pi)\delta\eta = e \tag{2.6.5}$$

if N_{vortex} is odd, while $Q = 0$ if N_{vortex} is even.

2.7 Transferred particle number

This quantized transfer of one electron charge may be accompanied by the non-quantized transfer of neutral electron-hole pairs. To assess this we calculate the expectation value of the transferred particle number, given by Eq. (2.6.3) upon substitution of the charge operator $e\sigma_y$ by unity:

$$N_{\text{particles}} = \frac{1}{4\pi^2}\int_{0+}^{\infty} d\omega\,\omega\,\mathrm{Tr}\,S^\dagger(\omega)S(\omega). \tag{2.7.1}$$

This integrand is an odd function of ω, so we cannot easily transform it to the time domain.

We proceed instead by calculating $S(\omega)$ from Eq. (2.5.3), in the approximation $\eta(t) \approx 2\arccos[-\tanh(t/t_{\text{inj}})]$, accurate when $W/\xi_0 \gg 1$. The result is

$$\begin{aligned} S(\omega) &= -\frac{2\pi\omega t_{\text{inj}}^2\sigma_0}{\sinh(\pi\omega t_{\text{inj}}/2)} - \frac{2\pi\omega t_{\text{inj}}^2\sigma_y}{\cosh(\pi\omega t_{\text{inj}}/2)} - 2\pi\delta(\omega) \\ \Rightarrow N_{\text{particles}} &= (84/\pi^4)\zeta(3) = 1.037. \end{aligned} \tag{2.7.2}$$

One can construct a special t-dependent phase variation* that makes $N_{\text{particles}}$ exactly equal to unity, by analogy with the "leviton" [104, 108], but even without any fine tuning the charge transfer is nearly noiseless.

2.8 Discussion

We have shown how the chiral motion of edge modes in a topological superconductor can be harnessed to braid a pair of non-Abelian anyons: one immobile in a bulk vortex, the other mobile in an edge vortex. The experimental layout of Fig. 2.2 is directly applicable to the recently reported

*The special time dependence $\eta(t) = \pi + 2\arctan(t/t_{\text{inj}})$ produces precisely one particle with charge e.

chiral Majorana fermion modes in quantum anomalous Hall insulator–superconductor structures [27, 109].

While the presence of a bulk vortex and the crossing of its branch cut is essential for the charge transfer, it is of the essence for braiding that no tunnel coupling or Coulomb coupling to the edge vortices is needed. This distinguishes the braiding experiment proposed here to tunnel probes of Majorana zero-modes that can also produce a quantized charge transfer [107]. In the quantum Hall effect attempts to use edge modes for braiding [80] have been inconclusive because of Coulomb coupling with bulk quasiparticles [110]. The superconductor offers a large gap, to suppress tunnel coupling, and a large capacitance, to suppress Coulomb coupling, which could make the edge mode approach to braiding a viable alternative to existing approaches using zero-modes bound to superconducting nanowires [89–93].

In the quantum Hall effect there is a drive to use quasiparticles in edge modes as "flying qubits" for quantum information processing [111]. Edge vortices in a topological superconductor could play the same role for topological quantum computation. The pair of edge vortices in the geometry of Fig. 2.1*a* carries a topologically protected qubit encoded in the fermion parity. The deterministic voltage-driven injection of edge vortices that we have proposed here could become a key building block for such applications.

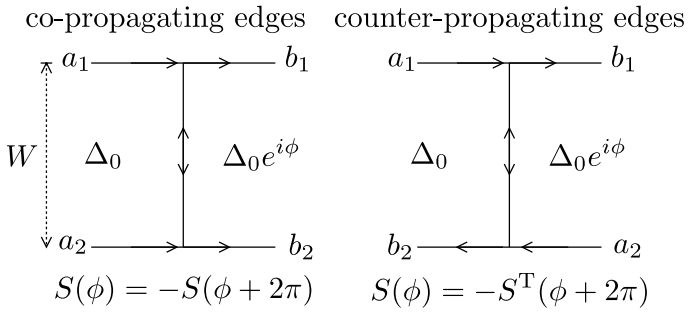

Figure 2.4: Two scattering geometries of a Josephson junction connecting chiral Majorana edge modes. We contrast the case of co-propagating modes in the left panel, with the case of counter-propagating modes in the right panel.

2.A Calculation of the scattering phase shift

We calculate the scattering phase shift at the Fermi level in the double Josephson junction geometry of Fig. 2.2. We first consider a single Josephson junction, shown schematically in Fig. 2.4. We specify the phase difference ϕ in the interval $(-2\pi, 2\pi)$, to accomodate the 4π-periodicity of the junction Hamiltonian (2.2.1).

2.A.1 Single Josephson junction

The scattering matrix S_{J} of the Josephson junction relates incoming and outgoing amplitudes via

$$S_{\mathrm{J}} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}. \tag{2.A.1}$$

At the Fermi level $S_{\mathrm{J}} \in \mathrm{SO}(2)$ is a 2×2 orthogonal matrix with determinant $+1$, of the general form

$$S_{\mathrm{J}} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} = e^{i\alpha\sigma_y}. \tag{2.A.2}$$

We seek the ϕ-dependence of the phase shift $\alpha(\phi)$, in particular the increment $\delta\alpha = \alpha(2\pi) - \alpha(0)$.

In our geometry of co-propagating edge modes (left panel in Fig. 2.4), the incoming modes are on one side of the junction and the outgoing modes are at the other side. Fu and Kane [99] studied a different geometry with

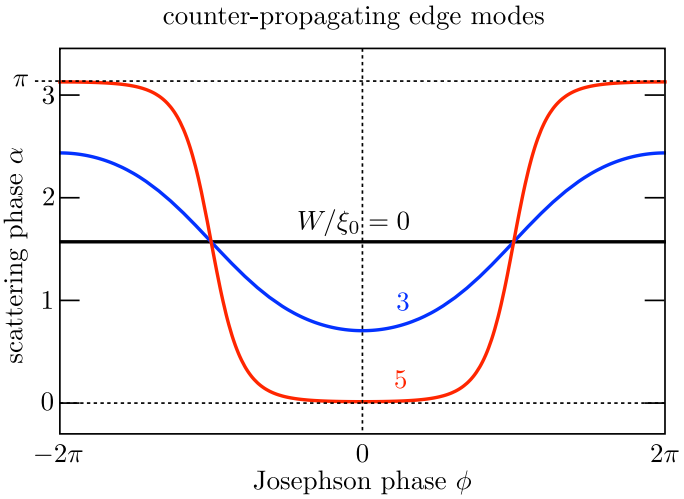

Figure 2.5: Plot of the ϕ-dependence of the scattering phase shift α of a Josephson junction between counter-propagating Majorana edge modes. The plot is calculated from Eq. (2.A.5) for three values of the ratio W/ξ_0.

counter-propagating modes (right panel), where the two incoming modes, as well as the two outgoing modes, are on opposite sides of the junction. As we shall see, the difference is crucial for the quantization of $\delta\alpha$.

Counter-propagating edge modes

For counter-propagating edge modes the SO(2) scattering matrix is [99]

$$S_{\mathrm{J}} = \begin{pmatrix} \tanh\beta & 1/\cosh\beta \\ -1/\cosh\beta & \tanh\beta \end{pmatrix}, \quad \beta = \frac{W}{\xi_0}\cos(\phi/2). \tag{2.A.3}$$

If the superconducting phase difference ϕ across the junction is advanced by 2π, a fermion crossing the junction experiences a phase shift of π. Hence the diagonal matrix elements of S_{J} change sign, while the off-diagonal elements do not change sign, as expressed by the symmetry relation

$$S_{\mathrm{J}}(\phi + 2\pi) = -S_{\mathrm{J}}^{\mathrm{T}}(\phi). \tag{2.A.4}$$

The scattering phase shift α in $S_{\mathrm{J}} = e^{i\alpha\sigma_y}$ from Eq. (2.A.3) equals

$$\alpha = \arccos(\tanh\beta) \in (0, \pi), \tag{2.A.5}$$

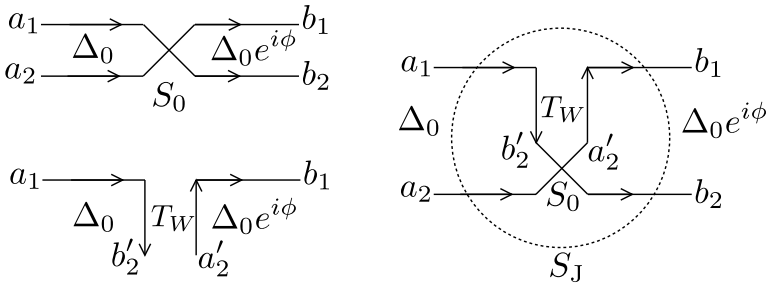

Figure 2.6: Scattering geometries that define S_0, T_W, and S_J.

plotted in Fig. 2.5. The increment

$$\delta\alpha \approx \pi - 4e^{-W/\xi_0} \tag{2.A.6}$$

approaches π for $W/\xi_0 \to \infty$, but it is not quantized. Also note that $\alpha(-\phi) = \alpha(\phi)$, so the net phase increment over a 4π period is zero. Both these results for counter-propagating modes change when we consider co-propagating modes.

Co-propagating edge modes

In the case of co-propagating edge modes, as in the left panel of Fig. 2.4, each element of the scattering matrix S_J relates amplitudes on opposite sides of the Josephson junction, so it should change sign when ϕ is advanced by 2π. Instead of Eq. (2.A.4) we thus have

$$S_J(\phi + 2\pi) = -S_J(\phi). \tag{2.A.7}$$

It follows that $\alpha(2\pi) = \alpha(0) + \pi$, modulo 2π, hence the phase increment

$$\delta\alpha = \pi \tag{2.A.8}$$

is *exactly* quantized, independent of the ratio W/ξ_0. The step profile $\alpha(\phi)$ does depend on this ratio, as we now calculate.

We first consider the $W \to 0$ limit, when the Josephson junction is a point contact as in Fig. 2.6, upper left panel. The two incoming Majorana operators γ_1, γ_2 form an electron operator $c = (\gamma_1 - i\gamma_2)/\sqrt{2}$ that is transmitted through the junction with a $\phi/2$ phase shift,

$$c_{\text{out}} = e^{i\phi/2} c_{\text{in}}. \tag{2.A.9}$$

The corresponding scattering matrix for the Majorana modes is

$$S_0 = e^{i(\phi/2)\sigma_y}. \tag{2.A.10}$$

For a finite width W we insert a line junction described by the Hamiltonian (2.2.1). The corresponding transfer matrix (lower left panel in Fig. 2.6) is

$$T_{\rm W} = e^{-\beta\sigma_x}, \quad \beta = \frac{W}{\xi_0}\cos(\phi/2). \tag{2.A.11}$$

Combination of S_0 and T_W (right panel in Fig. 2.6) produces upon mode matching the 2×2 scattering matrix $S_{\rm J}$ of the entire Josephson junction,

$$\begin{aligned} &T_W \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} b_2' \\ a_2' \end{pmatrix}, \quad S_0 \begin{pmatrix} b_2' \\ a_2 \end{pmatrix} = \begin{pmatrix} a_2' \\ b_2 \end{pmatrix}, \\ &\Rightarrow S_{\rm J} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}. \end{aligned} \tag{2.A.12}$$

The result is

$$\begin{aligned} S_{\rm J} = &\frac{1}{\cosh\beta + \cos(\phi/2)\sinh\beta} \\ &\times \begin{pmatrix} \cos(\phi/2)\cosh\beta + \sinh\beta & \sin(\phi/2) \\ -\sin(\phi/2) & \cos(\phi/2)\cosh\beta + \sinh\beta \end{pmatrix}. \end{aligned} \tag{2.A.13}$$

The corresponding scattering phase shift in $S_{\rm J} = e^{i\alpha\sigma_y}$ is

$$\alpha = \arccos\left(\frac{\cos(\phi/2) + \tanh\beta}{1 + \cos(\phi/2)\tanh\beta}\right) \times \mathrm{sign}\,(\phi). \tag{2.A.14}$$

It increases monotonically from $\alpha = -\pi$ at $\phi = -2\pi$ through $\alpha = 0$ at $\phi = 0$ to $\alpha = \pi$ at $\phi = 2\pi$. As shown in Fig. 2.7, the increase starts out linearly for $W/\xi_0 \ll 1$, and then becomes more and more step-function like with increasing W.

For $W/\xi_0 \gg 1$ the ϕ-dependence of α is described with exponential accuracy by

$$\alpha \approx \arccos(\tanh\beta) \times \mathrm{sign}\,(\phi), \tag{2.A.15}$$

as in Eq. (2.A.5), but now antisymmetric in ϕ. As a consequence the net phase increment over a 4π period equals 2π rather than zero.

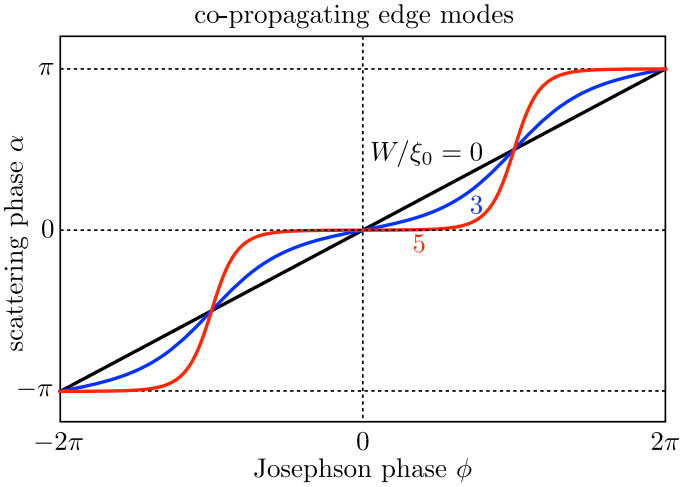

Figure 2.7: Same as Fig. 2.5, but for *co*-propagating Majorana edge modes, calculated from Eq. (2.A.14) .

2.A.2 Double Josephson junction

We combine two Josephson junctions in series with co-propagating edge modes, as in Fig. 2.2. We denote the result (2.A.13) by

$$S_{\mathrm{J}} = S_1 = e^{i\alpha_1\sigma_y}, \tag{2.A.16}$$

with a subscript 1 to indicate that this is the scattering matrix of the first Josephson junction (J_1 in Fig. 2.2).

If the second Josephson junction J_2 would be identical to the first, its scattering matrix would be $S_2(\phi) = S_1(-\phi) = S_1^{-1}$. We allow for a difference in the ratio W/ξ_0 at the two junctions, so more generally

$$S_2 = e^{-i\alpha_2\sigma_y}. \tag{2.A.17}$$

The parameter α_2 still increases by $\delta\alpha_2 = \pi$ for each 2π increment of ϕ, but it may do so with a different ϕ-dependence than α_1.

If there are no bulk vortices in the superconductor, the scattering matrix of the two junctions in series is simply the product S_2S_1. However, if the geometry is as in Fig. 2.2, with a pair of bulk vortices on opposite sides of the first Josephson junction, we have to insert a σ_z to account for each crossing of a branch cut, so the full scattering matrix is $S_2\sigma_z S_1\sigma_z$. The

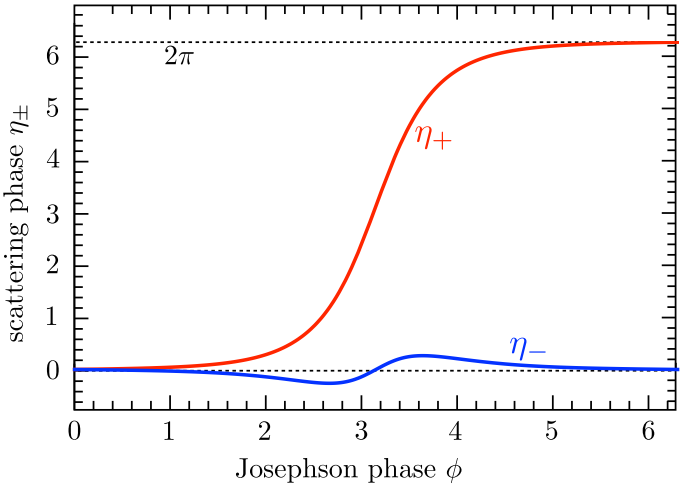

Figure 2.8: Plot of the ϕ-dependence of the two scattering phase shifts η_+ (with a bulk vortex in the region between the two Josephson junctions J_1 and J_2) and η_- (without a bulk vortex). The plot is calculated from Eqs. (2.A.14) and (2.A.18) for $W/\xi_0 = 3$ in the first junction and $W/\xi_0 = 5$ in the second jnction.

two cases can be combined as

$$S_\pm = e^{-i\eta_\pm \sigma_y}, \quad \eta_\pm = \alpha_2 \pm \alpha_1, \tag{2.A.18}$$

where S_+ and S_- refer, respectively, to the situation with or without the bulk vortices. More generally, if the region between Josephson junctions J_1 and J_2 has $N_{\rm vortex}$ bulk vortices, S_+ applies if $N_{\rm vortex}$ is odd while S_- applies if $N_{\rm vortex}$ is even.

When the phase ϕ across the Josephson junction varies from 0 to 2π both α_1 and α_2 advance from 0 to π. It follows that a 2π increment of ϕ induces a 2π increase of the scattering phase η_+, while η_- has no net increase. Fig. 2.8 illustrates the difference for a particular choice of parameters.

2.B Details of the numerical simulation

For the numerical simulation shown in Fig. 2.3 we applied the Kwant tight-binding code [105] to the Bogoliubov-De Gennes Hamiltonian of Qi,

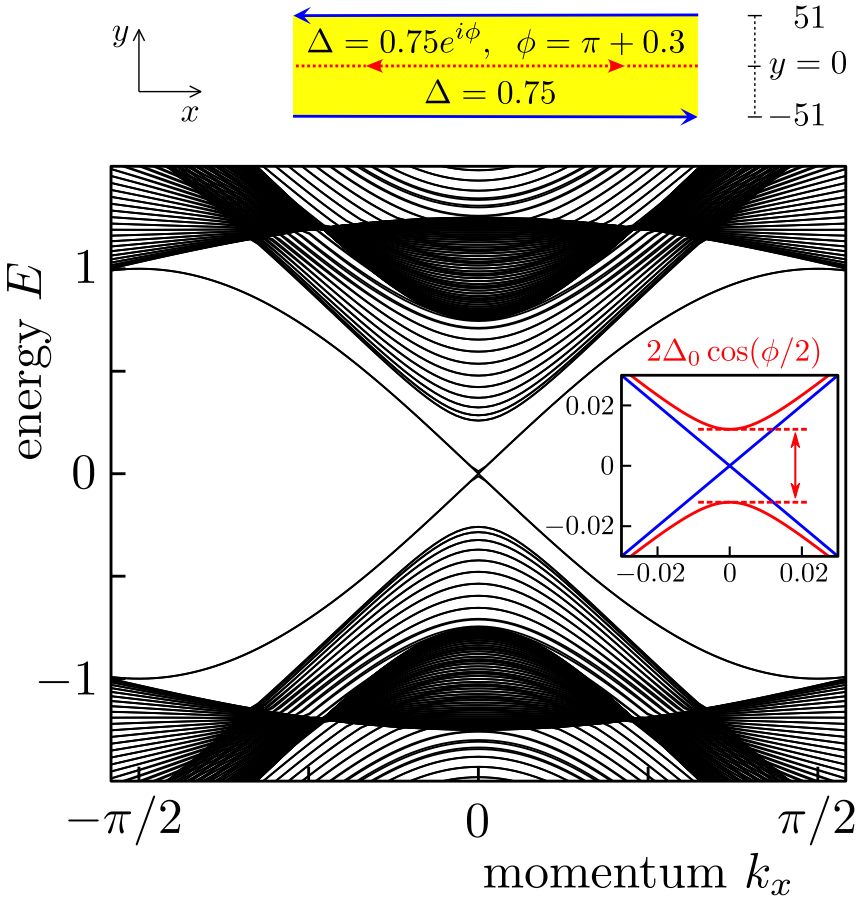

Figure 2.9: Dispersion relation of a superconducting strip in the region $1 < |y| < 51$, with a line junction along $y = 0$ ($\Delta = 0$, two lattice sites wide), separating regions with phase 0 and phase $\phi = \pi + 0.3$. Inset: Enlargement of the region near $k_x = 0$, $E = 0$. The blue modes are gapless chiral Majorana edge modes at the boundary of the superconductor, the red mode is a nonchiral Majorana mode in the line junction, with a gap of $2\Delta_0 \cos(\phi/2)$. The effective pair potential $\Delta_0 = 0.0808$ is much smaller than the bare value 0.75.

Hughes, and Zhang [98]:

$$\mathcal{H} = \begin{pmatrix} h_0(\boldsymbol{p}) - E_\mathrm{F} & i\Delta(\boldsymbol{r})\tau_y \\ -i\Delta^*(\boldsymbol{r})\tau_y & -h_0^*(-\boldsymbol{p}) + E_\mathrm{F} \end{pmatrix}, \tag{2.B.1a}$$

$$h_0(\boldsymbol{p}) = (C + Bp_x^2 + Bp_y^2)\tau_z + Ap_x\tau_x + Ap_y\tau_y. \tag{2.B.1b}$$

The blocks of $\mathcal{H}$ refer to the electron-hole degree of freedom, while the Pauli matrices τ_α act on the spin degree of freedom. This Hamiltonian was discretized on a two-dimensional square lattice. Lengths are measured in units of the lattice constant $a = 1$ and energies in units of the hopping matrix element $t_0 = 1$. We also set $\hbar = 1$, so that all parameters are dimensionless.

The electron block $h_0(\boldsymbol{p})$ describes a quantum anomalous Hall insulator. We took the parameters $A = 1$, $B = 0.5$, $C = -0.5$, $E_\mathrm{F} = 0$, when h_0 has Chern number 1. The insulator covers the region $-75 < x < 25$. The bulk is gapped while the edges support a single chiral edge mode. The mode at the $x = 25$ boundary moves in the $+y$ direction and a counterpropagating edge mode flows at the $x = -75$ boundary. (See Fig. 2.3, top panel, for the geometry.)

The pair potential Δ induces spin-singlet s-wave superconductivity in a strip $-25 < x < 25$ (so $W = 50$) streching from $y = -77$ to $y = 77$. We inserted two line junctions J_1, J_2 (each two lattice sites wide), separating three superconducting islands I_1, I_2, I_3, by means of the profile

$$\Delta = \begin{cases} 0.75 & \text{if } -77 < y < -27 \ \ (I_1), \\ 0 & \text{if } -27 < y < -25 \ \ (J_1), \\ 0.75\, e^{i\phi} & \text{if } -25 < y < 25 \ \ (I_2), \\ 0 & \text{if } 25 < y < 27 \ \ (J_2), \\ 0.75 & \text{if } 27 < y < 77 \ \ (I_3). \end{cases} \tag{2.B.2}$$

The effective gap $\Delta_0 \cos(\phi/2)$ in the Josephson junction was obtained directly from the excitation spectrum. (See Fig. 2.9.) We found $\Delta_0 = 0.0808$ — much smaller than the bare gap of 0.75. At $\phi = \pi$ the gap closes, producing a linear dispersion along the Josephson junction with velocity $v = 1$ — the same as the velocity of the edge modes. The corresponding coherence length is $\xi_0 = v/\Delta_0 = 12.38$, resulting in a ratio $W/\xi_0 = 4.04$.

A pair of vortices is inserted at positions $\boldsymbol{r}_1 = (1, -51)$ in I_1 and $\boldsymbol{r}_2 = (1, 1)$ in I_2. The vortex core does not coincide with a lattice point (which are at half-integer x, y), so we can keep a constant ampitude $|\Delta| = 0.75$ of the pair potential. Multiplication of $\Delta(\boldsymbol{r})$ by the function

$$f(\boldsymbol{r}) = \frac{z - z_1}{|z - z_1|} \frac{|z - z_2|}{z - z_2}, \quad z = x + iy, \tag{2.B.3}$$

ensures that the phase of the pair potential winds by $\pm 2\pi$ around each vortex.

The scattering matrix was calculated at an energy $E = 0.001$ that is slightly offset from the Fermi level at $E = 0$ to avoid the zero-mode

resonance in the vortex cores. The representation $S = e^{-i\eta\sigma_y}$ in the Majorana basis corresponds to $S = e^{-i\eta\sigma_z}$ in the electron-hole basis, so the scattering phase shift η can be calculated by comparing incident and transmitted electron wave functions along $x = 25$. The edge at $x = -75$ is decoupled from the superconductor and does not contribute to η.

At $\phi = 0$ we find $\eta \neq 0$, presumably because of additional phase shifts acquired when the Dirac mode splits into two Majorana modes and back at $\boldsymbol{r} = (25, -77)$ and $\boldsymbol{r} = (25, 77)$. In Fig. 2.3 we have plotted the phase increment $\eta(\phi) - \eta(0)$, to eliminate this ϕ-independent offset.

3 Electrical detection of the Majorana fusion rule for chiral edge vortices in a topological superconductor

3.1 Introduction

Vortices in a two-dimensional topological superconductor contain a midgap state, or *zero-mode*, that can be used to store quantum mechanical information in a nonlocal way, protected from local sources of decoherence [3, 4, 78, 112, 113]. The qubit degree of freedom is the fermion parity of any two widely separated vortices, which may or may not share an unpaired electron or hole (a fermionic quasiparticle) in the condensate of Cooper pairs. The pairwise exchange, or *braiding*, of vortices is a unitary transformation which can serve as a building block for a quantum computation [32, 114]. The merging, or *fusion*, of two vortices is the read-out operation [115]: The qubit is in the state $|1\rangle$ or $|0\rangle$ depending on whether or not the vortices leave behind a unpaired fermion. The fact that braiding operations do not commute, referred to as *non-Abelian statistics*, goes hand-in-hand with the fact that the fusion outcome is non-deterministic. As illustrated in Fig. 3.1, the fusion of two vortices σ produces a quantum superposition of states ψ and $\mathbb{I}$ with and without a quasiparticle excitation. This is the Majorana fusion rule* of non-Abelian anyons, symbolically written as $\sigma \otimes \sigma = \psi \oplus \mathbb{I}$.

Neither the braiding nor the fusion of vortices has been realized in the laboratory. This has motivated a variety of theoretical proposals for methods to demonstrate the appearance of non-Abelian anyons in a topological superconductor [90–93]. The obstacle that these proposals seek to remove, is the need to physically move the zero-modes around. Ref. [116] proposes an alternative approach: Substitute immobile bulk

*Because of a mapping onto the Ising model, the term "Ising fusion rule" is also used.

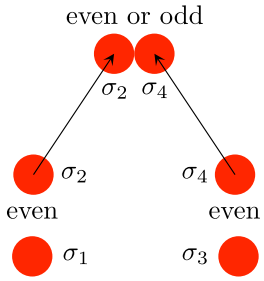

Figure 3.1: Schematic illustration of the fusion rule $\sigma_2 \otimes \sigma_4 = \psi \oplus \mathbb{I}$ of Majorana zero-modes (red dots, labeled σ_n). Pairs of zero-modes may or may not share a quasiparticle. In the former case the fermion parity is "odd" (indicated by ψ), in the latter case it is "even" (indicated by $\mathbb{I}$). The overall fermion parity is conserved, so if the fusion of σ_2 and σ_4 leaves behind a quasiparticle, then the fusion of σ_1 and σ_3 must also produce a quasiparticle.

vortices for mobile edge vortices. In that paper the braiding of vortices was considered. Here we turn to the fusion of edge vortices, in order to demonstrate the Majorana fusion rule.

Edge vortices are π-phase domain walls for Majorana fermions propagating along the edge of a topological superconductor [102]. Edge vortices may appear stochastically from quantum phase slips at a Josephson junction [95–97], but for our purpose we use the *deterministic* injector of Ref. [116]: A voltage pulse $V(t)$ of integrated magnitude $\int V(t)dt = h/2e$ applied over a Josephson junction injects an edge vortex at each end of the junction. The injection happens when the phase difference ϕ of the superconducting pair potential crosses π. At $\phi = \pi$ the effective gap $\Delta_0 \cos(\phi/2)$ in the junction changes sign [25]. By the same mechanism that is operative in the Kitaev chain [86], the gap inversion creates a zero-mode at each end of the junction, which then propagates away from the junction along the edge mode. The edge modes are chiral, meaning that the motion is in a single direction only. For our purpose we need that the propagation is in the same direction along both edges connected by a Josephson junction. The geometry of Fig. 3.2 shows one way to achieve this using a topological insulator/magnetic insulator/superconductor heterostructure [94, 99]. (In Fig. 3.3 we show an alternative realization using a Chern insulator/superconductor heterostructure [27, 98].)

In the next section 3.2 we describe the way in which the fusion process shown schematically in Fig. 3.1 can be implemented in the structure of Figs. 3.2 and 3.3. In the subsequent sections 3.3 and 3.4 we present an

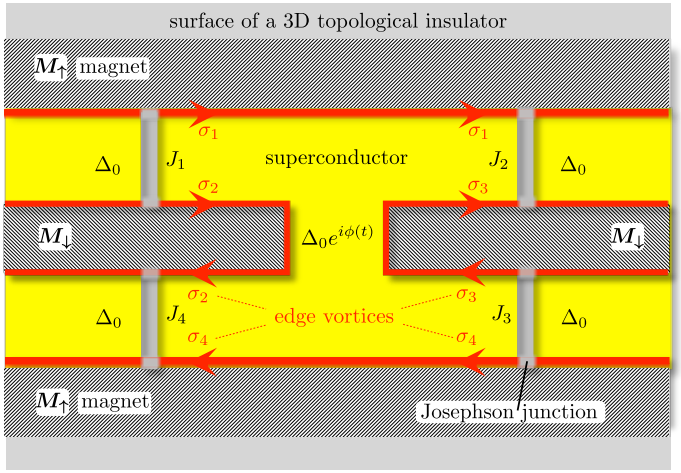

Figure 3.2: Geometry to create and fuse two pairs of edge vortices in a topological insulator/magnetic insulator/superconductor heterostructure. The edge vortices are created at Josephson junctions J_1 and J_3, by a 2π increment of the superconducting phase $\phi(t)$ on the central superconducting island. Each edge vortex contains a Majorana zero-mode and two zero-modes define a fermion parity qubit. The initial state $|J_1 J_3\rangle = |00\rangle$ has even-even fermion parity. When the edge vortices fuse at Josephson junctions J_2 and J_4 the final state $|J_2 J_4\rangle = (|00\rangle + i|11\rangle)/\sqrt{2}$ is in an equal-weight superposition of even-even and odd-odd parity states.

explicit calculation of the fermion parity of the final state, to demonstrate the equal-weight superposition of even and odd fermion parity implied by the Majorana fusion rule. Sec. 3.5 addresses an electrical signature of the fusion process: The sum $I_\mathrm{L} + I_\mathrm{R}$ of the currents at the two ends of the structure shows shot noise, because of the nondeterministic nature of the fusion process, but the difference $I_\mathrm{L} - I_\mathrm{R}$ is nearly noiseless, because of the correlated fermion parity. We conclude in Sec. 3.6.

3.2 Edge vortex injection and fusion in a four-terminal Josephson junction

The geometry of Fig. 3.2, with four incoming and four outgoing Majorana edge modes was introduced in Ref. [117] and studied recently in Refs. [33, 118, 119]. Those earlier works considered the injection of *fermions*:

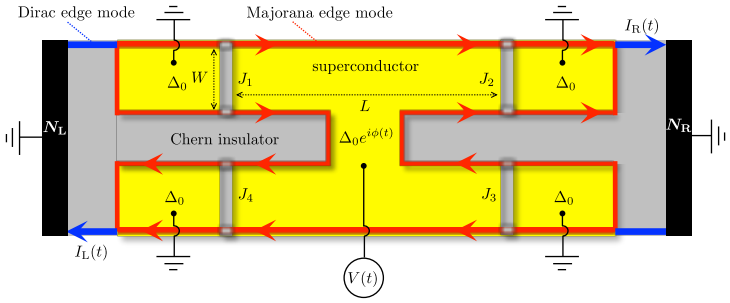

Figure 3.3: Same as Fig. 3.2, but now in a Chern insulator/superconductor heterostructure with normal metal contacts (N_L, N_R) to detect the charge produced upon fusion of the edge vortices. An integrated voltage pulse $\int V(t)dt = h/2e$ induces a 2π phase shift over the four Josephson junctions J_1, J_2, J_3, J_4, which results in a current pulse $I_\mathrm{L}(t)$, $I_\mathrm{R}(t)$ into the left and right contact. While I_L and I_R separately, as well as the sum $I_\mathrm{L} + I_\mathrm{R}$, exhibit shot noise, the difference $I_\mathrm{L} - I_\mathrm{R}$ becomes exactly noiseless for identical junctions J_1 and J_3.

electrons and holes injected into the Majorana edge modes from a normal metal contact. Here instead we consider the injection of *vortices*: π-phase domain walls injected into the edge modes by a Josephson junction. The injection happens in response to a voltage pulse $\int V(t)dt = h/2e$, which advances by 2π the phase $\phi(t)$ of the pair potential $\Delta_0 e^{i\phi}$. (Alternatively, an $h/2e$ flux bias achieves the same.) If the width W of the Josephson junction is large compared to the superconducting coherence length $\xi_0 = \hbar v_\mathrm{F}/\Delta_0$, the injection happens in a short time interval $t_\phi \simeq (\xi_0/W)\Delta t$ around $\phi(t) = \pi$, short compared the duration Δt of the voltage pulse [116].*

The edge vortices σ_n are anyons with a non-Abelian exchange statistics encoded in the Clifford algebra of Majorana operators γ_n,

$$\gamma_n \gamma_m + \gamma_m \gamma_n = \delta_{nm}. \tag{3.2.1}$$

Each edge vortex has a zero-mode and two zero-modes n, m encode a qubit degree of freedom in the fermion parity $P_{nm} = 2i\gamma_n\gamma_m$ with eigenvalues ± 1. Provided the vortices are non-overlapping, the qubit is protected from local sources of decoherence.

In the four-terminal Josephson junction of Fig. 3.2, one pair of edge vortices σ_1, σ_2 is injected at Josephson junction J_1 and a second pair

*This separation of time scales $t_\phi/\Delta t \simeq \xi_0/W \ll 1$ is why it is meaningful to distinguish the injection of vortices from the injection of fermions, since a Majorana fermion in an edge mode is equivalent to a pair of overlapping edge vortices.

σ_3, σ_4 is injected at Josephson junction J_3. Because the voltage pulse cannot create an unpaired fermion, the edge vortices are injected in a state $|\Psi\rangle$ of even fermion parity, $P_{12}|\Psi\rangle = |\Psi\rangle = P_{34}|\Psi\rangle$. Edge vortices σ_1 and σ_3 are fused at Josephson junction J_2 and vortices σ_2 and σ_4 are fused at junction J_4. The expectation value of the fermion parity upon fusion vanishes,

$$\begin{aligned} &\langle\Psi|P_{13}|\Psi\rangle = \langle\Psi|P_{12}P_{13}P_{12}|\Psi\rangle = -\langle\Psi|P_{13}P_{12}^2|\Psi\rangle = -\langle\Psi|P_{13}|\Psi\rangle \\ &\Rightarrow \langle\Psi|P_{13}|\Psi\rangle = 0, \end{aligned} \tag{3.2.2}$$

and similarly $\langle\Psi|P_{24}|\Psi\rangle = 0$. So the fusion of edge vortices at J_2 and J_3 leaves the edge modes in an equal weight superposition of odd and even fermion parity. This presence of multiple fusion channels is a defining property of non-Abelian anyons [3, 112, 113].

Because the overall fermion parity is conserved, the fusion outcomes at J_2 and J_3 must have the same fermion parity — either even-even or odd-odd. In the next two sections we present an explicit calculation of the fermion parity, to demonstrate that an $h/2e$ voltage pulse produces a superposition of even-even and odd-odd fermion parity states with identical probabilities P_{00} and $P_{11} = 1 - P_{00}$.

3.3 Scattering formula for the fermion parity

3.3.1 Construction of the fermion parity operator

We focus on the geometry of Fig. 3.3, with incoming and outgoing modes in the left lead (labeled L) and in the right lead (R). We seek the expectation value

$$\rho_\pi \equiv \left\langle e^{i\pi\mathcal{N}} \right\rangle = P_{00} - P_{11}, \tag{3.3.1}$$

of the fermion parity operator $e^{i\pi\mathcal{N}}$, with $\mathcal{N}$ the particle number operator of outgoing modes in one of the two leads. We will take the left lead for definiteness. In terms of the annilation operators $b_n(E)$ of outgoing modes n at excitation energy $E > 0$ this operator takes the form

$$\mathcal{N} = \sum_{n\in \mathrm{L}} \sum_{E>0} b_n^\dagger(E) b_n(E), \tag{3.3.2}$$

where we have discretized the energy. In the continuum limit $\sum_E \mapsto \int dE/2\pi$ and the Kronecker delta becomes a Dirac delta function, $\delta_{EE'} \mapsto 2\pi\delta(E - E')$.

Incoming and outgoing modes are related by a unitary scattering matrix,

$$b_n(E) = \sum_{m,E'} S_{nm}(E,E')a_m(E'), \tag{3.3.3}$$

$$\sum_{n'',E''} S^*_{n''n}(E'',E)S_{n''m}(E'',E') = \delta_{nm}\delta_{EE'}. \tag{3.3.4}$$

Note that the sums in these two equations run over positive and negative energies. Particle-hole symmetry relates

$$S_{nm}(-E,-E') = S^*_{nm}(E,E'). \tag{3.3.5}$$

We write Eq. (3.3.3) more compactly as $\boldsymbol{b} = \boldsymbol{S}\cdot\boldsymbol{a}$, collecting the mode and energy variables in vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. The unitarity relation (3.3.4) is then written as $\boldsymbol{S}^\dagger\boldsymbol{S} = 1$. In terms of a projection operator $\mathcal{P}_\mathrm{L}$ onto modes in lead L, and a projection operator $\mathcal{P}_+$ onto positive energies, the combination of Eqs. (3.3.2) and (3.3.3) reads

$$\mathcal{N} = \boldsymbol{a}^\dagger\cdot\boldsymbol{M}\cdot\boldsymbol{a}, \quad \boldsymbol{M} = \boldsymbol{S}^\dagger\mathcal{P}_\mathrm{L}\mathcal{P}_+\boldsymbol{S}. \tag{3.3.6}$$

The expectation value $\langle\cdots\rangle = \mathrm{Tr}\,(\rho_\mathrm{eq}\cdots)$ is with respect to an equilibrium distribution of the incoming modes,

$$\rho_\mathrm{eq} \propto \exp\left(-\beta\sum_n\sum_{E>0} Ea^\dagger_n(E)a_n(E)\right). \tag{3.3.7}$$

We denote $\beta = 1/k_\mathrm{B}T$ and have omitted the normalization constant (fixed by $\mathrm{Tr}\,\rho_\mathrm{eq} = 1$).

The combination of particle-hole symmetry,

$$a^\dagger_n(E) = a_n(-E), \tag{3.3.8}$$

with anticommutation,

$$\{a^\dagger_n(E), a_m(E')\} = \delta_{nm}\delta_{EE'}, \tag{3.3.9}$$

allows us to extend the sum $\sum_{E>0}$ in Eq. (3.3.7) to a sum over positive and negative energies,

$$\rho_\mathrm{eq} \propto \exp\left(-\tfrac{1}{2}\beta\sum_{n,E} Ea^\dagger_n(E)a_n(E)\right) \equiv e^{-\frac{1}{2}\beta\boldsymbol{a}^\dagger\cdot\boldsymbol{E}\cdot\boldsymbol{a}}. \tag{3.3.10}$$

In the second equation we introduced the diagonal operator $\boldsymbol{E}_{nm}(E,E') = E\delta_{nm}\delta_{EE'}$.

With this notation the average fermion parity is given by the ratio of two operator traces,

$$\rho_\pi = \frac{\mathrm{Tr}\left(e^{-\frac{1}{2}\beta \boldsymbol{a}^\dagger\cdot\boldsymbol{E}\cdot\boldsymbol{a}} e^{i\pi \boldsymbol{a}^\dagger\cdot\boldsymbol{M}\cdot\boldsymbol{a}}\right)}{\mathrm{Tr}\, e^{-\frac{1}{2}\beta \boldsymbol{a}^\dagger\cdot\boldsymbol{E}\cdot\boldsymbol{a}}}. \tag{3.3.11}$$

3.3.2 Klich formula for particle-hole conjugate Majorana operators

Fermionic operator traces of the form (3.3.11) have been studied by Klich and collaborators [120–122]. For Dirac fermion creation and annihilation operators $\boldsymbol{d}^\dagger, \boldsymbol{d}$ one has the simple expression [120]

$$\mathrm{Tr}\prod_k e^{\boldsymbol{d}^\dagger\cdot\boldsymbol{O}_k\cdot\boldsymbol{d}} = \mathrm{Det}\left(1+\prod_k e^{\boldsymbol{O}_k}\right). \tag{3.3.12}$$

The answer is different for self-conjugate Majorana operators $\boldsymbol{\gamma} = \boldsymbol{\gamma}^\dagger$, with anticommutator $\{\gamma_n, \gamma_m\} = \delta_{nm}$, when one has instead [122]

$$\left[\mathrm{Tr}\prod_k e^{\boldsymbol{\gamma}^\dagger\cdot\boldsymbol{O}_k\cdot\boldsymbol{\gamma}}\right]^2 = e^{\sum_k \mathrm{Tr}\,\boldsymbol{O}_k}\,\mathrm{Det}\left(1+\prod_k e^{\boldsymbol{O}_k-\boldsymbol{O}_k^{\mathrm{T}}}\right). \tag{3.3.13}$$

(The superscript T indicates the transpose of the matrix.)

The Majorana fermion modes in the topological superconductor are not self-conjugate, instead creation and annihilation operators $\boldsymbol{a}^\dagger, \boldsymbol{a}$ are related by the particle-hole symmetry relation (3.3.8). In view of Eq. (3.3.9) this implies that annihilation operators at energies $\pm E$ fail to anticommute:

$$\{a_n(E), a_m(-E')\} = \delta_{nm}\delta_{EE'}. \tag{3.3.14}$$

This unusual anticommutator expresses the Majorana nature of Bogoliubov quasiparticles [123].

To arrive at the analogue of Eq. (3.3.13) for particle-hole conjugate Majorana operators we rewrite the bilinear form $\boldsymbol{a}^\dagger\cdot\boldsymbol{O}\cdot\boldsymbol{a}$ such that the $\boldsymbol{a}, \boldsymbol{a}^\dagger$ operators appear only at positive energies:

$$\begin{aligned}\boldsymbol{a}^\dagger\cdot\boldsymbol{O}\cdot\boldsymbol{a} &= \sum_{n,m}\sum_{E,E'} a_n^\dagger(E) O_{nm}(E,E') a_m(E') \\ &= \sum_{n,m}\sum_{E,E'>0} \begin{pmatrix} a_n^\dagger(E) \\ a_n(E)\end{pmatrix} \mathcal{O}_{nm}(E,E') \begin{pmatrix} a_m(E') \\ a_m^\dagger(E')\end{pmatrix}.\end{aligned} \tag{3.3.15}$$

The matrix $\mathcal{O}$ imposes on $\boldsymbol{O}$ a 2×2 block structure,

$$\mathcal{O} = \begin{pmatrix} \boldsymbol{O}_{++} & \boldsymbol{O}_{+-} \\ \boldsymbol{O}_{-+} & \boldsymbol{O}_{--} \end{pmatrix}, \tag{3.3.16}$$

to encode the sign of the energy variables:

$$(\boldsymbol{O}_{ss'})_{nm}(E,' E') = O_{nm}(sE, s'E') \text{ for } s, s' \in \{+,-\} \text{ and } E, E' > 0. \tag{3.3.17}$$

We introduce the 2×2 Pauli matrix σ_x that acts on the block structure of $\mathcal{O}$ and define the generalized antisymmetrization

$$\begin{aligned} \mathcal{O}^{\mathrm{A}} &= \tfrac{1}{2}\mathcal{O} - \tfrac{1}{2}\sigma_x \mathcal{O}^{\mathrm{T}} \sigma_x \\ &= \frac{1}{2}\begin{pmatrix} \boldsymbol{O}_{++} - \boldsymbol{O}_{--}^{\mathrm{T}} & \boldsymbol{O}_{+-} - \boldsymbol{O}_{+-}^{\mathrm{T}} \\ \boldsymbol{O}_{-+} - \boldsymbol{O}_{-+}^{\mathrm{T}} & \boldsymbol{O}_{--} - \boldsymbol{O}_{++}^{\mathrm{T}} \end{pmatrix}. \end{aligned} \tag{3.3.18}$$

Only $\mathcal{O}^{\mathrm{A}}$ and $\operatorname{Tr}\mathcal{O} = \operatorname{Tr}\boldsymbol{O}$ contribute to the Majorana fermion operator trace,

$$\left[\operatorname{Tr}\prod_k e^{\boldsymbol{a}^\dagger \cdot \boldsymbol{O}_k \cdot \boldsymbol{a}}\right]^2 = e^{\sum_k \operatorname{Tr}\boldsymbol{O}_k} \operatorname{Det}\left(1 + \prod_k e^{2\mathcal{O}_k^{\mathrm{A}}}\right), \tag{3.3.19}$$

see App. 3.B. Eq. (3.3.19) is the desired analogue of Eq. (3.3.13) for particle-hole conjugate Majorana operators.

3.3.3 Fermion parity as the determinant of a scattering matrix product

For the average fermion parity ρ_π we apply Eq. (3.3.19) to the ratio of operator traces (3.3.11). We start from the block decomposition of $\boldsymbol{E}, \boldsymbol{S}$, and $\boldsymbol{M} = \boldsymbol{S}^\dagger \mathcal{P}_{\mathrm{L}} \mathcal{P}_+ \boldsymbol{S}$,

$$\begin{aligned} &\mathcal{E} = \begin{pmatrix} \boldsymbol{E} & 0 \\ 0 & -\boldsymbol{E} \end{pmatrix} = \boldsymbol{E}\sigma_z, \quad \mathcal{S} = \begin{pmatrix} \boldsymbol{S}_{++} & \boldsymbol{S}_{+-} \\ \boldsymbol{S}_{-+} & \boldsymbol{S}_{--} \end{pmatrix}, \\ &\mathcal{M} = \tfrac{1}{2}\mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}}(\sigma_0 + \sigma_z)\mathcal{S}. \end{aligned} \tag{3.3.20}$$

In the equation for $\mathcal{M}$ we substituted $\mathcal{P}_+ = \frac{1}{2}(\sigma_0 + \sigma_z)$, with σ_0 the 2×2 unit matrix.

The antisymmetrization of $\mathcal{E}$ is simple,

$$\mathcal{E}^{\mathrm{A}} \equiv \tfrac{1}{2}\mathcal{E} - \tfrac{1}{2}\sigma_x \mathcal{E}^{\mathrm{T}} \sigma_x = \boldsymbol{E}\sigma_z. \tag{3.3.21}$$

For the antisymmetrization of $\mathcal{M}$ we note that Eq. (3.3.5) implies $\sigma_x \mathcal{S} \sigma_x = \mathcal{S}^*$, hence

$$\sigma_x \mathcal{S}^{\mathrm{T}} \sigma_x = \mathcal{S}^\dagger \Rightarrow \mathcal{M}^{\mathrm{A}} = \tfrac{1}{2} \mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \sigma_z \mathcal{S}. \tag{3.3.22}$$

We thus arrive at

$$\rho_\pi^2 = e^{i\pi \operatorname{Tr} \boldsymbol{M}} \frac{\operatorname{Det}\left(1 + e^{-\beta \boldsymbol{E} \sigma_z} e^{i\pi \mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \sigma_z \mathcal{S}}\right)}{\operatorname{Det}\left(1 + e^{-\beta \boldsymbol{E} \sigma_z}\right)}. \tag{3.3.23}$$

The ratio of determinants is equivalent to a single determinant,

$$\begin{aligned} &\rho_\pi^2 = e^{i\pi \operatorname{Tr} \boldsymbol{M}} \operatorname{Det}\left(1 - \mathcal{F} + \mathcal{F} e^{i\pi \mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \sigma_z \mathcal{S}}\right), \\ &\mathcal{F} = (1 + e^{\beta \boldsymbol{E} \sigma_z})^{-1}, \;\; 1 - \mathcal{F} = (1 + e^{-\beta \boldsymbol{E} \sigma_z})^{-1}. \end{aligned} \tag{3.3.24}$$

To proceed we first rewrite the exponent of the trace of $\boldsymbol{M}$ as a determinant,

$$e^{i\pi \operatorname{Tr} \boldsymbol{M}} = e^{i\pi \operatorname{Tr} \mathcal{P}_{\mathrm{L}} \mathcal{P}_+} \tag{3.3.25a}$$

$$= \operatorname{Det}\left[-\sigma_z\right]_{\mathrm{LL}} = \operatorname{Det}\left[\sigma_z\right]_{\mathrm{LL}} \;\text{ with }\; \sigma_z \equiv 2\mathcal{P}_+ - 1, \tag{3.3.25b}$$

$$= \operatorname{Det}\left[-\tau_z\right]_{++} = \operatorname{Det}\left[\tau_z\right]_{++} \;\text{ with }\; \tau_z \equiv 2\mathcal{P}_{\mathrm{L}} - 1. \tag{3.3.25c}$$

The notation $[\cdots]_{\mathrm{LL}}$ indicates a projection onto mode indices in the left lead, and $[\cdots]_{++}$ indicates a projection onto positive energies.

We then evaluate the exponent of the scattering matrix product,

$$\begin{aligned} &e^{i\xi \mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \sigma_z \mathcal{S}} = \sigma_0 + i(\sin\xi) \mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \sigma_z \mathcal{S} + (\cos\xi - 1) \mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \mathcal{S}, \\ &\Rightarrow e^{i\pi \mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \sigma_z \mathcal{S}} = \sigma_0 - 2\mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \mathcal{S}, \end{aligned} \tag{3.3.26}$$

since $(\mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \sigma_z \mathcal{S})^{2n} = \mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \mathcal{S}$ and $(\mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \sigma_z \mathcal{S})^{2n-1} = \mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \sigma_z \mathcal{S}$, for $n = 1, 2, 3, \ldots$. It follows that

$$\rho_\pi^2 = e^{i\pi \operatorname{Tr} \boldsymbol{M}} \operatorname{Det}\left(1 - 2\mathcal{F} \mathcal{S}^\dagger \mathcal{P}_{\mathrm{L}} \mathcal{S}\right) \tag{3.3.27a}$$

$$= e^{i\pi \operatorname{Tr} \boldsymbol{M}} \operatorname{Det}\left(1 - 2\mathcal{P}_{\mathrm{L}} \mathcal{S} \mathcal{F} \mathcal{S}^\dagger\right) \tag{3.3.27b}$$

$$= e^{i\pi \operatorname{Tr} \boldsymbol{M}} \operatorname{Det}\left[1 - 2\mathcal{S} \mathcal{F} \mathcal{S}^\dagger\right]_{\mathrm{LL}} \tag{3.3.27c}$$

$$= \operatorname{Det}\left[\sigma_z\right]_{\mathrm{LL}} \operatorname{Det}\left[\mathcal{S}(1 - 2\mathcal{F}) \mathcal{S}^\dagger\right]_{\mathrm{LL}} \tag{3.3.27d}$$

$$= \operatorname{Det}\left[\sigma_z \mathcal{S} \tanh(\tfrac{1}{2}\beta \mathcal{E}) \mathcal{S}^\dagger\right]_{\mathrm{LL}}. \tag{3.3.27e}$$

In Eq. (3.3.27b) we used the Sylvester identity $\operatorname{Det}(1 - AB) = \operatorname{Det}(1 - BA)$, in Eq. (3.3.27c) we used $\operatorname{Det}(1 - \mathcal{P}_{\mathrm{L}} A) = \operatorname{Det}[1 - A]_{\mathrm{LL}}$, in Eq. (3.3.27d)

we used $\mathcal{S}\mathcal{S}^\dagger = 1$, and in (3.3.27e) we used that $\mathrm{Det}\,[A]_{\mathrm{LL}}\mathrm{Det}\,[B]_{\mathrm{LL}} = \mathrm{Det}\,[AB]_{\mathrm{LL}}$ if A or B commutes with $\mathcal{P}_{\mathrm{L}}$.

In what follows we restrict ourselves to zero temperature, when $\mathcal{F} \mapsto \mathcal{P}_-$ projects onto negative energies and $\tanh(\frac{1}{2}\beta\mathcal{E}) \mapsto \sigma_z$. Eq. (3.3.27e) then reduces to

$$\rho_\pi^2 = \mathrm{Det}\left[\sigma_z \mathcal{S} \sigma_z \mathcal{S}^\dagger\right]_{\mathrm{LL}}, \tag{3.3.28}$$

the determinant of a scattering matrix product projected onto mode indices in the left lead. An alternative projection onto positive energies is possible:

$$\begin{aligned}
\rho_\pi^2 &= e^{i\pi\,\mathrm{Tr}\,\boldsymbol{M}}\mathrm{Det}\left(1 - 2\mathcal{P}_-\mathcal{S}^\dagger\mathcal{P}_{\mathrm{L}}\mathcal{S}\right) && \text{(3.3.29a)}\\
&= e^{i\pi\,\mathrm{Tr}\,\boldsymbol{M}}\mathrm{Det}\left(1 - 2\mathcal{P}_+\mathcal{S}^\dagger\mathcal{P}_{\mathrm{L}}\mathcal{S}\right) && \text{(3.3.29b)}\\
&= \mathrm{Det}\,[-\tau_z]_{++}\,\mathrm{Det}\big[\mathcal{S}^\dagger(1 - 2\mathcal{P}_{\mathrm{L}})\mathcal{S}\big]_{++}, && \text{(3.3.29c)}
\end{aligned}$$

(In Eq. (3.3.29b) we used particle-hole symmetry, $\mathcal{S} = \sigma_x\mathcal{S}^*\sigma_x$, and $\sigma_x\mathcal{P}_-\sigma_x = \mathcal{P}_+$.) Because τ_z commutes with $\mathcal{P}_+$, Eq. (3.3.29c) may be combined into a a single determinant,

$$\rho_\pi^2 = \mathrm{Det}\big[\tau_z\mathcal{S}^\dagger\tau_z\mathcal{S}\big]_{++}. \tag{3.3.30}$$

Equations (3.3.28) and (3.3.30) express the average fermion parity of a scattering state as the determinant of a product of scattering matrices projected onto a submatrix in mode space, Eq. (3.3.28), or in energy space, Eq. (3.3.30).* Both equations give the square ρ_π^2 rather than ρ_π itself. Since we wish to show that $\rho_\pi = 0$, that is not a limitation for the present study.

3.3.4 Simplification in the adiabatic regime

The energy dependence of the scattering matrix is characterized by the inverse of two time scales of the Josephson junction: the dwell time $\tau_{\mathrm{dwell}} \simeq L/v_{\mathrm{F}}$ in the superconducting island and the characteristic time scale

$$t_\phi = (\xi_0/W)(d\phi/dt)^{-1} \tag{3.3.31}$$

for the variation of the superconducting phase shift. (The time t_ϕ is the "vortex injection time" t_{inj} of Ref. [116].) While $S(E, E')$ depends on the

*To avoid a possible confusion we note that, because of the projection, the product rule $\mathrm{Det}\,(AB) = (\mathrm{Det}\,A)(\mathrm{Det}\,B)$ cannot be applied to $\mathrm{Det}[AB]_{++}$ or $\mathrm{Det}[AB]_{\mathrm{LL}}$, unless A or B commutes with the projector.

average energy $\bar{E} = (E + E')/2$ on the scale $1/\tau_{\rm dwell}$, it depends on the energy difference $\delta E = E - E'$ on the scale $1/\tau_\phi$.

In the adiabatic regime $\tau_{\rm dwell} \ll \tau_\phi$ the scattering matrix $S(E, E')$ for $\bar{E} \lesssim 1/\tau_\phi \ll 1/\tau_{\rm dwell}$ is only a function of δE,

$$S(E, E') = \int_{-\infty}^{\infty} dt\, e^{i(E-E')t} S_{\rm F}(t) + \mathcal{O}(\tau_{\rm dwell}/\tau_\phi). \tag{3.3.32}$$

The unitary matrix $S_{\rm F}(t)$ is the "frozen" scattering matrix at the Fermi level, calculated for a fixed value $\phi \equiv \phi(t)$ of the superconducting phase.

The fermion parity determinant can be simplified in the adiabatic regime, because only energies within $1/\tau_\phi$ from the Fermi level contribute. This is most easily seen from Eq. (3.3.28), which is the determinant of the scattering matrix product $\Omega = \sigma_z \mathcal{S} \sigma_z \mathcal{S}^\dagger$, projected onto the left lead. A matrix element of Ω,

$$\Omega_{nm}(E, E') = (\operatorname{sign} E) \sum_{n', E''} (\operatorname{sign} E'') S_{nn'}(E, E'') S^*_{mn'}(E', E'') \tag{3.3.33}$$

is only nonzero for $|E - E'| \lesssim 1/\tau_\phi$. Moreover, $\Omega_{nm}(E, E') \approx \delta_{nm}\delta_{EE'}$ for $|E| \gtrsim 1/\tau_\phi$. Hence the determinant of Ω is fully determined by energies in the range $-1/\tau_\phi \lesssim E, E' \lesssim 1/\tau_\phi$, where $S(E, E')$ may be approximated by the frozen scattering matrix (3.3.32).

For computational purposes it is more convenient to rewrite the determinant (3.3.28) in the form (3.3.30), because the scattering matrix product $\tau_z \mathcal{S} \tau_z \mathcal{S}^\dagger$ is a convolution in energy space when $S(E, E')$ is a function of $E - E'$. The convolution is readily evaluated in the time domain, resulting in an expression for the fermion parity

$$\rho_\pi^2 = \operatorname{Det} [Q]_{++}, \tag{3.3.34}$$

in terms of the determinant of the projection onto $E, E' > 0$ of the matrix

$$Q(E, E') = \int_{-\infty}^{\infty} dt\, e^{i(E-E')t} Q(t), \quad Q(t) = \tau_z S_{\rm F}^\dagger(t) \tau_z S_{\rm F}(t). \tag{3.3.35}$$

In the next section we shall show how to evaluate this determinant.

3.4 Vanishing of the average fermion parity

We apply the formalism that we developed in Sec. 3.3 to the four-terminal Josephson junction of Sec. 3.2, in order to demonstrate that the 2π phase

shift produces a state with an equal weight $P_{00} = P_{11}$ of even-even and odd-odd fermion parity in the left and right leads. We work in the adiabatic regime, when $\rho_\pi = P_{00} - P_{11}$ is given by Eqs. 3.3.34 and (3.3.35) in terms of the "frozen" scattering matrix $S_\mathrm{F}(t)$, for a fixed phase $\phi(t)$.

3.4.1 Frozen scattering matrix of the Josephson junction

The frozen scattering matrix $S_\mathrm{F} \in \mathrm{SO}(4)$ is calculated in App. 3.A, resulting in

$$S_\mathrm{F} = \begin{pmatrix} e^{-i\alpha_4\nu_y} & 0 \\ 0 & e^{-i\alpha_2\nu_y} \end{pmatrix} \cdot \Pi \cdot \begin{pmatrix} e^{i\alpha_1\nu_y} & 0 \\ 0 & e^{i\alpha_3\nu_y} \end{pmatrix}, \quad \Pi = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.4.1}$$

The Pauli matrix ν_y acts on the two Majorana modes in each lead. The scattering phase α_n depends on the superconducting phase difference ϕ through the relation [116]

$$\alpha_n = \arccos\left(\frac{\cos(\phi/2) + \tanh\beta_n}{1 + \cos(\phi/2)\tanh\beta_n}\right) \times \mathrm{sign}\,(\phi), \quad \beta_n = \frac{W_n}{\xi_0}\cos(\phi/2). \tag{3.4.2}$$

A 2π increment of ϕ corresponds to a π increment of α_n, irrespective of the width W_n of the Josephson junction or the superconducting coherence length $\xi_0 = \hbar v_\mathrm{F}/\Delta_0$.

We need to evaluate the matrix product $\tau_z S_\mathrm{F}^\dagger \tau_z S_\mathrm{F}$, where the Pauli matrix

$$\tau_z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.4.3}$$

is defined with respect to the block structure of modes in the left and right lead. Because of the identity

$$\Pi\tau_z\Pi = \begin{pmatrix} \nu_z & 0 \\ 0 & \nu_z \end{pmatrix}, \tag{3.4.4}$$

this matrix product is block-diagonal,

$$Q(t) = \tau_z S_\mathrm{F}^\dagger(t)\tau_z S_\mathrm{F}(t) = -\begin{pmatrix} \nu_z e^{2i\nu_y\alpha_1(t)} & 0 \\ 0 & \nu_z e^{2i\nu_y\alpha_3(t)} \end{pmatrix}, \tag{3.4.5}$$

independent of α_2 and α_4.

3.4.2 Reduction of the fermion parity to a Toeplitz determinant

Instead of taking a single 2π phase increment it is more convenient to assume a sequence of 2π phase shifts with period Δt. Then $\alpha_n(t)$ varies periodically in time with $\alpha_n(t+\Delta t) = \pi + \alpha_n(t)$. We Fourier transform to the energy domain,

$$\begin{aligned} \boldsymbol{T}_n(k,k') &= \frac{1}{\Delta t}\int_0^{\Delta t} dt\, e^{2\pi i(k-k')t/\Delta t} e^{2i\alpha_n(t)\nu_y}, \\ T_n(k,k') &= \frac{1}{\Delta t}\int_0^{\Delta t} dt\, e^{2\pi i(k-k')t/\Delta t} e^{2i\alpha_n(t)}, \end{aligned} \tag{3.4.6}$$

and restrict $k, k' \in \{1, 2, 3, \ldots\}$ to positive integers. The infinite matrix $T_n(k,k')$ has constant diagonals, so it is a Toeplitz matrix. Eq. (3.3.30) becomes the product of Toeplitz determinants,

$$\rho_\pi^2 = (\mathrm{Det}\,\boldsymbol{T}_1)(\mathrm{Det}\,\boldsymbol{T}_3) = |\mathrm{Det}\,T_1|^2\,|\mathrm{Det}\,T_3|^2. \tag{3.4.7}$$

The Toeplitz matrices T_n are banded matrices which extend over a large number of order W/ξ_0 of diagonals around the main diagonal. This follows from the fact that the π increment of $\alpha(t)$ happens in the time interval $t_\phi = (\xi_0/W)(\Delta t/2\pi)$ which is much shorter than Δt for $\xi_0 \ll W$. The ratio $t_\phi/\Delta t$ governs the exponential decay of the Toeplitz matrix elements as one moves away from the main diagonal, according to

$$|T_n(k,k')| \simeq \exp(-c_{\mathrm{decay}}|k-k'|), \quad c_{\mathrm{decay}} = \frac{\pi^2 t_\phi}{\Delta t} = \frac{\pi\xi_0}{2W}. \tag{3.4.8}$$

3.4.3 Fisher-Hartwig asymptotics

In a general formulation, the function $b(\theta)$ defines the $K \times K$ Toeplitz matrix

$$B_K(k,k') = \int_0^{2\pi} e^{i(k-k')\theta} b(\theta)\,\frac{d\theta}{2\pi}, \quad k,k' \in \{1,2,\ldots K\}. \tag{3.4.9}$$

If b is smooth and nonvanishing on the unit circle $0 < \theta < 2\pi$, it has a well-defined winding number

$$\nu = \frac{1}{2\pi i}\int_0^{2\pi} \frac{b'(\theta)}{b(\theta)}\,d\theta. \tag{3.4.10}$$

The number ν may be non-integer, or even complex, if b has a jump discontinuity at $\theta = 0$.

The Fisher-Hartwig asymptotics [124, 125] determines the large-K limit of the determinant of B_K from the decomposition $b(\theta) = b_0(\theta)e^{i\nu\theta}$, where b_0 has zero winding number. In the most general case the function b_0 may have (integrable) singularities, but if we assume it is smooth the asymptotics reads

$$\begin{aligned} \operatorname{Det} B_K \simeq & \exp\left(\frac{K}{2\pi}\int_0^{2\pi} \ln b_0(\theta)\, d\theta\right) \\ & \times \begin{cases} K^{-\nu^2} & \text{for non-integer } \nu, \\ e^{-|\nu| c_{\mathrm{decay}} K} & \text{for integer } \nu. \end{cases} \end{aligned} \tag{3.4.11}$$

The coefficient c_{decay} in the exponent is the decay rate $|B_K(k,k')| \simeq \exp(-c_{\mathrm{decay}}|k-k'|)$ of the Toeplitz matrix elements as we move away from the diagonal.

Applied to $b(t) = e^{2i\alpha(t)}$, $\theta = 2\pi t/\Delta t$, we have $\nu = 1$, $b_0(t) = e^{2i\alpha(t)-2\pi i t/\Delta t}$. The Toeplitz determinant

$$\operatorname{Det} B_K \simeq e^{-c_{\mathrm{decay}} K} \exp\left(\frac{2iK}{\Delta t}\int_0^{\Delta t} \alpha(t)dt - i\pi K\right) \tag{3.4.12}$$

vanishes exponentially in the limit $K \to \infty$, with decay rate $c_{\mathrm{decay}} = \pi\xi_0/W$ determined by the ratio of the superconducting coherence length ξ_0 and the width W of the Josephson junction.

For the evaluation of the fermion parity, the band width $K/\Delta t$ is limited by the energy range $|\bar{E}| \lesssim 1/t_{\mathrm{dwell}}$ where the dependence of the scattering matrix $S(E,E')$ on the average energy $\bar{E} = (E+E')/2$ may be neglected. We thus conclude that

$$\begin{aligned} |\rho_\pi| \simeq \exp(-2c_{\mathrm{decay}}K) &\simeq \exp\left(-\frac{2\pi\xi_0}{W}\frac{\Delta t}{t_{\mathrm{dwell}}}\right) \\ &\simeq \exp\left(-\frac{4\pi^2 t_\phi}{t_{\mathrm{dwell}}}\right), \end{aligned} \tag{3.4.13}$$

which is exponentially small in the adiabatic regime $t_\phi \gg t_{\mathrm{dwell}}$.

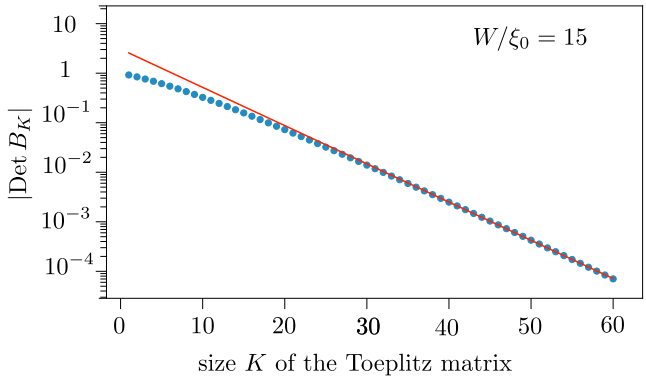

Figure 3.4: Decay of the Toeplitz determinant (data points), compared with the exponential decay expected from Eq. (3.4.12). The constant c_{decay} was calculated separately from $|B_K(k,k')| \simeq \exp(-c_{\mathrm{decay}}|k-k'|)$. The estimate $c_{\mathrm{decay}} = \pi\xi_0/W$ is off by 15%.

3.5 Transferred charge

3.5.1 Average charge

The average charge $\langle Q_{\mathrm{L}}\rangle$, $\langle Q_{\mathrm{R}}\rangle$ transferred into the left or right lead during one 2π increment of ϕ is given, in the adiabatic regime, by the superconducting analogue of Brouwer's formula [106, 107]:

$$\begin{aligned}\langle Q_{\mathrm{L}}\rangle &= \frac{ie}{4\pi}\int_{-\infty}^{\infty} dt\,\mathrm{Tr}\, S_{\mathrm{F}}^{\dagger}(t)\begin{pmatrix}\nu_y & 0\\ 0 & 0\end{pmatrix}\frac{\partial}{\partial t}S_{\mathrm{F}}(t),\\ \langle Q_{\mathrm{R}}\rangle &= \frac{ie}{4\pi}\int_{-\infty}^{\infty} dt\,\mathrm{Tr}\, S_{\mathrm{F}}^{\dagger}(t)\begin{pmatrix}0 & 0\\ 0 & \nu_y\end{pmatrix}\frac{\partial}{\partial t}S_{\mathrm{F}}(t).\end{aligned} \tag{3.5.1}$$

Substitution of Eq. (3.4.1) gives

$$\begin{aligned}\langle Q_{\mathrm{L}}\rangle &= \frac{e}{2\pi}\int_{-\infty}^{\infty} dt\,\frac{d}{dt}\alpha_4(t),\\ \langle Q_{\mathrm{R}}\rangle &= \frac{e}{2\pi}\int_{-\infty}^{\infty} dt\,\frac{d}{dt}\alpha_2(t).\end{aligned} \tag{3.5.2}$$

Because both α_2 and α_4 increase by π when ϕ is incremented by 2π, see Eq. (3.4.2), we conclude that

$$\langle Q_{\mathrm{L}}\rangle = \langle Q_{\mathrm{R}}\rangle = \frac{e}{2}. \tag{3.5.3}$$

While the average transferred charge per cycle is exactly $e/2$, the average particle number is close to but not exactly equal to $1/2$ — indicating that there is a small contribution from charge-neutral particle-hole pairs.*

3.5.2 Charge correlations

Fluctuations in the transferred charge are described by the second moments $\langle Q_{\rm L}^2\rangle$, $\langle Q_{\rm R}^2\rangle$, and $\langle Q_{\rm L}Q_{\rm R}\rangle$. Scattering matrix formulas for these correlators are derived in App. 3.C. In the adiabatic regime one has

$$\mathrm{var}(Q_{\rm L}) \equiv \langle Q_{\rm L}^2\rangle - \langle Q_{\rm L}\rangle^2 = \frac{e^2}{8\pi^2}\int_{0^+}^{\infty} d\omega\,\omega\,\mathrm{Tr}\,\Sigma_{\rm L}^\dagger(\omega)\Sigma_{\rm L}(\omega), \tag{3.5.4a}$$

$$\mathrm{var}(Q_{\rm R}) \equiv \langle Q_{\rm R}^2\rangle - \langle Q_{\rm R}\rangle^2 = \frac{e^2}{8\pi^2}\int_{0^+}^{\infty} d\omega\,\omega\,\mathrm{Tr}\,\Sigma_{\rm R}^\dagger(\omega)\Sigma_{\rm R}(\omega), \tag{3.5.4b}$$

$$\begin{aligned}\mathrm{covar}(Q_{\rm L}Q_{\rm R}) &\equiv \tfrac{1}{2}\langle Q_{\rm L}Q_{\rm R}\rangle + \tfrac{1}{2}\langle Q_{\rm R}Q_{\rm L}\rangle - \langle Q_{\rm L}\rangle\langle Q_{\rm R}\rangle \\ &= \frac{e^2}{16\pi^2}\int_{0^+}^{\infty} d\omega\,\omega\,\mathrm{Tr}\,\left[\Sigma_{\rm L}^\dagger(\omega)\Sigma_{\rm R}(\omega) + \Sigma_{\rm R}^\dagger(\omega)\Sigma_{\rm L}(\omega)\right],\end{aligned} \tag{3.5.4c}$$

in terms of the matrices

$$\Sigma_{\rm L}(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\omega t}\,\Sigma_{\rm L}(t),\quad \Sigma_{\rm L}(t) = S_{\rm F}^\dagger(t)\begin{pmatrix}\nu_y & 0\\ 0 & 0\end{pmatrix} S_{\rm F}(t), \tag{3.5.5a}$$

$$\Sigma_{\rm R}(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\omega t}\,\Sigma_{\rm R}(t),\quad \Sigma_{\rm R}(t) = S_{\rm F}^\dagger(t)\begin{pmatrix}0 & 0\\ 0 & \nu_y\end{pmatrix} S_{\rm F}(t). \tag{3.5.5b}$$

The lower limit 0^+ in the ω-integrals (3.5.4) avoids a spurious contribution $\propto \delta(\omega)$.

From the expression (3.4.1) for $S_{\rm F}(t)$ we find

$$\begin{aligned}\mathrm{Tr}\,\Sigma_{\rm L}^\dagger(\omega)\Sigma_{\rm L}(\omega) &= \mathrm{Tr}\,\Sigma_{\rm R}^\dagger(\omega)\Sigma_{\rm R}(\omega)\\ &= \tfrac{1}{2}|Z_+(\omega)|^2 + \tfrac{1}{2}|Z_+(-\omega)|^2 + \tfrac{1}{2}|Z_-(\omega)|^2 + \tfrac{1}{2}|Z_-(-\omega)|^2,\end{aligned} \tag{3.5.6a}$$

$$\begin{aligned}\mathrm{Tr}\,\Sigma_{\rm L}^\dagger(\omega)\Sigma_{\rm R}(\omega) &= \mathrm{Tr}\,\Sigma_{\rm R}^\dagger(\omega)\Sigma_{\rm L}(\omega)\\ &= \tfrac{1}{2}|Z_+(\omega)|^2 + \tfrac{1}{2}|Z_+(-\omega)|^2 - \tfrac{1}{2}|Z_-(\omega)|^2 - \tfrac{1}{2}|Z_-(-\omega)|^2,\end{aligned} \tag{3.5.6b}$$

$$Z_\pm(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\omega t} e^{i\alpha_1(t)\pm i\alpha_3(t)}. \tag{3.5.6c}$$

*A calculation along the lines of Ref. [116] of the average number of quasiparticles transferred per cycle into the left or the right lead gives $\langle N_{\rm L}\rangle = \langle N_{\rm R}\rangle = 42\zeta(3)/\pi^4 = 0.518$.

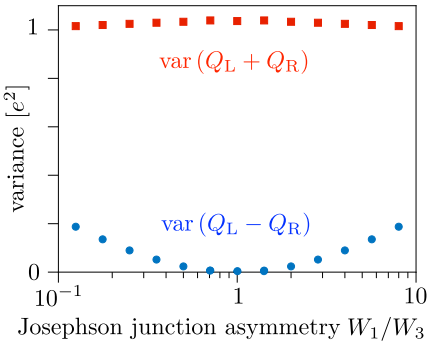

Figure 3.5: Variance of the sum and difference of the transferred charges upon fusion of the edge vortices in Josephson junctions J_2 and J_4, as a function of the asymmetry in the width of the injecting Josephson junctions J_1 and J_3.

The dependence on α_2 and α_4 drops out.

Without further calculation we see that for $\alpha_1 = \alpha_3$ the contribution of $Z_-(\omega)$ to the correlators (3.5.4) vanishes, hence $\mathrm{covar}(Q_\mathrm{L}Q_\mathrm{R}) = \mathrm{var}(Q_\mathrm{L}) = \mathrm{var}(Q_\mathrm{R})$. This implies that the charge difference $Q_\mathrm{L} - Q_\mathrm{R}$ is zero without fluctuations,

$$\mathrm{var}\,(Q_\mathrm{L} - Q_\mathrm{R}) = \mathrm{var}\,(Q_\mathrm{L}) + \mathrm{var}\,(Q_\mathrm{R}) - 2\,\mathrm{covar}(Q_\mathrm{L}Q_\mathrm{R}) = 0. \tag{3.5.7}$$

The charges Q_L and Q_R do fluctuate individually, with a variance close to $e^2/4$, and so does the sum $Q_\mathrm{L} + Q_\mathrm{R}$, with a variance close to e^2. These values can be calculated precisely for the time dependence [116]

$$\alpha(t) \approx \arccos\left[\tanh\left(\frac{W}{\xi_0}\frac{\pi - \phi(t)}{2}\right)\right] \approx \arccos[-\tanh(t/2t_\phi)], \tag{3.5.8}$$

which is an accurate representation of Eq. (3.4.2) for $W/\xi_0 \gg 1$. We find

$$Z_+(\omega) = 2\pi\delta(\omega) - \frac{8\pi\omega t_\phi^2}{\sinh(\pi\omega t_\phi)} + \frac{8\pi\omega t_\phi^2}{\cosh(\pi\omega t_\phi)}, \quad Z_-(\omega) = 2\pi\delta(\omega), \tag{3.5.9}$$

$$\Rightarrow \mathrm{var}\,(Q_\mathrm{L}) = \mathrm{var}\,(Q_\mathrm{R}) = \tfrac{1}{4}\mathrm{var}\,(Q_\mathrm{L} + Q_\mathrm{R}) = \frac{21\zeta(3)}{\pi^4}\,e^2 = 0.259\,e^2. \tag{3.5.10}$$

For $\alpha_1 \neq \alpha_3$ we can evaluate the integrals numerically using the time dependence

$$\alpha_n = \arccos\left[-\tanh(t/2t_n)\right], \tag{3.5.11}$$

increasing from 0 to π in a time $t_n = (\xi_0/W_n)(\Delta t/2\pi)$ around $t = 0$. Results for var $(Q_\mathrm{L} \pm Q_\mathrm{R})$ are shown in Fig. 3.5. The shot noise for the charge difference remains suppressed for a moderately large deviation from unity of W_1/W_3.

3.6 Conclusion

We have shown how the method of time-resolved and "on-demand" injection of edge vortices proposed in Ref. [116] can be used to demonstrate the non-Abelian fusion rule of Majorana zero-modes. The signature of the correlated but non-deterministic outcome of the fusion of two pairs of edge vortices is a fluctuating electrical current I_L and I_R through two Josephson junctions, induced by a 2π phase shift of the pair potential. While the sum $I_\mathrm{L} + I_\mathrm{R}$ has average e per cycle and variance close to e^2, the difference $I_\mathrm{L} - I_\mathrm{R}$ vanishes without fluctuations in a symmetric structure (and remains much below e^2 for moderate asymmetries).

The four-terminal structure of chiral Majorana edge modes that we have studied has been investigated before in the context of the injection of fermions [33, 117–119]. A Majorana fermion that splits into partial waves at opposite edges defines a nonlocally encoded *charge qubit*: a coherent superposition of an electron and a hole.* In contrast, the injection of vortices at opposite edges is a nonlocal encoding of the *fermion parity*. The difference could be significant for quantum information processing if the fermion parity qubit is more robust against decoherence than the charge qubit. We surmise that zero-modes in edge vortices are better protected against charge noise and other local sources of decoherence than Majorana fermions — basically because a Majorana fermion is charge neutral on average but does exhibit quantum fluctuations of the charge.

Much further research is needed to substantiate the potential of edge vortices as carriers of quantum information, but we feel that they have much to offer at least for the demonstration of basic operations in topological quantum computation: the braiding operation of Ref. [116] and the non-deterministic fusion operation considered here.

*The splitting of a Majorana fermion into partial waves does not provide a local encoding of the fermion parity because a measurement at one edge can detect the presence or absence of a fermion.

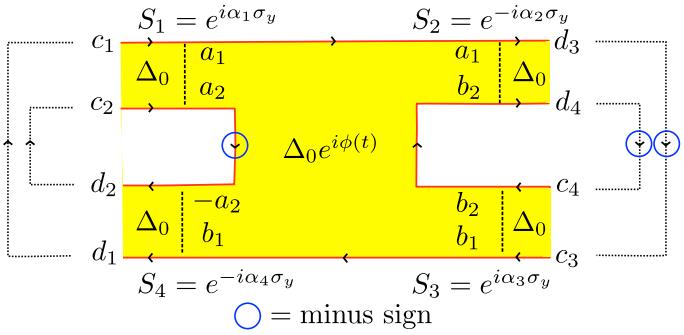

Figure 3.6: Labeling of incoming and outgoing Majorana edge modes in a four-terminal Josephson junction.

3.A Calculation of the frozen scattering matrix

Consider first the stationary scattering problem, when the four-terminal Josephson junction from Fig. 3.3 has a time-independent phase difference ϕ. This gives the "frozen" scattering matrix $S_{\mathrm{F}}(E, \phi)$, which we evaluate at the Fermi level ($E = 0$).

As calculated in Ref. [116], each of the four terminals (width W_n) has at the Fermi level a scattering matrix in SO(2) given by

$$\begin{aligned} S_n &= \begin{pmatrix} \cos\alpha_n & \sin\alpha_n \\ -\sin\alpha_n & \cos\alpha_n \end{pmatrix} = e^{i\alpha_n\nu_y} \quad \text{for} \quad n = 1, 3, \\ S_n &= \begin{pmatrix} \cos\alpha_n & -\sin\alpha_n \\ \sin\alpha_n & \cos\alpha_n \end{pmatrix} = e^{-i\alpha_n\nu_y} \quad \text{for} \quad n = 2, 4. \end{aligned} \tag{3.A.1}$$

The Pauli matrix ν_y acts on the two Majorana modes at a Josephson junction. The angles α_n are given as a function of ϕ and the ratio W_n/ξ_0 by Eq. (3.4.2) from the main text.

Referring to the labeling of modes from Fig. 3.6, we have the linear

relations

$$\begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix} = S_{\mathrm{F}} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix}, \tag{3.A.2a}$$

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = S_1 \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \quad \begin{pmatrix} d_3 \\ d_4 \end{pmatrix} = S_2 \begin{pmatrix} a_1 \\ b_2 \end{pmatrix},$$

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = S_3 \begin{pmatrix} c_3 \\ c_4 \end{pmatrix}, \quad \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = S_4 \begin{pmatrix} b_1 \\ -a_2 \end{pmatrix}. \tag{3.A.2b}$$

The minus sign for the coefficient a_2 in the last equality accounts for the π Berry phase of a circulating Majorana edge mode. As indicated by the dotted lines in Fig. 3.6, the edge modes are segments of three closed loops. We choose a gauge where the minus sign in each loop is acquired on the downward branch, indicated by the blue circle. This only affects the branch with amplitude a_2, because the other two downward branches are outside of the scattering region.

Elimination of the a_n and b_n variables gives

$$S_{\mathrm{F}} = \begin{pmatrix} -\sin\alpha_1 \sin\alpha_4 & \cos\alpha_1 \sin\alpha_4 & \cos\alpha_3 \cos\alpha_4 & \cos\alpha_4 \sin\alpha_3 \\ \cos\alpha_4 \sin\alpha_1 & -\cos\alpha_1 \cos\alpha_4 & \cos\alpha_3 \sin\alpha_4 & \sin\alpha_3 \sin\alpha_4 \\ \cos\alpha_1 \cos\alpha_2 & \cos\alpha_2 \sin\alpha_1 & \sin\alpha_2 \sin\alpha_3 & -\cos\alpha_3 \sin\alpha_2 \\ \cos\alpha_1 \sin\alpha_2 & \sin\alpha_1 \sin\alpha_2 & -\cos\alpha_2 \sin\alpha_3 & \cos\alpha_2 \cos\alpha_3, \end{pmatrix}, \tag{3.A.3}$$

which may be written more compactly as Eq. (3.4.1). One can check that $S_{\mathrm{F}} \in \mathrm{SO}(4)$, in particular, it has determinant $+1$ as it should be in the absence of a Majorana zero-mode [126].*

In the adiabatic regime the scattering matrix $S(E, E')$ of the time-dependent problem is related to the frozen scattering matrix $S_{\mathrm{F}}(E, \phi)$ via

$$S(E + \tfrac{1}{2}\omega, E - \tfrac{1}{2}\omega) \approx \int_{-\infty}^{\infty} dt\, e^{i\omega t} S_{\mathrm{F}}(E, \phi(t)). \tag{3.A.4}$$

Near the Fermi level we may furthermore neglect the dependence on the average energy, approximating

$$S(E, E') \approx \int_{-\infty}^{\infty} dt\, e^{i(E-E')t} S_{\mathrm{F}}(0, \phi(t)). \tag{3.A.5}$$

*If we would not have accounted for the sign change of a_2 the determinant of S_{F} would have been -1.

3.B Derivation of the Klich formula

The operator trace (3.3.19) for particle-hole conjugate Majorana operators $a(E) = a^\dagger(-E)$ can be derived from the Klich formula (3.3.13) for self-conjugate Majorana operators $\gamma = \gamma^\dagger$, by performing a unitary transformation:

$$\begin{pmatrix}\gamma_n(E)\\ \gamma'_n(E)\end{pmatrix} = U \begin{pmatrix}a_n(E)\\ a_n^\dagger(E)\end{pmatrix}, \quad U = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -i & i\end{pmatrix}. \tag{3.B.1}$$

At positive energies the γ operators satisfy the Clifford algebra of Majorana operators,

$$\begin{aligned}\{\gamma_n(E),\gamma_m(E')\} &= \{\gamma'_n(E),\gamma'_m(E')\}\\ &= \{\gamma_n(E),\gamma'_m(E')\} = \delta_{nm}\delta_{EE'}, \quad E,E'>0.\end{aligned} \tag{3.B.2}$$

Note that

$$\gamma_n(E)^2 = \gamma'_n(E)^2 = 1/2. \tag{3.B.3}$$

The bilinear form (3.3.15) of the a operators transforms into

$$\boldsymbol{a}^\dagger \cdot \boldsymbol{O} \cdot \boldsymbol{a} = \sum_{n,m}\sum_{E,E'>0} \begin{pmatrix}\gamma_n(E)\\ \gamma'_n(E)\end{pmatrix} \tilde{\mathcal{O}}_{nm}(E,E') \begin{pmatrix}\gamma_m(E')\\ \gamma'_m(E')\end{pmatrix}, \tag{3.B.4}$$

with $\tilde{\mathcal{O}} = U\mathcal{O}U^\dagger$. Because only positive energies appear in Eq. (3.B.4), we may apply the anticommutator (3.B.2), which implies that the traceless symmetric part of $\tilde{\mathcal{O}}$ drops out. Only the trace $\operatorname{Tr}\tilde{\mathcal{O}} = \operatorname{Tr}\mathcal{O}$ and the antisymmetric part $(\tilde{\mathcal{O}} - \tilde{\mathcal{O}}^{\mathrm{T}})/2$ contribute,

$$\boldsymbol{a}^\dagger \cdot \boldsymbol{O} \cdot \boldsymbol{a} = \tfrac{1}{2}\boldsymbol{\gamma} \cdot (\tilde{\mathcal{O}} - \tilde{\mathcal{O}}^{\mathrm{T}}) \cdot \boldsymbol{\gamma} + \tfrac{1}{2}\operatorname{Tr}\boldsymbol{O}. \tag{3.B.5}$$

After these preparations we can apply Klich's original formula [122],

$$\begin{aligned}\left[\operatorname{Tr}\prod_k \exp(\boldsymbol{a}^\dagger \cdot \boldsymbol{O}_k \cdot \boldsymbol{a})\right]^2 &= \exp\left(\sum_k \operatorname{Tr}\boldsymbol{O}_k\right)\\ &\quad \times \operatorname{Det}\left(1 + \prod_k \exp(\tilde{\mathcal{O}}_k - \tilde{\mathcal{O}}_k^{\mathrm{T}})\right).\end{aligned} \tag{3.B.6}$$

Finally we invert the unitary transformation,

$$U^\dagger\tilde{\mathcal{O}}U = \mathcal{O}, \quad U^\dagger\tilde{\mathcal{O}}^T U = (U^{\mathrm{T}}U)^\dagger\mathcal{O}^{\mathrm{T}}(U^{\mathrm{T}}U) = \sigma_x\mathcal{O}^{\mathrm{T}}\sigma_x, \tag{3.B.7}$$

to arrive at

$$\left[\operatorname{Tr}\prod_k \exp(\boldsymbol{a}^\dagger\cdot\boldsymbol{O}_k\cdot\boldsymbol{a})\right]^2 = \exp\left(\sum_k \operatorname{Tr}\boldsymbol{O}_k\right) \times \operatorname{Det}\left(1+\prod_k \exp(\mathcal{O}_k-\sigma_x\mathcal{O}_k^{\mathrm{T}}\sigma_x)\right), \tag{3.B.8}$$

which is Eq. (3.3.19).

3.C Scattering formulas for charge correlators

3.C.1 General expressions for first and second moments

Moments of the transferred charge in the left lead are given by the expectation value

$$\langle Q_{\mathrm{L}}^p\rangle = \left\langle \left(\boldsymbol{a}^\dagger\cdot\boldsymbol{Q}\cdot\boldsymbol{a}\right)^p\right\rangle, \quad \boldsymbol{Q}=\boldsymbol{S}^\dagger\mathcal{P}_{\mathrm{L}}\mathcal{P}_+ e\nu_y\boldsymbol{S}. \tag{3.C.1}$$

In comparison with the number operator (3.3.6) there is a matrix $e\nu_y$ which is the charge operator in the Majorana basis. (It would be $e\nu_z$ in the particle-hole basis.) The expectation value $\langle\cdots\rangle = \operatorname{Tr}(\rho_{\mathrm{eq}}\cdots)$ is with respect to an equilibrium distribution of the $\boldsymbol{a}$ operators, with density matrix (3.3.7).

Because of the Majorana commutator (3.3.14), we have both the usual type-I average

$$\langle a_n^\dagger(E)a_m(E')\rangle = \delta_{nm}\delta(E-E')f(E), \quad f(E)=(1+e^{\beta E})^{-1}, \tag{3.C.2}$$

and the unusual type-II average

$$\langle a_n(E)a_m(E')\rangle = \delta_{nm}\delta(E+E')f(-E), \quad f(-E)=1-f(E). \tag{3.C.3}$$

Averages of strings of a and $a^\dagger$ operators are obtained by summing over all pairwise averages of both types I and II, signed by the permutation.*

*An equivalent procedure [123] is to first use the relation $a_n(-E)=a_n^\dagger(E)$ to rewrite the expectation value such that only positive energies appear, and then apply Wick's theorem as usual.

We assume zero temperature, when $f(E) = \mathcal{P}_-$ and $1 - f(E) = \mathcal{P}_+$ are step functions of energy.

The first moment of the transferred charge contains a single type-I average,

$$\langle Q_{\rm L} \rangle = \operatorname{Tr} \mathcal{P}_- \boldsymbol{Q} = \int_0^\infty \frac{dE}{2\pi} \int_{-\infty}^0 \frac{dE'}{2\pi} \operatorname{Tr} S^\dagger(E,E') e\nu_y \mathcal{P}_{\rm L} S(E,E'). \tag{3.C.4}$$

The variance contains a term with two type-I averages and a term with two type-II averages,

$$\begin{aligned} \operatorname{var}(Q_{\rm L}) = {} & \operatorname{Tr} \mathcal{P}_- \boldsymbol{Q} \mathcal{P}_+ \boldsymbol{Q} \\ & - \int_0^\infty \frac{dE}{2\pi} \int_{-\infty}^0 \frac{dE'}{2\pi} \sum_{n,m} Q_{nm}(-E,-E') Q_{nm}(E,E'). \end{aligned} \tag{3.C.5}$$

The particle-hole symmetry relation (3.3.5) of the scattering matrix implies that

$$Q_{nm}(-E,-E') = -(\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} \mathcal{P}_- e\nu_y \boldsymbol{S})_{mn}(E',E). \tag{3.C.6}$$

Substitution into Eq. (3.C.5) gives

$$\operatorname{var}(Q_{\rm L}) = \operatorname{Tr} \mathcal{P}_- \boldsymbol{Q} \mathcal{P}_+ \boldsymbol{Q} + \operatorname{Tr} \mathcal{P}_- \boldsymbol{Q}' \mathcal{P}_+ \boldsymbol{Q}, \tag{3.C.7}$$

with $\boldsymbol{Q}'$ as in Eq. (3.C.1) upon replacement of $\mathcal{P}_+$ by $\mathcal{P}_-$. Since $\mathcal{P}_+ + \mathcal{P}_- = 1$, this reduces to

$$\operatorname{var}(Q_{\rm L}) = \operatorname{Tr} \mathcal{P}_- (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} e\nu_y \boldsymbol{S}) \mathcal{P}_+ (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} \mathcal{P}_+ e\nu_y \boldsymbol{S}). \tag{3.C.8}$$

It is convenient to eliminate the second $\mathcal{P}_+$ projector from Eq. (3.C.8). This can be done via particle-hole symmetry, which implies that

$$\begin{aligned} & \operatorname{Tr} \mathcal{P}_- (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} e\nu_y \boldsymbol{S}) \mathcal{P}_+ (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} \mathcal{P}_+ e\nu_y \boldsymbol{S}) \\ & \qquad = \operatorname{Tr} (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} \mathcal{P}_+ e\nu_y \boldsymbol{S})^{\rm T} \mathcal{P}_+ (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} e\nu_y \boldsymbol{S})^{\rm T} \mathcal{P}_- \\ & \qquad = \operatorname{Tr} (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} \mathcal{P}_- e\nu_y \boldsymbol{S}) \mathcal{P}_- (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} e\nu_y \boldsymbol{S}) \mathcal{P}_+ \\ & \qquad = \operatorname{Tr} \mathcal{P}_- (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} e\nu_y \boldsymbol{S}) \mathcal{P}_+ (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} \mathcal{P}_- e\nu_y \boldsymbol{S}). \end{aligned} \tag{3.C.9}$$

Hence

$$\tfrac{1}{2} \operatorname{Tr} \mathcal{P}_- (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} e\nu_y \boldsymbol{S}) \mathcal{P}_+ (\boldsymbol{S}^\dagger \mathcal{P}_{\rm L} (\mathcal{P}_- - \mathcal{P}_+) e\nu_y \boldsymbol{S}) = 0, \tag{3.C.10}$$

and adding this to Eq. (3.C.8) we arrive at

$$\begin{aligned}\mathrm{var}\,(Q_\mathrm{L}) &= \frac{1}{2}\mathrm{Tr}\,\mathcal{P}_-(\boldsymbol{S}^\dagger\mathcal{P}_\mathrm{L}e\nu_y\boldsymbol{S})\mathcal{P}_+(\boldsymbol{S}^\dagger\mathcal{P}_\mathrm{L}e\nu_y\boldsymbol{S})\\ &= \frac{1}{2}e^2\int_0^\infty\frac{dE}{2\pi}\int_{-\infty}^0\frac{dE'}{2\pi}\,\mathrm{Tr}\,\Sigma_\mathrm{L}^\dagger(E,E',)\Sigma_\mathrm{L}(E,E'), \qquad (3.\mathrm{C}.11)\\ \boldsymbol{\Sigma}_\mathrm{L} &= \boldsymbol{S}^\dagger\mathcal{P}_\mathrm{L}\nu_y\boldsymbol{S}.\end{aligned}$$

The expressions for the other correlators are analogous,

$$\mathrm{var}\,(Q_\mathrm{R}) = \frac{1}{2}e^2\int_0^\infty\frac{dE}{2\pi}\int_{-\infty}^0\frac{dE'}{2\pi}\,\mathrm{Tr}\,\Sigma_\mathrm{R}^\dagger(E,E')\Sigma_\mathrm{R}(E,E'), \qquad (3.\mathrm{C}.12)$$

$$\begin{aligned}\boldsymbol{\Sigma}_\mathrm{R} &= \boldsymbol{S}^\dagger\mathcal{P}_\mathrm{R}\nu_y\boldsymbol{S},\\ \mathrm{covar}\,(Q_\mathrm{L}Q_\mathrm{R}) &= \frac{1}{4}e^2\int_0^\infty\frac{dE}{2\pi}\int_{-\infty}^0\frac{dE'}{2\pi}\,\mathrm{Tr}\,\Big[\Sigma_\mathrm{L}^\dagger(E,E')\Sigma_\mathrm{R}(E,E')\\ &\quad+\Sigma_\mathrm{R}^\dagger(E,E')\Sigma_\mathrm{L}(E,E')\Big]. \qquad (3.\mathrm{C}.13)\end{aligned}$$

Eq. (3.C.13) gives the symmetrized covariance,

$$\mathrm{covar}(Q_\mathrm{L}Q_\mathrm{R}) \equiv \tfrac{1}{2}\langle Q_\mathrm{L}Q_\mathrm{R}\rangle + \tfrac{1}{2}\langle Q_\mathrm{R}Q_\mathrm{L}\rangle - \langle Q_\mathrm{L}\rangle\langle Q_\mathrm{R}\rangle, \qquad (3.\mathrm{C}.14)$$

appropriate for a calculation of $\mathrm{var}\,(Q_\mathrm{L}\pm Q_\mathrm{R})$.

3.C.2 Adiabatic approximation

The general expressions (3.C.4) and (3.C.11)–(3.C.13) can be simplified in the adiabatic regime, when near the Fermi level $S(E,E')$ depends only on the energy difference $\omega = E - E'$. We use the identity

$$\int_0^\infty dE\int_{-\infty}^0 dE'\,F(E-E') = \int_{0^+}^\infty d\omega\,\omega F(\omega). \qquad (3.\mathrm{C}.15)$$

The lower integration limit 0^+ eliminates a possibly singular delta function in $F(\omega)$, which should not enter in the excitation spectrum.

For the average transferred charge (3.C.4) we thus have

$$\langle Q_\mathrm{L}\rangle = \frac{1}{4\pi^2}\int_{0^+}^\infty d\omega\,\omega\,\mathrm{Tr}\,S^\dagger(\omega)e\nu_y\mathcal{P}_\mathrm{L}S(\omega). \qquad (3.\mathrm{C}.16)$$

As explained in Ref. [116], this is equivalent to the Brouwer formula (3.5.1): Because of

$$[S^\dagger(\omega)\nu_y\mathcal{P}_\mathrm{L}S(\omega)]^\mathrm{T} = -S^\dagger(-\omega)\nu_y\mathcal{P}_\mathrm{L}S(-\omega) \tag{3.C.17}$$

the integrand in Eq. (3.C.16) is an even function of ω, hence the integration can be extended to $\int_{-\infty}^{\infty} d\omega$, and then transformation to the time domain gives Eq. (3.5.1).

For the second moments we use that the kernels $\Sigma(E, E') \mapsto \Sigma(\omega)$ are functions of $\omega = E - E'$ when $S(E, E') \mapsto S(\omega)$,

$$\begin{aligned}
\Sigma_{\mathrm{L,R}}(E, E') &= \int_{-\infty}^{\infty} \frac{dE''}{2\pi}\, S^\dagger(E'', E)\mathcal{P}_{\mathrm{L,R}}\nu_y S(E'', E') \\
\Rightarrow \Sigma_{\mathrm{L,R}}(\omega) &= \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi}\, S^\dagger(\omega' - \omega)\mathcal{P}_{\mathrm{L,R}}\nu_y S(\omega') \\
&= \int_{-\infty}^{\infty} dt\, e^{i\omega t} S^\dagger(t)\mathcal{P}_{\mathrm{L,R}}\nu_y S(t).
\end{aligned} \tag{3.C.18}$$

The Fourier transform is defined as

$$S(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\omega t} S(t). \tag{3.C.19}$$

Note that for the representation (3.C.18) of $\Sigma(\omega)$ as a single time integral it was essential that we eliminated the $\mathcal{P}_+$ projector from the scattering matrix product.

Application of Eqs. (3.C.15) and (3.C.18) to Eqs. (3.C.11)–(3.C.13) then gives the formulas (3.5.4) from the main text.

4 Bounds on nonlocal correlations in the presence of signaling and their application to topological zero modes

4.1 Introduction

In a Bell test [127, 128], Alice and Bob measure pairs of particles (possibly having a common source in their past) and then communicate in order to calculate the correlations between these measurements. The strength of empirical correlations enables one to characterize the underlying theory. In quantum mechanics, the above procedure corresponds to local measurements of Hermitian operators A_0/A_1 on Alice's side and B_0/B_1 on Bob's side. The correlators are defined using the quantum expectation value $c_{ij} = \langle A_i B_j \rangle$ and, when the operators have eigenvalues ± 1, it can be shown that the CHSH parameter obeys

$$|\mathcal{B}| \equiv |c_{00} + c_{10} + c_{01} - c_{11}| \leq 2\sqrt{2}, \tag{4.1.1}$$

which is known as the Tsirelson bound [129]. Stronger bounds on the correlators (i.e., bounds from which the Tsirelson bound can be derived) were proposed, e.g., by Uffink [130] and independently by Tsirelson, Landau and Masanes (TLM) [131–133]. The latter implies that

$$|c_{00}c_{10} - c_{01}c_{11}| \leq \sum_{j=0,1} \sqrt{(1 - c_{0j}^2)(1 - c_{1j}^2)}. \tag{4.1.2}$$

The TLM inequality is known to be necessary and sufficient for the correlators c_{ij} to be realizable in quantum mechanics [131–133] (implying, in particular, that if a set of correlators satisfies Eq. (4.1.2), it necessarily satisfies Eq. (4.1.1); the converse is not true). Importantly, when calculating

$\mathcal{B}$ in any local-realistic model it turns out that $|\mathcal{B}| \leq 2$, which is a famous variant of Bell's theorem known as the Clauser-Horne-Shimony-Holt (CHSH) inequality [134], which provides a measurable distinction between correlations achievable in local-realistic models and in quantum theory. These bounds, however, are not enough for fully characterizing the Alice-Bob quantum correlations. For the latter task, the Navascues-Pironio-Acin (NPA) hierarchical scheme of semidefinite programs was proposed [135].

All the above works plausibly assuming that Alice's and Bob's measurements are described by spatially local and Hermitian operators, implying that $[A_i, B_j] = 0$ for all i, j. As such, they cannot lead to superluminal signaling between Alice and Bob.

Trying, on the one hand, to generalize some of the above results, and on the other hand to pin-point the core reason they work so well, we relax below these two assumptions and examine the consequences of complex-valued correlations emerging from non-Hermitian non-commuting Alice and Bob operators. We thus allow a restricted form of signaling between the parties (similar to the one in [136]), but we maintain the Hilbert space structure, as well as other core ingredients of quantum mechanics. Surprisingly, the Tsirelson bound and TLM inequality remain valid in this generalized setting. Apart from that, we find intriguing relations between nonlocality, local correlations of Alice (or Bob), and signaling in the case of Hermitian yet non-commuting observables.

Considering non-Hermitian non-commuting observables may seem far from any sensible model. To alleviate this impression, we study an explicit example of a parafermionic system, which is a proper quantum system that provides a natural setting for comparing commuting and non-commuting sets of observables. The natural observables in the parafermionic system happen to be non-Hermitian. Parafermions (or rather parafermionic zero modes) are topological zero modes that generalize the better-known Majorana zero modes [6, 21, 113]. Parafermions can be realized in various quasi-one-dimensional systems [40, 137–142], see [143] for a comprehensive review. Similarly to the case of Majoranas, observables in a system of parafermions are inherently non-local as they comprise at least two parafermionic operators hosted at different spatial locations. In the case of Majoranas, this nonlocality is known to have manifestations through the standard CHSH inequality [144]. We do not follow the investigation line of Ref. [144], but rather investigate a different aspect of nonlocality, which is absent for Majoranas yet present for parafermions.

Specifically, we construct two examples. In the first, the system of parafermions is split into two spatially separated parts, A and B, with commuting observables $[A_i, B_j] = 0$. In the second example, Alice's and

Bob's parts are still spatially separated; the local permutation properties of A_0, A_1, as well as those of B_0, B_1 are exactly the same as in the first example, yet $[A_i, B_j] \neq 0$. This property alone has the potential to contradict relativistic causality since we have spatially separated observables which do not commute and thus allow for superluminal signaling (thus these systems can indeed simulate the case of non-Hermitian signaling operators). However, as we explain in Sec. 4.3.2, in order to measure their respective observables, Alice and Bob in our system must share a common region of space, which resolves the paradox. In this sense, Alice and Bob can be thought of as two experimenters acting on the same system. Therefore, the system of parafermions does not constitute a system in which the spatial and quantum mechanical notions of locality disagree. However, it simulates such a system (with spatial locality interpreted in a very naive way). Using these examples we investigate the theoretical bounds on correlations. We find that both systems obey the derived bounds. However, the maximal achievable correlations in the truly local system (first example) are significantly weaker than those of the non-local one.

Before we present our results in the next sections, one comment is due. One may think that investigating Bell-CHSH correlations with $[A_i, B_j] \neq 0$ is an abuse of notation. Originally introduced for distinguishing local-realistic theories from the standard quantum theory, the Bell-CHSH inequalities imply the use of conditional probabilities $P(a, b|i, j)$ that are defined in both. With $[A_i, B_j] \neq 0$, the correlators that have the same operator form are expressed not through probability distributions $P(a, b|i, j)$ but rather through quasiprobability distributions $W(a, b|i, j)$, cf. Appendix 4.C. Therefore, a formal replacement of commuting operators with non-commuting ones may seem an illegitimate operation in this context. We would like to emphasize that the key to comparing properties of different theories is considering objects that are defined in these theories in an operationally identical way. This is the reason that local-realistic theories are compared to quantum mechanics not in terms of the joint probability distribution $P(a_0, a_1, b_0, b_1)$ (that does not exist in quantum theory when $[A_i, A_j] \neq 0$ and/or $[B_i, B_j] \neq 0$) but in terms of $P(a, b|i, j)$ conditioned on the choice of observables: $P(a, b|i, j)$ are defined in both theories and can be measured by the same measurement procedure. Since our aim here is to compare the standard quantum theory with that allowing for $[A_i, B_j] \neq 0$, working in the language of correlators that are defined and can be measured (even if they are complex) by means of weak measurements in both theories [145] is a natural decision. We thus compare nonlocal theories having a Hilbert space structure, rather

than a probabilistic structure (common, e.g., to local hidden variables theories and quantum mechanics, but not to the post-quantum theories discussed here). However, in the case of the standard quantum theory, the correlation functions (and thus our new bounds) can be expressed in terms of $P(a,b|i,j)$, making them new bounds on the possible probability distributions in the standard quantum theory.

In what follows, we start in Sec. 4.2 by defining an operator-based (rather than probability-based) notion of complex correlations arising in nonlocal, non-Hermitian systems admitting signaling and then find the generalized inequalities bounding them. Importantly, this notion has an operational sense in terms of weak measurements, as discussed in the Appendix 4.C.2. In Sec. 4.3, we review parafermionic systems and show they can simulate such non-Hermitian signaing systems. We then numerically prove they are indeed bounded by the proposed bounds. Sec. 4.4 concludes the chapter. Some technical details appear in the Appendices.

4.2 Analytic results for correlations of general non-Hermitian non-commuting operators

Below we prove a number of bounds on quantum correlations of non-Hermitian non-commuting operators. We generalize the Tsirelson and the TLM bounds (Theorems 1 and 2, which have been previously derived for Hermitian commuting operators, see [146]) and derive previously unknown bounds (Theorems 3 and 4, which are applicable to the Hermitian, non-signaling case as well). Here we introduce the bounds and discuss them, while their proofs are deferred to Sec. 4.2.1. The bounds are expressed in terms of Pearson correlation functions of operators X and Y defined as

$$C(X,Y)=\frac{\langle XY^\dagger\rangle-\langle X\rangle\langle Y\rangle^\dagger}{\Delta_X\Delta_Y}, \tag{4.2.1}$$

where $\Delta_X=\sqrt{\langle XX^\dagger\rangle-|\langle X\rangle|^2}$ is the variance of X (which is assumed to be non-zero), and averaging is performed with respect to some state $|\psi\rangle$ in the Hilbert space. This definition is a straightforward generalization of the usual Pearson correlation between commuting Hermitian operators. The Pearson correlations reduce to the standard $c_{XY}=\langle XY\rangle$ for Hermitian X and Y on states $|\psi\rangle$ such that $\langle X\rangle=\langle Y\rangle=0$ and $\Delta_X=\Delta_Y=1$. We note that $C(X,Y)$ is ill-defined when $\Delta_X=0$ or $\Delta_Y=0$; yet, as we show

in Sec. 4.2.1, $|C(X,Y)| \leq 1$ everywhere, including the vicinity of such special points.

For the case of commuting operators X, Y, the definition of $C(X,Y)$ can be expressed in terms of the joint probability distributions, and our below bounds can be thought of as restricting the possible probability distributions in quantum theory. When X and Y do not commute, this is not the case, which defies the notions that conventionally underlie Bell inequalities. However, our aim here is not to analyze complex local hidden variables models but rather to examine general models which are manifestly nonlocal. In particular, we wish to analyze whether known bounds on quantum correlations remain effective when generalized to cases of non-Hermitian signaling operators. We argue that these complex correlations are physically meaningful because there is an empirical protocol for measuring them. That operational meaning of the above correlations in terms of weak measurements is given in Appendix 4.C.2. Alternatively, for the case of non-commuting observables, $C(X,Y)$ can be expressed in terms of quasiprobability distibutions, and thus our bounds restrict possible quasiprobability distributions in that case. We discuss this in detail in Appendix 4.A.

We now discuss the bounds on Alice-Bob correlations.

Theorem 1. ***(Generalized Tsirelson bound).*** *Define* $\mathcal{B} \stackrel{\text{def}}{=} C(A_0,B_0) + C(A_1,B_0) + C(A_0,B_1) - C(A_1,B_1)$ *as the complex-valued Bell-CHSH parameter of any operators* A_i *and* B_j*. The following holds*

$$\begin{aligned} |\mathcal{B}| &= \sqrt{\mathrm{Re}(\mathcal{B})^2 + \mathrm{Im}(\mathcal{B})^2} \\ &\leq \sqrt{2}\left[\sqrt{1+\mathrm{Re}(\eta)} + \sqrt{1-\mathrm{Re}(\eta)}\right] \\ &\leq 2\sqrt{2}, \end{aligned} \tag{4.2.2}$$

where η *is either* $C(A_0,A_1)$ *or* $C(B_0,B_1)$ *(the one having the larger* $|\mathrm{Re}(\eta)|$ *among them will give rise to a tighter inequality).*

Despite the fact that $C(X,Y) \neq c_{XY}$, the Bell-CHSH parameter defined through $C(X,Y)$ obeys the same Tsirelson bound as for c_{XY} in Eq. (4.1.1). Moreover, the proof of the Tsirelson bound for $C(X,Y)$ is valid independently of whether $[A_i,B_j]=0$. A somewhat tighter bound (the middle row of Eq. (4.2.2)) is obtained in terms of η that expresses on-site correlations on Alice's or Bob's side. This is also insensitive to whether $[A_i,B_j]=0$.

Theorem 2. ***(Generalized TLM bound).*** *The following holds for any*

operators A_i, B_j, $i,j \in \{0,1\}$,

$$\left|C(B_0,A_0)^\dagger C(B_0,A_1) - C(B_1,A_0)^\dagger C(B_1,A_1)\right| \leq \sum_{j=0,1} \sqrt{(1-|C(B_j,A_0)|^2)(1-|C(B_j,A_1)|^2)}. \quad (4.2.3)$$

Similarly to the previous theorem, this bound is insensitive to whether $[A_i,B_j]=0$ and has the same form as the standard TLM bound, Eq. (4.1.2), modulo replacing $C(X,Y)$ with real-valued c_{XY}.

We note in passing that our bounds apply to both operators with bounded and unbounded spectrum. Implementing Bell tests in mesoscopic systems often requires dealing with operators having an unbounded spectrum, cf. Ref. [147]. Our theorems 1 and 2 may thus be useful for studies in such systems.

Theorem 3. ***(Relation between nonlocality, local correlations, and signaling)***. *Let $\mathcal{B}$ be the complex-valued Bell-CHSH parameter defined in Theorem 1. Then,*

$$\left(\frac{\mathrm{Re}(\eta)}{2}\right)^2 + \left(\frac{\mathrm{Re}(\mathcal{B})}{2\sqrt{2}}\right)^2 + \left(\frac{\mathrm{Im}(\mathcal{B})}{2\sqrt{2}}\right)^2 \leq 1. \quad (4.2.4)$$

This bound is also valid independently of $[A_i,B_j]=0$. In the case of Hermitian A_i, B_j that obey $[A_i,B_j]=0$, $C(A_i,B_j)$ is real, implying $\mathrm{Im}(\mathcal{B})=0$. If A_i and B_j are Hermitian but do not mutually commute, there can appear imaginary components to $C(A_i,B_j)$ and $\mathcal{B}$. Therefore, this relation may be interpreted as a constraint on non-local correlations (represented by $\mathrm{Re}(\mathcal{B})/(2\sqrt{2})$), local on-site correlations ($\mathrm{Re}(\eta)/2$), and signaling (represented by $\mathrm{Im}(\mathcal{B})/(2\sqrt{2}) \neq 0$). These three quantities are thus confined to the unit ball.

Theorem 4. *Let $\mathcal{B}$ be the complex-valued Bell-CHSH parameter defined in Theorem 1. In the case of isotropic correlations, $C(A_i,B_j)=(-1)^{ij}\varrho$ (such that $\mathcal{B}=4\varrho$) for some complex-valued ϱ,*

$$|\eta|^2 + \left(\frac{\mathrm{Re}(\mathcal{B})}{2\sqrt{2}}\right)^2 + \left(\frac{\mathrm{Im}(\mathcal{B})}{2\sqrt{2}}\right)^2 \leq 1. \quad (4.2.5)$$

Note that Eq. (4.2.5) provides a tighter bound than Eq. (4.2.4). However, Eq. (4.2.5) is proved under the rather restrictive assumption of $C(A_i,B_j) =$

$(-1)^{ij}\varrho$. This is a valid assumption within non-signaling theories in the following sense. Reference [148] argued that the standard Bell-CHSH parameter $\mathcal{B} = c_{00} + c_{10} + c_{01} - c_{11}$ for ± 1-valued observables in a non-signaling theory (not necessarily classical or quantum) can always be maximized on a state satisfying $c_{ij} = (-1)^{ij}\rho$ with a real ρ. While the statement of Ref. [148] was proved for the standard correlations c_{XY} (and not our $C(X,Y)$) and maximizing the l.h.s. of Eq. (4.2.5) is not equivalent to maximizing $|\mathcal{B}|$, one might hope that the possibility of arranging $C(A_i, B_j) = (-1)^{ij}\varrho$ is related to non-signaling, and the bound of Eq. (4.2.5) would discriminate the cases of $[A_i, B_j] = 0$ and $[A_i, B_j] \neq 0$. We provide some numerical evidence for the last statement in Sec. 4.3.

4.2.1 Proofs of analytic bounds

Lemma 1. ***(Generalized uncertainty relations, see Ref. [149] for elaboration on the term)***. *Denote by $X_1, \ldots, X_n$, a number of operators. Let C be an $n \times n$ Hermitian matrix whose ij-th entry is*

$$C(X_i, X_j) = \frac{\langle X_i X_j^\dagger \rangle - \langle X_i \rangle \langle X_j \rangle^\dagger}{\Delta_{X_i} \Delta_{X_j}}, \tag{4.2.6}$$

where $\Delta_X = \sqrt{\langle XX^\dagger \rangle - |\langle X \rangle|^2}$ is the uncertainty in X (which is assumed to be non-zero). Then $C \succeq 0$, i.e., it is positive semidefinite.

Proof. Denote $|\psi\rangle$, the underlying quantum state. For any n-dimensional vector, $v^T = [v_1, \ldots, v_n]$, it follows that

$$v^T D C D^T v = \langle \phi | \phi \rangle \geq 0, \tag{4.2.7}$$

where D is a (positive semidefinite) diagonal matrix whose entries are $D_{ii} = \Delta_{X_i}$, and $|\phi\rangle = \sum_{i=1}^n v_i \left(X_i - \langle X_i \rangle\right) |\psi\rangle$. Therefore, $DCD^T \succeq 0$ and so is $C \succeq 0$. □

Applying this lemma to two operators, X_1, X_2, one obtains an inequality $|C(X_1, X_2)| \leq 1$, implying that the correlation functions are bounded even near $\Delta_{X_{1,2}} = 0$.

Theorem 1. *Proof.* Construct the matrix C for the operators A_0, A_1,

and B_j,

$$\begin{bmatrix} C(B_j,B_j) & C(B_j,A_1) & C(B_j,A_0) \\ C(B_j,A_1)^\dagger & C(A_1,A_1) & C(A_0,A_1) \\ C(B_j,A_0)^\dagger & C(A_0,A_1)^\dagger & C(A_0,A_0) \end{bmatrix} = \begin{bmatrix} 1 & C(B_j,A_1) & C(B_j,A_0) \\ C(B_j,A_1)^\dagger & 1 & \eta \\ C(B_j,A_0)^\dagger & \eta^\dagger & 1 \end{bmatrix} \succeq 0, \quad (4.2.8)$$

where $\eta \stackrel{\text{def}}{=} C(A_0,A_1)$. By the Schur complement condition for positive semidefiniteness this is equivalent to

$$C^A \stackrel{\text{def}}{=} \begin{bmatrix} 1 & \eta \\ \eta^\dagger & 1 \end{bmatrix} \succeq \begin{bmatrix} C(B_j,A_1)^\dagger \\ C(B_j,A_0)^\dagger \end{bmatrix} \begin{bmatrix} C(B_j,A_1) & C(B_j,A_0) \end{bmatrix}. \quad (4.2.9)$$

Let $v_j^T = [(-1)^j, 1]$. The above inequality implies

$$2(1+(-1)^j \mathrm{Re}(\eta)) = v_j^T C^A v_j \geq \left| C(B_j,A_0) + (-1)^j C(B_j,A_1) \right|^2. \quad (4.2.10)$$

This together with the triangle inequality yield

$$|\mathcal{B}| \leq \sum_{j=0,1} \left| C(B_j,A_0) + (-1)^j C(B_j,A_1) \right| \leq \sqrt{2} \sum_{j=0,1} \sqrt{1+(-1)^j \mathrm{Re}(\eta)}, \quad (4.2.11)$$

which completes the proof. Note that by swapping the roles of A and B, a similar inequality is obtained where $\eta = C(B_0,B_1)$. □

Theorem 2. *Proof.* The inequality (4.2.9) implies

$$(1-|C(B_j,A_0)|^2)(1-|C(B_j,A_1)|^2) - \left| \eta - C(B_j,A_0)^\dagger C(B_j,A_1) \right|^2 \geq 0, \quad (4.2.12)$$

which follows from the non-negativity of the determinant of the matrix obtained by subtracting the right hand side from the left hand side in (4.2.9). Therefore,

$$\left| \eta - C(B_j,A_0)^\dagger C(B_j,A_1) \right| \leq \sqrt{(1-|C(B_j,A_0)|^2)(1-|C(B_j,A_1)|^2)}. \quad (4.2.13)$$

This and the triangle inequality give rise to the theorem,

$$\begin{aligned}\left|C(B_0,A_0)^\dagger C(B_0,A_1)-C(B_1,A_0)^\dagger C(B_1,A_1)\right| &\\ \leq \sum_{j=0,1}\left|\eta - C(B_j,A_0)^\dagger C(B_j,A_1)\right| &\\ \leq \sum_{j=0,1}\sqrt{(1-|C(B_j,A_0)|^2)(1-|C(B_j,A_1)|^2)}. \quad \square \quad (4.2.14)\end{aligned}$$

Theorem 3. *Proof.* We have seen that

$$|\mathcal{B}| \leq \sqrt{2}\left(\sqrt{1+\mathrm{Re}(\eta)}+\sqrt{1-\mathrm{Re}(\eta)}\right). \quad (4.2.15)$$

Therefore,

$$|\mathcal{B}|^2 = \mathrm{Re}(\mathcal{B})^2+\mathrm{Im}(\mathcal{B})^2 \leq 4\left(1+\sqrt{1-\mathrm{Re}(\eta)^2}\right). \quad (4.2.16)$$

Because, $\sqrt{1-a} \leq 1-a/2$ for $a \in [0,1]$, it follows that

$$\mathrm{Re}(\mathcal{B})^2+\mathrm{Im}(\mathcal{B})^2 \leq 8-2\mathrm{Re}(\eta)^2, \quad (4.2.17)$$

from which the theorem follows. □

Theorem 4. *Proof.* In case the isotropy holds, i.e., $C(A_i,B_j) = C(B_j,A_i)^* = (-1)^{ij}\varrho$, (4.2.12) reads

$$\left|\eta-(-1)^j|\varrho|^2\right|^2 \leq (1-|\varrho|^2)^2, \quad (4.2.18)$$

and thus

$$|\eta|^2-2(-1)^j|\varrho|^2\mathrm{Re}(\eta) \leq 1-2|\varrho|^2. \quad (4.2.19)$$

Averaging both sides in this inequality over $j=0,1$, and rearranging give

$$|\eta|^2+2|\varrho|^2 \leq 1. \quad (4.2.20)$$

Finally, substituting $\varrho = \mathcal{B}/4$ into (4.2.20) yields the theorem. □

4.3 Investigating the bounds in the system of parafermions

Parafermions provide a unique test system for the bounds proven in the previous section. First, the natural observables in a system of parafermions

are non-Hermitian. Second, in this system the non-commutativity between Alice's and Bob's operators can be switched on and off without changing anything else about the algebra of operators, enabling a clean investigation of the effect of Alice-Bob non-commutativity. Finally, there have been a number of proposals for experimental implementations of parafermions [40, 137–142], which opens the way for experimental verification of our predictions.

The structure of the section is as follows. In Sec. 4.3.1, we give a brief introduction to the physics of parafermions and the algebra of their operators. In Sec. 4.3.2, we construct the observables of Alice and Bob. Those not interested in the physics of parafermions may skip directly to Eqs. (4.3.8–4.3.10) detailing the permutation relations of the observables and Eqs. (4.3.11–4.3.16) introducing their explicit matrix representation. In Sec. 4.3.3, we provide the results of the numerical investigation of bounds (4.2.2–4.2.5).

4.3.1 Parafermion physics and algebra

Parafermionic zero modes can be created in a variety of settings [40, 137–142]. In different settings, they have subtly different properties. We focus on parafermions implemented with the help of fractional quantum Hall (FQH) edges proximitized by a superconductor [40, 137, 140]. The setup employs two FQH puddles of the same filling factor ν (grey regions in Fig. 4.1a) separated by vacuum. This gives rise to two counter-propagating chiral FQH edges. The edges can be gapped either by electron tunneling between them (T domains) or by proximity-induced superconducting pairing of electrons at the edges (SC domains). Domain walls between the domains of two types host parafermionic zero modes $\alpha_{s,j}$ with $s = R/L = \pm 1$ denoting whether a parafermion belongs to the right- or left-propagating edge respectively, and j denoting the domain wall number.

Parafermion operators have the following properties:

$$\alpha_{s,j}^{2/\nu} = \alpha_{s,j}\alpha_{s,j}^\dagger = \alpha_{s,j}^\dagger\alpha_{s,j} = 1, \tag{4.3.1}$$

$$\alpha_{s,j}\alpha_{s,k} = \alpha_{s,k}\alpha_{s,j}e^{i\pi\nu s\mathrm{sgn}(k-j)}, \tag{4.3.2}$$

$$\alpha_{R,j}\alpha_{L,k} = \alpha_{L,k}\alpha_{R,j}\begin{cases} e^{i\pi\nu} & , k \neq j, \\ 1 & , k = j \text{ are even}, \\ e^{2i\pi\nu} & , k = j \text{ are odd}, \end{cases} \tag{4.3.3}$$

where sgn is the sign function. These properties are valid for $\nu = 1/(2m+1)$, $m \in \mathbb{Z}_+$ considered in Refs. [40, 137] and for $\nu = 2/3$ considered in

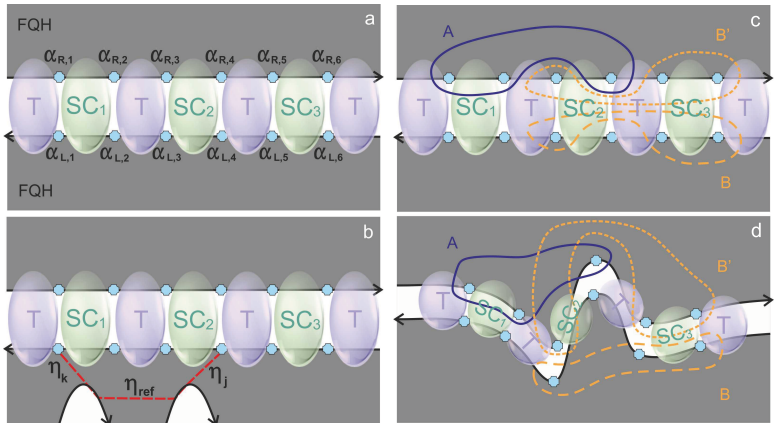

Figure 4.1: a, b — A physical setup for creating and measuring parafermions. a — Setup for implementing parafermions (represented in cyan) with two fractional quantum Hall (FQH) edges (arrows) supporting a series of electron-tunneling-gapped (T) and superconductivity-gapped (SC) domains. b — Setup for measuring parafermionic observables with the help of two additional FQH edges (curved arrows) as in Ref. [150] (cf. Appendix 4.B). c, d — Parafermionic observables and their mutual locality. c — Grouping parafermions into groups belonging to Alice (A) and Bob (B/B'). d — A and B do not have common parafermions, are mutually local, and can be made arbitrarily distant in space. While A and B' do not have common parafermions, they are not mutually local: for Alice to measure A while Bob can measure B', there should be a region of the upper FQH puddle accessible both to Alice and Bob.

Ref. [140]. In the case of $\nu = 1$, parafermions reduce to Majorana operators and $\alpha_{R,j} = \alpha_{L,j}$.

The physics of parafermions is associated with degenerate ground states of the system. Namely, beyond hosting Cooper pairs, each superconducting domain SC_j can host a certain charge Q_j (mod$2e$) quantized in the units of charge of FQH quasiparticles νe. Thus each Q_j has $d = 2/\nu$ distinct values, and the ground state degeneracy of a system as in Fig. 4.1a is therefore $d^{N_{\mathrm{SC}}}$, where N_{SC} is the number of SC domains. Parafermionic operators $\alpha_{s,j}$ act in this degenerate space of ground states and represent the effect of adding a FQH quasiparticle to the system from a FQH puddle corresponding to s at domain wall j. Various observables in the system of parafemions can be expressed through unitary operators $\alpha_{s,j}\alpha^\dagger_{s,k}$. In particular, Q_j themselves can be expressed through $e^{i\pi s(Q_j/e-\nu/2)} = (-1)^{2/\nu}\alpha^\dagger_{s,2j-1}\alpha_{s,2j}$. One can show that $\left(\alpha_{s,j}\alpha^\dagger_{s,k}\right)^d = -e^{2i\pi/\nu}$, which implies that $\alpha_{s,j}\alpha^\dagger_{s,k}$ has d distinct eigenvalues, all having the form $-e^{i\pi\nu(r+1/2)}$ with $r \in \mathbb{Z}$.

Unitary operators $\alpha_{s,j}\alpha^\dagger_{s,k}$ are thus natural "observables" in the system despite being non-Hermitian. The permutation relations of such operators immediately follow from Eqs. (4.3.1-4.3.3). Despite being spatially disconnected, such operators composed of different pairs of parafermions may not commute, e.g.,

$$\alpha_{R,2}\alpha^\dagger_{R,4}\alpha_{R,3}\alpha^\dagger_{R,5} = \alpha_{R,3}\alpha^\dagger_{R,5}\alpha_{R,2}\alpha^\dagger_{R,4}e^{2i\pi\nu}. \tag{4.3.4}$$

It is interesting to note that in the case of Majoranas ($\nu = 1$), none of these two unique properties would hold: the operators $i\alpha_{s,j}\alpha^\dagger_{s,k}$ would be Hermitian, while two such operators having no common Majoranas would commute.

4.3.2 Alice's and Bob's observables

For a parfermionic system with three SC domains (as in Fig. 4.1) with a fixed total charge, the ground state is d^2-degenerate, which allows to split it into two distinct subsystems: SC_1 and SC_3 domains, each having degeneracy d as each can have d distinct values of charge Q_j. The charge of SC_2 domain is determined by the state of SC_1 and SC_3 in order for the total charge to be fixed. This system is thus a natural candidate for studying quantum correlations between two subsystems. To this end, we introduce observables accessible to Alice,

$$A_0 = \alpha_{R,2}\alpha^\dagger_{R,4}, \quad A_1 = \alpha_{R,1}\alpha^\dagger_{R,4}, \tag{4.3.5}$$

and two different sets of observables accessible to Bob:

$$B_0 = \alpha_{L,3}\alpha_{L,5}^{\dagger}, \quad B_1 = \alpha_{L,3}\alpha_{L,6}^{\dagger}, \tag{4.3.6}$$

and

$$B_0' = \alpha_{R,3}\alpha_{R,6}^{\dagger}, \quad B_1' = \alpha_{R,3}\alpha_{R,5}^{\dagger}. \tag{4.3.7}$$

They have identical local algebra, yet different commutation properties of Alice's and Bob's observables:

$$A_0 A_1 = A_1 A_0 e^{-i\pi\nu}, \tag{4.3.8}$$

$$B_0 B_1 = B_1 B_0 e^{-i\pi\nu}, \quad B_0' B_1' = B_1' B_0' e^{-i\pi\nu}, \tag{4.3.9}$$

$$[A_j, B_k] = 0, \quad A_j B_k' = B_k' A_j e^{2i\pi\nu}. \tag{4.3.10}$$

The non-commutation of A and B' observables would imply the possiblity of superluminal signaling had the observables been truly spatially separate (which is not the case, as we explain below). Therefore, we call the set of A and B a non-signaling set, and the set of A and B' a signaling set of observables.

Naively, Alice's observables are local with respect to either set of Bob's observables, cf. Fig. 4.1c. Indeed, A and either the B or B' set use different parafermions, which can be made arbitrarily distant from each other, cf. Fig. 4.1d. However, the locality issue in this system is subtler as in order to probe an observable of the form $\alpha_{s,j}\alpha_{s,k}^{\dagger}$, one needs to enable FQH quasiparticle tunneling to both parafermions simultaneously (see Appendix 4.B). At the same time, quasiparticles can tunnel to a parafermion only from the FQH puddle corresponding to the parafermion index s, not through vacuum and not from the other puddle. Therefore, as can be seen from Fig. 4.1d, the A and B sets are indeed mutually local, while A and B' are not. The ability of Alice to measure observables in A and of Bob to measure observables in B', requires them to have access to a common region of the upper FQH puddle. Thus, the system does not violate the laws of quantum mechanics, nor exhibits superluminal signaling. Nevertheless, it presents a unique opportunity for comparing correlations of commuting and non-commuting (but otherwise equivalent) sets of observables.

The standard tool for studying quantum correlations is given by Bell inequalities. However, since the observables considered here have more than two eigenvalues, we require CHSH-like inequalities suitable for multi-outcome measurements. We study the inequalities introduced in Theorems 1–4, as well as an inequality from Ref. [151]. These inequalities

involve correlators of the form $\langle A_j B_k^\dagger \rangle$ and $\langle A_j \, (B_k')^\dagger \rangle$. Since $[A_j, B_k] = 0$, $\langle A_j B_k^\dagger \rangle$ can be experimentally obtained by performing strong measurements of A_j and B_k separately according to the protocol of Appendix 4.B and then calculating the correlations. Alternatively, these correlations can be measured with weak measurements [67, 68]. The non-commutativity of A_j and B_k' does not allow for a strong-measurement-based approach in the case of $\langle A_j \, (B_k')^\dagger \rangle$. However, this correlator can be measured with the help of weak measurements as described in Appendix 4.C.

From now on we focus on parafermions implemented using $\nu = 2/3$ FQH puddles. Using permutation relations (4.3.8–4.3.10) supplemented by the permutation relations of B_j and B_k', as well as $\left(\alpha_{s,j}\alpha_{s,k}^\dagger\right)^3 = 1$, one can derive an explicit matrix representation for observables (4.3.5–4.3.7):

$$A_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{2\pi i/3} & 0 \\ 0 & 0 & e^{-2\pi i/3} \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.3.11}$$

$$A_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.3.12}$$

$$B_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{2\pi i/3} & 0 \\ 0 & 0 & e^{-2\pi i/3} \end{pmatrix}, \tag{4.3.13}$$

$$B_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \tag{4.3.14}$$

$$B_0' = \begin{pmatrix} 0 & e^{-2\pi i/3} & 0 \\ 0 & 0 & e^{2\pi i/3} \\ 1 & 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \tag{4.3.15}$$

$$B_1' = \begin{pmatrix} 0 & e^{-2\pi i/3} & 0 \\ 0 & 0 & e^{2\pi i/3} \\ 1 & 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 & 1 \\ e^{-2\pi i/3} & 0 & 0 \\ 0 & e^{2\pi i/3} & 0 \end{pmatrix}. \tag{4.3.16}$$

This is used in the numerical investigation in the next section.

Bound:		I_3	**Tsirelson type**	**TLM-type, l.h.s/r.h.s.**	**L.h.s. of (4.2.4)**	**L.h.s. of (4.2.5)**
Theoretical maximum		2.91 [152]	$2\sqrt{2}$ (≈ 2.83)	1	1	1 (if assumptions hold)
Parafermion maximum	non-signaling (A+B)	2.60	2.44	0.71	0.74	1.00
	signaling (A+B')	2.60	2.82	1.00	1.00	1.56

Table 4.1: Characterization of various bounds on non-local correlations for the signaling and non-signaling sets of parafermionic observables. Maximal values achieved with parafermionic observables ('Parafermion maximum') are compared to the theoretical maximum. The quantities considered are I_3 (4.3.17, l.h.s.), generalized Tsirelson (4.2.2, l.h.s.), generalized TLM (4.2.3, l.h.s/r.h.s.), relation (4.2.4, l.h.s.), and relation (4.2.5, l.h.s.).

4.3.3 Numerical results for correlations of parafermions

Here we numerically investigate the bounds on correlations presented above (4.2.2–4.2.5) and the CHSH-like inequality derived in Ref. [151]. The inequality of Ref. [151] states that for a local-realistic system

$$I_3 = Q_{00} + Q_{01} - Q_{10} + Q_{11} \leq 2, \tag{4.3.17}$$

where $Q_{jk} = \mathrm{Re}\left[\langle A_j B_k\rangle\right] + \frac{1}{\sqrt{3}}\mathrm{Im}\left[\langle A_j B_k\rangle\right]$ for $i \geq j$, and $Q_{01} = \mathrm{Re}\left[\langle A_0 B_1\rangle\right] - \frac{1}{\sqrt{3}}\mathrm{Im}\left[\langle A_0 B_1\rangle\right]$. The observables are assumed to have possible values (for the quantum case that we are interested in, eigenvalues) $e^{2\pi i r/3}$, $r \in \mathbb{Z}$, which is the case for the observables defined in Eqs. (4.3.11–4.3.16). In the standard quantum theory, i.e., for quantum observables such that $[A_j, B_k] = 0$, the maximum attainable value is known to be ≈ 2.91 [152].

For all the inequalities investigated, we calculated the corresponding correlations $C(A_i, B_j)$ or $\langle A_j B_k\rangle$, and maximized the relevant expressions numerically over all possible states $|\psi\rangle$. The expressions maximized were the left-hand side of bounds (4.3.17, 4.2.2, 4.2.4, 4.2.5) and the ratio of the left-hand side to the right-hand side of inequality (4.2.3). The numerical maximization was performed independently via Wolfram Mathematica (functions `NMaximize` for finding the global maximum and `FindMaximum` for investigating local maxima) and Python (package `scipy.optimize`, functions `basinhopping` for finding the global maximum with `SLSQP` optimization method for investigating local maxima). One aspect deserves

mentioning. Correlation functions $C(A_i, B_j)$ defined in Eq. (4.2.6) are not well-defined in all of the Hilbert space as the denominator can turn out to be zero. However, the points where it does, constitute a set of measure zero among all the states. Moreover, in the vicinity of these special points, $C(A_i, B_j)$ does not diverge but stays bounded as $|C(A_i, B_j)| \leq 1$; however, the limiting value as one approaches the special point depends on the direction of approach. Therefore, with careful treatment, these special points do not constitute a problem for investigation. Namely, we replaced $\Delta_{A_i} \to \Delta_{A_i} + \epsilon^2$, $\Delta_{B_j} \to \Delta_{B_j} + \epsilon^2$ with a small cutoff ϵ, and checked that our results do not change as $\epsilon \to 0$. Furthermore, the states $|\psi\rangle$ on which the maximum values in Table 4.1 are achieved are such that $\Delta_{A_i}, \Delta_{B_j} \neq 0$ for all A_i, B_j.

The results of our investigation are presented in Table 4.1. First, we note that the l.h.s. of Eq. (4.3.17) does not distinguish the signaling and non-signaling sets of observables. Second, our bounds (4.2.2-4.2.4) are obeyed by both sets. However, the signaling set saturates the bounds much better than the non-signaling one. Finally, the bound of Theorem 4, (4.2.5), is saturated by the non-signaling set and *violated* by the signaling one. This does not contradict the proof, which assumes $C(A_i, B_j) = (-1)^{ij}\varrho$. In fact, this property is not satisfied by the states $|\psi\rangle$ maximizing the l.h.s. of (4.2.5) for either of the sets. However, this numerical evidence together with the fact that $C(A_i, B_j) = (-1)^{ij}\varrho$ correlations might be special for non-signaling theories (cf. the discussion after Theorem 4) imply that Eq. (4.2.5) may be a good bound for distinguishing signaling and non-signaling quantum theories. We provide further evidence for the last statement in Appendix 4.D.

4.4 Discussion

Our analytic results have important implications for understanding quantum correlations. It is known that the standard CHSH parameter has distinct bounds for classical local ($|\mathcal{B}| \leq 2$) and non-local ($|\mathcal{B}| \leq 4$) hidden variable theories, while the standard quantum theory obeys the Tsirelson bound (4.1.1). Our variation of the Tsirelson bound (4.2.2) is closely related to the original Tsirelson bound. In particular, for Hermitian observables $X = A_i, B_j$ such that $XX^\dagger = 1$ and states $|\psi\rangle$ such that $\langle\psi|X|\psi\rangle = 0$, our Bell-CHSH parameter $|\mathcal{B}|$ (4.2.2) coincides with the original one. At the same time, our proof shows that the Tsirelson bound (4.2.2), as well as the TLM bound (4.2.3), do not distinguish between the standard and non-local signaling quantum theories. This implies that the Hilbert space

structure is much more restrictive than it was previously thought (see, e.g., Eq. 4.2.8 which underlies our proofs). Naively, one could expect that the possibility of signaling would allow nonlocal correlations to be stronger than quantum, because one party can directly affect from a distance the others' outcomes and in particular make them more correlated with hers. However, the limited kind of signaling we have introduced here, still within a quantum-like structure, is insufficient for this purpose.

At the same time, understanding the bounds on correlations in the standard quantum theory, that explicitly takes into account the absence of signaling, may be beneficial both for deepening its understanding, further testing its validity, and deriving bounds on protocols for quantum information processing. Our numerical results with parafermions provide a candidate for such a bound, Eq. (4.2.5). Indeed, the "non-signaling" parafermionic set stayed within the bound, while the "signaling" one violated it. Moreover, Ref. [148] argued that the assumptions we used to prove theorem 4 hold generally for the states maximizing the standard Bell-CHSH parameter in non-signaling theories (not in the sense that any maximizing state satisfies the assumptions, but in the sense that it is always possible to find a state that maximizes the standard Bell-CHSH parameter and satisfies the assumptions). Therefore, we believe that inequality (4.2.5) deserves further investigation.

4.A Relation between the correlation functions $C(X,Y)$ and joint probability distributions

For the standard case of commuting operators X and Y, it is possible to express correlators $C(X,Y)$ defined in Eq. (4.2.1) through the joint probability distribution $P(x,y)$ of outcomes of X and Y measurements. Indeed, for commuting X and Y, it is possible to find their common eigenbasis $|xy\lambda\rangle$, where $X|xy\lambda\rangle = x|xy\lambda\rangle$ and similarly for Y; λ represents possible additional quantum numbers. Then any state allows for a decomposition

$$|\psi\rangle = \sum_{x,y,\lambda} \alpha_{xy\lambda}|xy\lambda\rangle. \tag{4.A.1}$$

The probability of one observer obtaining x in a measurement of X, while the other obtains y in a measurement of Y is given by

$$P(x,y) = \mathrm{Tr}|\psi\rangle\langle\psi|\mathcal{P}_x^{(X)}\mathcal{P}_y^{(Y)} = \sum_{\lambda} |\alpha_{ab\lambda}|^2, \tag{4.A.2}$$

where $\mathcal{P}_x^{(X)}$ and $\mathcal{P}_y^{(Y)}$ are the projectors onto the eigenspaces of X and Y respectively. Then $\langle XY^\dagger\rangle = \sum_{x,y} xy^* P(x,y)$, $\langle X\rangle = \sum_x xP(x,y)$ etc. This allows for expressing $C(X,Y)$ as a non-linear functional of the probability distribution $P(x,y)$. Therefore, for the case of commuting Alice-Bob observables, $[A_i, B_j] = 0$ our bounds (4.2.2,4.2.3) can be considered restrictions on the possible joint probability distributions $P(a,b|i,j)$ in the quantum theory, defined exactly as in Eq. (4.A.2) modulo a replacement $X \to A_i$ and $Y \to B_j$.

For the case of non-commuting X and Y, one cannot define a joint eigenbasis, but rather eigenbases $|x\lambda\rangle$ of X and $|y\tilde{\lambda}\rangle$ of Y. One can still expand any state

$$|\psi\rangle = \sum_{x} \alpha_{x\lambda}|x\lambda\rangle \tag{4.A.3}$$

and define

$$\begin{aligned} W(x,y) &= \mathrm{Tr}|\psi\rangle\langle\psi|\mathcal{P}_x^{(X)}\mathcal{P}_y^{(Y)} \\ &= \sum_{x',\lambda,\lambda',\tilde{\lambda}} \alpha_{x'\lambda'}\alpha_{x\lambda}^*\langle x\lambda|y\tilde{\lambda}\rangle\langle y\tilde{\lambda}|x'\lambda'\rangle. \end{aligned} \tag{4.A.4}$$

Moreover, $\langle XY^\dagger\rangle = \sum_{x,y} xy^* W(x,y)$, $\langle X\rangle = \sum_x xW(x,y)$ etc., leading to exactly the same expression of $C(X,Y)$ in terms of $W(x,y)$ as previously

in terms of $P(x,y)$. However, $W(x,y)$ is not a probability distribution as the r.h.s. of Eq. (4.A.4) can acquire complex values. $W(x,y)$ is a quasiprobability distribution (somewhat similar to the Wigner function) in the case of non-commuting X and Y. Therefore, when $[A_i, B_j] \neq 0$ can be considered as restrictions on the possible joint quasiprobability distributions $W(a,b|i,j)$.

4.B Measuring parafermionic observables

A system combining parafermions with charging energy was introduced in Ref. [150]. In such a system there is a charging energy associated with the total system charge $Q_{\rm tot} = \sum_j Q_j + Q_0$, where $Q_0 = 2en_C$ is the charge of the proximitizing superconductor, and n_C is the number of Cooper pairs in it. However, no energy cost is associated with different distributions of a given total charge over different SC domains. Therefore, the ground state of such a system has degeneracy $d^{N_{\rm SC}-1}$, where the reduction by a factor of d corresponds to fixing the system's total charge. The properties of operators $\alpha_{s,j}\alpha^\dagger_{s,k}$ acting in this reduced subspace are identical to those in the original system of parafermions with unrestricted total charge.

Introducing charging energy allows for designing a relatively simple protocol for measuring $\alpha_{s,j}\alpha^\dagger_{s,k}$ (both parafermions have the same s!) [150]. A sketch of the measurement setup is shown in Fig. 4.1b. Two additional FQH edges (belonging to one of the puddles) are required in this setup. Tunneling of FQH quasiparticles is allowed directly between the two edges with tunnelling amplitude $\eta_{\rm ref}$ or between each edge and the corresponding parafermion $\alpha_{s,j/k}$ with amplitude $\eta_{j/k}$. As changing the charge of the parafermionic system is energetically costly, the leading non-trivial process resulting from coupling of the edges to the parafermions is co-tunneling of quasiparticles: a quasiparticle is transferred between the edges, while the parafermion state is changed via $\alpha_{s,j}\alpha^\dagger_{s,k}$ and the effective tunneling amplitude is $\eta_{\rm cot} \simeq -\eta_k\eta_j^*/E_C$, where E_C is the charging energy. The two processes, direct and parafermion-mediated tunneling of a quasiparticle between the edges, interfere quantum-mechanically. When a voltage bias V is applied between the edges, the tunneling current between the edges is sensitive to this interference:

$$I_{\rm T} \propto |V|^{2\nu-1}\,{\rm sgn}V \times \left(|\eta_{\rm ref}|^2 + |\eta_{\rm cot}|^2 + 2\kappa{\rm Re}\left[\eta^*_{\rm ref}\eta_{\rm cot}\alpha_{s,j}\alpha^\dagger_{s,k}\right]\right), \quad (4.B.1)$$

where κ is the interference suppression factor due to finite temperature

and other effects, $\mathrm{Re}\,[A] = \left(A + A^\dagger\right)/2$, and $|V|$ is assumed to be much larger than the temperature T of the probing edges. As a result, by measuring I_T, one can measure the operator $\mathrm{Re}\left[e^{i\varphi}\alpha_{s,j}\alpha_{s,k}^\dagger\right]$ with phase φ depending on the phases of η_ref and η_cot. Thus, one can measure the system in the eigenstates of $\alpha_{s,j}\alpha_{s,k}^\dagger$ employing the fact that the eigenvalues of the $\alpha_{s,j}\alpha_{s,k}^\dagger$ are discrete: for a generic φ, distinct eigenvalues of $\alpha_{s,j}\alpha_{s,k}^\dagger$ correspond to distinct eigenvalues of $\mathrm{Re}\left[e^{i\varphi}\alpha_{s,j}\alpha_{s,k}^\dagger\right]$. Alternatively, through tuning the phase φ, one can measure independently $\mathrm{Re}\left[\alpha_{s,j}\alpha_{s,k}^\dagger\right]$ and $\mathrm{Im}\left[\alpha_{s,j}\alpha_{s,k}^\dagger\right] = \mathrm{Re}\left[e^{-i\pi/2}\alpha_{s,j}\alpha_{s,k}^\dagger\right]$, and combine the measurement results for calculating the expectation value $\langle\alpha_{s,j}\alpha_{s,k}^\dagger\rangle$.

4.C How to measure correlations of non-commuting observables

4.C.1 Measuring correlations of non-commuting parafermionic observables

Here we discuss how one can measure the correlators $\langle A_j\,(B'_k)^\dagger\rangle$ for non-commuting parafermionic observables. The procedure outlined in Appendix 4.C.2 enables one to measure $\langle\{A_j,(B'_k)^\dagger\}\rangle$, where $\{X,Y\}$ denotes the anti-commutator of operators X and Y, using weak measurements [67, 68]. For the observables defined in Sec. 4.3.2, the following permutation relation holds: $A_j\,(B'_k)^\dagger = (B'_k)^\dagger\,A_j e^{-2i\pi\nu}$. Therefore, $\langle\{A_j,(B'_k)^\dagger\}\rangle = \langle A_j\,(B'_k)^\dagger\,(1+e^{2i\pi\nu})\rangle = 2\langle A_j\,(B'_k)^\dagger\rangle e^{i\pi\nu}\cos\pi\nu$, and measuring $\langle\{A_j,(B'_k)^\dagger\}\rangle$ is sufficient for measuring $\langle A_j\,(B'_k)^\dagger\rangle$.

The rest of this appendix is dedicated to designing weak measurements of the required type and adapting the protocol of Appendix 4.C.2 to measuring parafermionic observables. Note that this measurement method is specific to the particular implementation of parafermions. We start with the measurement protocol discussed in Appendix 4.B. Suppose one of the additional FQH edges involved in the protocol has voltage V applied to it, while the other edge is grounded. The current injected to the first edge is $I_\mathrm{in} = \nu e^2 V/h$, while the tunneling current between the edges is I_T, cf. Eq. (4.B.1). Suppose one measures the current for time t, so that the number of quasiparticles injected into the system is $N = I_\mathrm{in}t/(\nu e)$. The number of quasiparticles q tunneling within the time window will be fluctuating around the average $\langle q\rangle = pN = I_\mathrm{T}t/(\nu e)$ with $p = I_\mathrm{T}/I_\mathrm{in}$. The

expression for I_{T} in Eq. (4.B.1) is valid as long as $|I_{\mathrm{T}}| \ll |I_{\mathrm{in}}|$. In this regime, tunneling of different quasiparticles can be considered independent, and thus the probability of observing tunneling of q quasiparticles should be approximated well by the binomial distribution

$$P(q) = C_N^q p^q (1-p)^{N-q}, \quad C_N^q = \frac{N!}{q!(N-q)!}. \tag{4.C.1}$$

If one measures for a sufficiently long time, i.e., $N \gg 1$, the binomial distribution is well-approximated by the Gaussian distribution

$$P(q) \approx \frac{1}{\sqrt{2\pi N p(1-p)}} \exp\left(-\frac{(q-pN)^2}{2Np(1-p)}\right). \tag{4.C.2}$$

Depending on the eigenvalue $-e^{i\pi\nu(r+1/2)}$ of the measured observable $\alpha_{s,j}\alpha^{\dagger}_{s,k}$, the tunneling probability $p = p_0 + \delta p_r$, with

$$p_0 \propto |\eta_{\mathrm{ref}}|^2 + |\eta_{\mathrm{cot}}|^2, \tag{4.C.3}$$

$$\delta p_r \propto -2\kappa\, |\eta_{\mathrm{ref}}|\, |\eta_{\mathrm{cot}}| \cos\left(\pi\nu r + \pi\nu/2 + \varphi\right), \tag{4.C.4}$$

where $\varphi = \arg(\eta^*_{\mathrm{ref}}\eta_{\mathrm{cot}})$, cf. Eq. (4.B.1). From now on we assume $|\eta_{\mathrm{cot}}| \ll |\eta_{\mathrm{ref}}|$, $p_0 \ll 1$ and $p_0 N \gg 1$. Then the average number of tunneled quasiparticles is $\langle q\rangle_r = p_0 N + \delta p_r N$, while the size of fluctuations in the measured values of q is of the order σ:

$$\sigma = \sqrt{2Np(1-p)} = \sqrt{2Np_0}\left(1 + O(|\eta_{\mathrm{cot}}/\eta_{\mathrm{ref}}|, p_0)\right).$$

The parameter determining the distinguishability of different r, and thus the measurement strength, is $\delta p_r N/\sigma \propto \left|\frac{\eta_{\mathrm{cot}}}{\eta_{\mathrm{ref}}}\right| \sqrt{N}$. For sufficiently large $\left|\frac{\eta_{\mathrm{cot}}}{\eta_{\mathrm{ref}}}\right| \sqrt{N}$, the scheme thus implements a strong measurement, while the limit $\left|\frac{\eta_{\mathrm{cot}}}{\eta_{\mathrm{ref}}}\right| \sqrt{N} \ll 1$ implies a weak measurement.

Denoting the initial state of parafermions as $\sum_r \psi_r |r\rangle$ and using some approximations, one can derive the state of the system after switching on the tunnel couplings for time t,

$$|\tilde{\Phi}\rangle = \sum_{r,q,\lambda} f_\lambda(q,r)\psi_r \left(\frac{\eta_{\mathrm{ref}} - \eta_{\mathrm{cot}} e^{i\pi\nu(r+1/2)}}{\left|\eta_{\mathrm{ref}} - \eta_{\mathrm{cot}} e^{i\pi\nu(r+1/2)}\right|}\right)^q |r\rangle |q,\lambda\rangle, \tag{4.C.5}$$

where λ represents additional quantum numbers of the edges. It follows from Eq. (4.C.2) that:

$$\sum_\lambda f^*_\lambda(q,r) f_\lambda(q,r) = \mathcal{N}^2 \exp\left[-\frac{(q-\langle q\rangle_r)^2}{2Np_0}\right]\left[1 + O\left(\left|\frac{\eta_{\mathrm{cot}}}{\eta_{\mathrm{ref}}}\right|, p_0\right)\right]$$

, with normalization factor $\mathcal{N} = (2\pi N p_0)^{-1/4}$. Having not performed the calculation, we make a plausible assumption that also

$$\sum_\lambda f_\lambda^*(q,r) f_\lambda(q,r') = \mathcal{N}^2 \times \exp\left[-\left(q - \frac{\langle q\rangle_r + \langle q\rangle_{r'}}{2}\right)^2 \times \frac{1}{2Np_0} - \frac{(\langle q\rangle_r - \langle q\rangle_{r'})^2}{8Np_0}\right] \times \left[1 + O\left(\left|\frac{\eta_{\rm cot}}{\eta_{\rm ref}}\right|, p_0\right)\right]. \quad (4.C.6)$$

Further assuming the limit $\left|\frac{\eta_{\rm cot}}{\eta_{\rm ref}}\right| p_0 N \ll 1$, we can neglect $\eta_{\rm cot} e^{i\pi\nu(r+1/2)}$ in Eq. (4.C.5) and obtain that for our purposes one can replace $|\tilde{\Phi}\rangle$ with

$$|\Phi\rangle = \mathcal{N}\sum_{r,q} \psi_r \exp\left[-\frac{(q - \langle q\rangle_r)^2}{4Np_0}\right] \times \left[1 + O\left(\left|\frac{\eta_{\rm cot}}{\eta_{\rm ref}}\right| p_0 N, \left|\frac{\eta_{\rm cot}}{\eta_{\rm ref}}\right|, p_0\right)\right] |r\rangle|q\rangle, \quad (4.C.7)$$

which brings us to weak measurements of the type considered in Appendix 4.C.2.

Consider now two weak measurements accessing A_j and $(B'_k)^\dagger$ performed one after the other, with the number of quasiparticles tunneled in each of the measurements being q_1 and q_2. Repeating the calculation of Appendix 4.C.2, we obtain

$$\langle (q_1 - p_0 N)(q_2 - p_0 N)\rangle \propto \langle \mathrm{Re}\left[e^{i\varphi} A_j\right] \mathrm{Re}\left[e^{i\varphi'} (B'_k)^\dagger\right]\rangle \times \left[1 + O\left(\left|\frac{\eta_{\rm cot}}{\eta_{\rm ref}}\right| p_0 N, \left|\frac{\eta_{\rm cot}}{\eta_{\rm ref}}\right|, p_0\right)\right]. \quad (4.C.8)$$

Using Eq. (4.C.16), one sees that by choosing different phases φ, φ', one can measure $\langle\{A_j, (B'_k)^\dagger\}\rangle = 2\langle A_j (B'_k)^\dagger\rangle e^{i\pi\nu}\cos\pi\nu$.

4.C.2 Measuring correlations of non-commuting observables with weak measurements

Here we discuss how to measure the averages $\langle\{A, B\}\rangle$ of non-Hermitian non-commuting A and B, where $\{A, B\} = AB + BA$, with the help of weak measurements. Our protocol uses essentially the same measurement

procedure as in Refs. [153–155], and is similar in spirit (yet has important differences) to Refs. [156, 157]. We note in passing that by more elaborate methods, one can measure also the expectation value of a commutator [145]. However, measuring the anti-commutator will suffice for our purposes. We first discuss how to measure correlations of Hermitian non-commuting observables, and then generalize the scheme to non-Hermitian observables.

Suppose one wants to measure the average $\langle\{A,B\}\rangle = \langle\psi|\{A,B\}|\psi\rangle$, where A and B are *Hermitian* non-commuting operators, and $|\psi\rangle$ is some quantum state. Introduce the eigenbases of A and B: $A|a\rangle = a|a\rangle$, $B|b\rangle = b|b\rangle$. Any system state $|\psi\rangle$ can then be written as $|\psi\rangle = \sum_a \psi_a|a\rangle = \sum_{a,b}\psi_a|b\rangle\langle b|a\rangle$ with some coefficients ψ_a. We assumed that there is no degeneracy in the spectra of A and B; generalization of the below consideration for the case with degeneracy is straightforward.

Consider two detectors, D_1 and D_2 each having coordinate Q_j and momentum P_j operators, $[P_j, Q_k] = -i\delta_{jk}$, with j and k having values 1 and 2. Prepare the system and detectors in initial state

$$|\Phi_{in}\rangle = |\psi\rangle|D_{1,in}\rangle|D_{2,in}\rangle, \tag{4.C.9}$$

$$|D_{j,in}\rangle = \mathcal{N}\int dq_j \exp\left(-\frac{q_j^2}{2\sigma^2}\right)|q_j\rangle, \tag{4.C.10}$$

where $|q_j\rangle$ is an eigenstate of Q_j with eigenvalue q_j, and $\mathcal{N} = \left(\pi\sigma^2\right)^{-1/4}$.

The Hamiltonian describing the system and the detectors is

$$H(t) = \lambda_1(t)H_1 + \lambda_2(t)H_2, \tag{4.C.11}$$

$$H_1 = P_1 A, \quad H_2 = P_2 B, \tag{4.C.12}$$

where the coupling constants $\lambda_j(t) = 0$ except for $\lambda_1(t) = g/T$ for $t \in (0;T)$ and $\lambda_2(t) = g/T$ for $t \in (T;2T)$. Then after the system has interacted with the detectors, their state is

$$\begin{aligned}|\Phi\rangle = e^{-igH_2}e^{-igH_1}|\Phi_{in}\rangle \\ = \mathcal{N}^2\sum_{a,b}\int dq_1 dq_2 \psi_a\langle b|a\rangle|b\rangle|q_1\rangle|q_2\rangle \\ \times\exp\left(-\frac{(q_1-ga)^2}{2\sigma^2} - \frac{(q_2-gb)^2}{2\sigma^2}\right).\end{aligned} \tag{4.C.13}$$

Measuring Q_1 and Q_2 of the detectors and calculating their correlations then yields the desired quantity. Indeed,

$$\begin{aligned}\langle\Phi|Q_1Q_2|\Phi\rangle &= \mathcal{N}^4 \sum_{a,a',b} \psi_a^*\psi_{a'}\langle a|b\rangle\langle b|a'\rangle \\ &\quad\times \int dq_2 q_2 \exp\left(-\frac{(q_2-gb)^2}{\sigma^2}\right) \\ &\quad\times \int dq_1 q_1 \exp\left(-\frac{(q_1-g(a+a')/2)^2}{\sigma^2} - \frac{g^2(a-a')^2}{4\sigma^2}\right) \\ &= \sum_{a,a',b} \alpha_a^*\alpha_{a'}\langle a|b\rangle\langle b|a'\rangle \frac{g^2}{2} b(a+a') \exp\left(-\frac{g^2(a-a')^2}{4\sigma^2}\right). \quad (4.C.14)\end{aligned}$$

Provided that $g\,|a-a'| \ll 2\sigma$ for all a, a' (which is the condition for weakness of the measurement), one obtains

$$\begin{aligned}&\langle\Phi|Q_1Q_2|\Phi\rangle \\ &\qquad= \frac{g^2}{2}\sum_{a,a',b} \psi_a^*\langle a|\,(a|b\rangle b\langle b| + |b\rangle b\langle b|a')\,\psi_{a'}|a'\rangle \\ &\qquad= \frac{g^2}{2}\langle\psi|\{A,B\}|\psi\rangle. \quad (4.C.15)\end{aligned}$$

Suppose now one wants to measure $\langle\{A,B\}\rangle = \langle\psi|\{A,B\}|\psi\rangle$ for *non-Hermitian* A and B. Define the real and imaginary part of each operator: $R_A = (A+A^\dagger)$, $I_A = i(A^\dagger - A)/2$, and similarly for B. It is easy to see that $\{A,B\} = \{R_A,R_B\} - \{I_A,I_B\} + i\{I_A,R_B\} + i\{R_A,I_B\}$. Then

$$\begin{aligned}\langle\{A,B\}\rangle = \langle\{R_A,R_B\}\rangle - \langle\{I_A,I_B\}\rangle \\ + i\langle\{I_A,R_B\}\rangle + i\langle\{R_A,I_B\}\rangle. \quad (4.C.16)\end{aligned}$$

Each of the averages in the r.h.s. can be measured using the protocol for Hermitian observables outlined above. Then combining them according to Eq. (4.C.16) yields the desired correlation of non-Hermitian non-commuting observables.

4.D Extra numerical data on the bounds for correlations in the system of parafermions

In the main text, Table 4.1, we provided the results of testing the bounds on correlations for two sets of observables, non-signaling (A_0, A_1, B_0, B_1) and signaling (A_0, A_1, B'_0, B'_1). Here, in Table 4.2, we present the results for several more sets of observables. Namely, we checked what happens when the roles of Alice's operators A_0 and A_1 are exchanged, and similarly for Bob. Apart from that, we also tested the sets involving $B_2 = B_0^\dagger B_1 = \alpha_{L,5}\alpha_{L,6}^\dagger$ and $B'_2 = B_0'^\dagger B'_1 = \alpha_{R,6}\alpha_{R,5}^\dagger$; $[B_2, A_j] = [B'_2, A_j] = 0$, with A_j, B_j, B'_j defined in Eqs. (4.3.11–4.3.16). In all the sets we tested, all Alice's and Bob's operators commute when Bob uses unprimed observables; some of Alice's operators do not commute with some of the Bob's observables when Bob uses primed observables, B'_j.

Note that all the sets we have tested obey all bounds except for relation (4.2.5). The latter is obeyed by all the non-signaling sets (when Bob uses B_j observables) and violated by all the signaling sets (when Bob uses B'_j observables). This strengthens the numerical evidence that relation (4.2.5) is a good candidate for quantifying the effect of signaling on quantum correlations.

In principle, the system of parafermions has many more possible sets of observables. First, assigning different parafermions to Alice and Bob, one can have different local algebras at Alice's and Bob's sites, as well as different Alice-Bob commutation relations. We investigate them in part by switching the order of A_0 and A_1 etc. or replacing B_1 with B_2 in Table 4.2. While this does not exhaust all the possibilities, the numerical results we do have, indicate that our conclusions are likely to hold in the cases we did not check. An even richer set of algebras can be accessed by using operators beyond quadratic in parafermions, e.g., $\left(\alpha_{s,j}\alpha_{s,k}^\dagger\right)^n$ or $\alpha_{s,j}^2\alpha_{s,k}^\dagger\alpha_{s,l}^\dagger$, as well as arbitrary linear combinations of quadratic operators, e.g., $x\alpha_{s,j}\alpha_{s,k}^\dagger + y\alpha_{s,j}\alpha_{s,l}^\dagger$. While investigating our bounds with these would be an interesting non-trivial check, we believe that the more important task is understanding and proving the role of Theorem 4 and bound (4.2.5) in the general context.

Bound:			I_3	**Tsirelson type**	**TLM-type, l.h.s/r.h.s.**	**L.h.s. of (4.2.4)**	**L.h.s. of (4.2.5)**
Theoretical maximum			2.91 [152]	$2\sqrt{2}$ ($\approx$ 2.83)	1	1	**1 (if assumptions hold)**
Parafermion maximum	**Alice's operators**	**Bob's operators**					
	A_0, A_1	B_0, B_1	2.60	2.44	0.71	0.74	1.00
	A_0, A_1	B'_0, B'_1	2.60	2.82	1.00	1.00	1.56
	A_1, A_0	B_1, B_0	2.60	2.44	0.71	0.74	1.00
	A_1, A_0	B'_1, B'_0	2.60	2.82	1.00	1.00	1.56
	A_0, A_1	B_1, B_0	2.60	2.22	0.71	0.62	1.00
	A_0, A_1	B'_1, B'_0	2.60	2.71	1.00	0.97	1.56
	A_0, A_1	B_0, B_2	2.60	2.44	0.71	0.74	1.00
	A_0, A_1	B'_0, B'_2	2.00	2.23	1.00	0.75	1.50
	A_0, A_1	B_2, B_0	2.60	2.22	0.71	0.62	1.00
	A_0, A_1	B'_2, B'_0	2.00	2.44	1.00	0.75	1.50

Table 4.2: Characterization of bounds on non-local correlations for various sets of parafermionic observables. Similarly to the Table 4.1, the maximal values achieved with parafermionic observables ('Parafermion maximum') are compared to the theoretical maximum.

5 A diagrammatic approach to variational quantum ansatz construction

5.1 Introduction

Despite promises of exponential speedups, quantum algorithms require optimization to achieve an advantage over their classical counterparts on state of the art supercomputers for problems of interest. This is the case both in the Noisy Intermediate-Scale Quantum era [11], where coherence times in quantum devices prohibit all but the shortest experiments to be performed, and in first-generation fault-tolerant devices, where a single non-Clifford rotation requires thousands of additional qubits and hundreds of error correcting cycles [158]. In the field of digital quantum simulation, the variational quantum eigensolver (VQE) [159] has emerged as a competitive class of algorithms for generating approximate ground states of quantum systems, due to its relatively low circuit length. These algorithms consist of parametrizing a quantum circuit with a small number of classical control variables, which may be tuned to minimize the energy of the state produced by the circuit, given a target Hamiltonian. As the manifold of obtainable states for a given VQE will only ever be an exponentially small region in the larger Hilbert space, optimizing VQE design is critical to obtain good approximations of the system's ground state [47, 47, 54]. This has spurred much recent work in optimizing VQEs based on the unitary coupled cluster expansion [47, 54, 161], or on the quantum approximate optimization algorithm [162, 163]. The efficiency of coupled cluster methods is based on the principle of size-extensivity. This means that the ansatz systematically accounts for ground state correlations, as ensured in perturbative language by the linked-cluster theorem [57]. However, to be realized as a quantum circuit size-extensive ansatzes typically require expansion via Trotter-Suzuki-based methods [58, 59]. At low circuit depth, these expansions introduce significant errors. Alleviating this issue would help to ensure the efficiency of the VQE algorithm.

In this chapter, we develop a Trotterization-free diagrammatic method to generate size-extensive VQEs. We start by designing a class of VQE ansatzes, based on the stabilizer formalism in quantum error correction, which provably tightly span the entire Hilbert space of N_q qubits. We then demonstrate how one may compress an arbitrary variational ansatz to account for symmetries of a target Hamiltonian. We further show how to construct a hierarchy of ansatz generators, allowing one to trade between circuit length and accuracy in a practical manner by choosing only those generators that contribute well to solving the problem. We motivate the construction of one particular such hierarchy from a general perturbative analysis of weakly coupled target Hamiltonians, for which we develop a simple-to-use diagrammatic formalism. We find that our geometrically tight stabilizer ansatz may be compressed to a practical size using this perturbative scheme. The analogue of the linked-cluster theorem for such compressed digital ansatzes is stated and proven, ensuring the size-extensivity of the construction. We also propose some possible modifications to our perturbative scheme to account for circuit depth and locality. We compare the performance of these constructions on simulations of the transverse-field Ising model in three different physical regimes (weak-coupling, strong-coupling, and critical). We find that strictly following the perturbative approach is beneficial in the weak-coupling regime, but restricting the ansatz to lowest-order gives better convergence in the strong-coupling regime — even though such ansatzes are seemingly less-informed about the strong-coupling physics.

5.2 Variational quantum eigensolvers

A variational quantum eigensolver (VQE) is an algorithm executed on a quantum register that aims to approximate the minimum eigenvalue E_0 of a target Hamiltonian H on $\mathbb{C}^{2^{N_q}}$ by finding low energy states $|\psi\rangle \in \mathbb{C}^{2^{N_q}}$ variationally. To be precise, this algorithm minimizes $\langle\psi|H|\psi\rangle$ over a variational ansatz:

Definition 1. *A variational ansatz on N_p parameters corresponds to a pair $(U, |\vec{0}\rangle)$, where U is a smooth map from the* ***parameter space*** *$\vec{\theta} \in \mathbb{R}^{N_p}$ to the unitary operator $U(\vec{\theta})$ on $\mathbb{C}^{2^{N_q}}$, and $|\vec{0}\rangle \in \mathbb{C}^{2^{N_q}}$ is the* ***starting state****, which is acted on to generate the* ***variational state*** *$|\psi(\vec{\theta})\rangle = U(\vec{\theta})|\vec{0}\rangle$, with variational energy $E(\vec{\theta}) = \langle\psi(\vec{\theta})|H|\psi(\vec{\theta})\rangle$.*

As a brief example, let us define the following toy two-qubit variational ansatz:

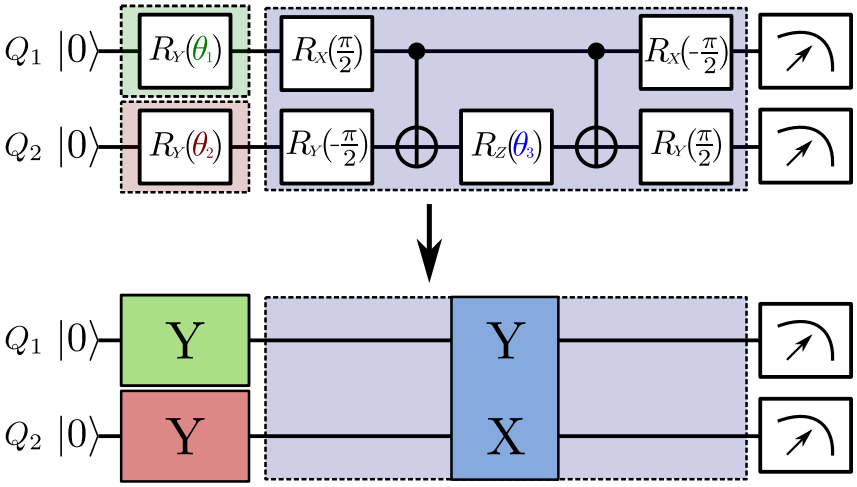

Figure 5.1: (top) A circuit to implement the YYX two-qubit variational ansatz in terms of Pauli rotations $R_A(\theta) = e^{iA_j\theta}$ on single qubits j ($A = X, Y, Z$) and CNOT gates. Shaded regions denote subcircuits to implement the three separate unitary rotations in Eq. (5.2.1), as color coded with the variational parameters θ_i. (below) The above circuit in a compressed notation, treating each rotation U as a single gate labeled by the elements of the rotation generators (Eq. (5.2.4)) on each qubit.

Example 2. *The 3-parameter* ***YYX*** *variational ansatz* $(U_{YYX}, |00\rangle)$ *is defined on two qubits* $\{Q_1, Q_2\}$*, with the starting state* $|00\rangle$ *in the computational (Z) basis, and*

$$U_{YYX}(\theta_1, \theta_2, \theta_3) := e^{i\theta_3 Y_1 X_2} e^{i\theta_2 Y_2} e^{i\theta_1 Y_1}. \tag{5.2.1}$$

A quantum circuit that implements this toy ansatz is given in Fig. 5.1, using standard methods [60] to decompose the two-qubit $e^{i\theta_2 Y_1 X_2}$ term in terms of single-qubit rotations and CNOT gates.

VQEs are appealing because they reduce the computational complexity of searching the (exponentially large) N_q-qubit Hilbert space to the complexity of searching the parameter space (which may be made arbitrarily small). However, this comes at a cost, as none of the states $|\psi(\vec{\theta})\rangle$ may be close (in energy or overlap) to the target ground state. The variance in the energy $\langle\psi|H|\psi\rangle$ of states $|\psi\rangle$ randomly drawn (i.e. with Haar measure)

from an N_q-qubit Hilbert space is given by

$$\frac{\|H\|_F^2 - \text{Trace}[H]^2}{4^{N_q}} \leq \frac{\|H - \text{Trace}[H]\|_S^2}{2^{N_q}}, \tag{5.2.2}$$

with $\|\cdot\|_F$ the Frobenius norm and $\|\cdot\|_S$ the spectral norm. This implies that the probability of a random state having energy close to the ground state energy of H scales as $e^{-2^{N_q}}$, while one expects the volume of space explored by a VQE to grow only as e^{N_p}. This, and similar results for derivatives of the energy with respect to variational parameters [160], imply that random ansatz choice has little to no chance of success for finding ground state energies. Instead, a variational ansatz should be designed to cover as much of the N_q-qubit Hilbert space as possible, in a way that maximises the chance of finding low-energy states (or states that overlap well with the true ground state).

A full VQE protocol must also concern itself with optimizing the minimization procedure, especially to prevent being stuck in local minima or barren plateaus [160]. One should further take care to make the resulting quantum circuit as hardware efficient [164, 165] as possible. Hardware-efficiency is an active field of research and dependent upon the physical implementation of the quantum computer, and recent work has gone into optimizing the minimization procedure of a VQE [54, 166], including the choice of cost function to minimize (e.g. to target excited states [167, 168]). In this chapter, we focus instead on studying the variational ansatzes themselves. We first focus on constructing 'geometrically efficient' variational ansatzes. Then we tailor these to target specific Hamiltonians based on a perturbative approach. This generic approach is in complement with previous work on ansatz design targeting specific (classically hard) problems of interest in e.g. optimization [162] and quantum chemistry [47].

To pin down a working definition of 'fundamentally digital' quantum ansatzes, we will use the following conditions (similar to those stated in [54, 160, 166, 169]):

Definition 3. *A variational ansatz* $(U, |\vec{0}\rangle)$ *is a* ***product ansatz*** *if it is a product of units* U_i*,*

$$U(\vec{\theta}) = \prod_{i=1}^{N_u} U_i(\theta_{n_i}), \tag{5.2.3}$$

where each U_i *has a generator* T_i*:*

$$U_i(\theta_{n_i}) = e^{iT_i\theta_{n_i}}. \tag{5.2.4}$$

If $n_i > n_j$ whenever $i > j$, we call the ansatz ***ordered****, and if each generator is a Pauli operator - $T_i \in \mathbb{P}^{N_q} := \{I, X, Y, Z\}^{\otimes N_q}$ - we call the ansatz a* ***Pauli-type ansatz****.*

We take the product in Eq. (5.2.3) from right to left (i.e. $U_1(\theta_{n_1})$ acts first on the state $|\vec{0}\rangle$). As we allow $n_i = n_j$ when $i \neq j$, we may have strictly more unitaries than parameters: $N_u \geq N_p$. In the rest of the chapter, we will refer to Pauli-type ansatzes as fundamentally digital: note that Pauli rotations can be directly implemented in a quantum circuit via the techniques of [60]. When used in a VQE, Pauli-type ansatzes also have the advantage that some derivatives of the variational energy may be obtained 'for free' [169].

Example 4. *The YYX toy ansatz is a Pauli-type ansatz, with generators $T_1 = Y_1$, $T_2 = Y_2$, $T_3 = Y_1X_2$.*

5.2.1 Variational manifolds

Although tailoring a VQE to a Hamiltonian is essential for its success [160], interesting statements may be made about the variational ansatz prior to fixing such a target, by focusing on the manifold of states it explores.

Definition 5. *The* ***variational manifold*** *$\mathcal{M}(U, |\vec{0}\rangle)$ of a variational ansatz $(U, |\vec{0}\rangle)$ is the set $\{|\psi(\vec{\theta})\rangle = U(\vec{\theta})|\vec{0}\rangle, (\vec{\theta}) \in \mathbb{R}^{N_p}\} \subset \mathbb{C}^{2^{N_q}}$.*

We note that, despite being a 'manifold generated by unitary rotations', $\mathcal{M}(U, |\vec{0}\rangle)$ does not have a structure of a Lie group. This is because we only apply U once to create the variational state; a state $U(\vec{\theta})U(\vec{\theta'})|\vec{0}\rangle$ may not correspond to any state $U(\vec{\theta''})|\vec{0}\rangle$ (and most often will not). If U is a product ansatz, one can defined a Lie group $\mathcal{L}(U) \subset \mathrm{U}(2^{N_q})$ from the set of generators T_i. The manifold $\mathcal{L}(U)|\vec{0}\rangle$ then contains $\mathcal{M}(U, |\vec{0}\rangle)$ as a submanifold, though it is almost always larger. Indeed, when $e^{i\theta T_i}$ defines a universal gate set, $\mathcal{L}(U) = \mathrm{U}(2^{N_q})$ and $\mathcal{L}(U)|\vec{0}\rangle$ is the entire set of N_q-qubit states, which is not terribly informative about the structure of $\mathcal{M}(U, |\vec{0}\rangle)$.

As a rough guide, the bigger the variational manifold the better; simply adding more manifold to an ansatz can never shift it further from the target ground state. However, measuring the size of a variational manifold is made somewhat difficult by dimensionality concerns. The (real) dimension $D_{\mathcal{M}(U,|\vec{0}\rangle)}$ of $\mathcal{M}(U, |\vec{0}\rangle)$ is at most N_p, but it may not achieve this upper bound, and $\mathcal{M}(U, |\vec{0}\rangle)$ may contain boundary regions of lower dimension. (Curiously, the minimal subspace of $\mathbb{C}^{2^{N_q}}$ containing $\mathcal{M}(U, |\vec{0}\rangle)$ may be of

much higher dimension than N_p.) As $\mathcal{M}(U, |\vec{0}\rangle)$ inherits a metric from $\mathbb{C}^{2^{N_q}}$, one can use this to define a Borel measure $d|\psi\rangle$, and thus define the area of the manifold:

$$A_{\mathcal{M}(U,|\vec{0}\rangle)} = \int_{\mathcal{M}(U,|\vec{0}\rangle)} d|\psi\rangle. \tag{5.2.5}$$

When the map $(\vec{\theta}) \to |\psi(\vec{\theta})\rangle$ is invertible on some range of parameters, its Jacobian J is full-rank, and the manifold area may be calculated as

$$A_{\mathcal{M}(U,|\vec{0}\rangle)} = \int d^{N_p}\theta \sqrt{\det(J^\dagger J)}. \tag{5.2.6}$$

However, when evaluating this integral one must take care to avoid double-counting points $\vec{\theta} \neq \vec{\theta}'$ when $|\psi(\vec{\theta})\rangle = |\psi(\vec{\theta}')\rangle$.

Example 6. *For the YYX toy ansatz, one may calculate*

$$J^\dagger J = \begin{pmatrix} 1 & 0 & -\sin(2\theta_2) \\ 0 & 1 & 0 \\ -\sin(2\theta_2) & 0 & 1 \end{pmatrix}. \tag{5.2.7}$$

The variational manifold $\mathcal{M}(U_{YYX}, |00\rangle)$ double-covers the Hilbert space, as

$$|\psi(\theta_3 - \pi/2, \pi/2 - \theta_2, \theta_1 - \pi/2)\rangle = |\psi(\theta_3, \theta_2, \theta_1)\rangle \tag{5.2.8}$$

(no other identifications exist). Following this identification, one can evaluate $A_{\mathcal{M}(U_{YYX},|\vec{0}\rangle)} = \pi^2$.

5.3 Stabilizer ansatzes

Clearly the largest space that can be spanned by any variational ansatz is the entire Hilbert space. The minimal number of (real) parameters required to achieve this spanning is $2(2^{N_q} - 1)$, and it is an interesting question whether this may be provably achieved. In this section we answer this question in the affirmative, constructing a class of ansatzes from sequential layers of $n = 1, \ldots, N_q$-qubit stabilizer groups [170] (defined in App. 5.A). Although such a construction has impractically large overhead, one may use this construction as a base to generate tractable variational ansatzes with the methods developed in Sec. 5.4 and Sec. 5.5.

Definition 7. *A **stabilizer ansatz** $(U, |\vec{0}\rangle)$ on N_q qubits is constructed by choosing for each $n = 1, \ldots, N_q$:*

1. *A $[n-1, n-1]$ stabilizer group $\mathcal{S}^{(n)}$, and*
2. *A single-qubit starting state $|s_n\rangle$ for the n-th qubit, and*
3. *Two single-qubit Pauli operators $R_0^{(n)}, R_1^{(n)}$, such that $\langle s_n|R_i|s_n\rangle = 0$, and $\mathrm{Trace}[R_0 R_1] = 0$.*

Then, one takes $|\vec{0}\rangle = \otimes_{n=1}^{N_q} |s_n\rangle$, and $U = \prod_{n=1}^{N_q} U^{(n)}$, where

$$U^{(n)} = \prod_{j=0,1} \prod_{S \in \mathcal{S}^{(n)}} e^{i\theta^n_{S,j} R_j^{(n)} S}. \tag{5.3.1}$$

The definition above allows for any choice of the $[n-1, n-1]$ stabilizer groups $\mathcal{S}^{(n)}$, including ones with non-commuting elements between different $\mathcal{S}^{(n)}$. However, we use the following prototypical example throughout the rest of this chapter.

Example 8. *The **quantum combinatorial ansatz**, or **QCA**, is a stabilizer ansatz with $|s_i\rangle = |0\rangle$, $R_0^{(n)} = X$, $R_1^{(n)} = Y$, and $\mathcal{S}^{(n)} = \langle X_i, i = 1, \ldots, n-1\rangle$.*

A compressed circuit for the quantum combinatorial ansatz on 3 qubits is given in Fig. 5.2

Theorem 9. *A stabilizer ansatz $(U, |\vec{0}\rangle)$ spans the entire Hilbert space of N_q-qubit states with the minimal number of parameters.*

Proof — That the number of parameters is minimal may be immediately calculated,

$$N_p = \sum_{n=1}^{N_q} 2 \times 2^{n-1} = 2(2^{N_q} - 1). \tag{5.3.2}$$

We then prove that the ansatz spans the entire Hilbert space by induction. The stabilizer group $\mathcal{S}^{(n)}$ gives a basis $|p\rangle$ for the $n-1$ qubit Hilbert space. Then, as $[R_j^{(n)} S, R_j^{(n)} S'] = 0$, one may rewrite $U^{(n)}$ as

$$U^{(n)} = \prod_{j=0,1} \exp\left[i \sum_{S \in \mathcal{S}^{(n)}} \theta^n_{S,j} R_j^{(n)} S \right]. \tag{5.3.3}$$

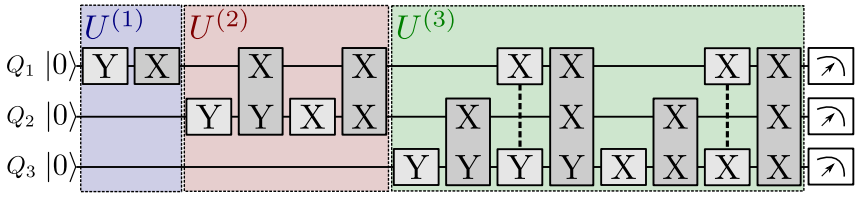

Figure 5.2: A circuit for the QCA on 3 qubits. For simplicity, we label each circuit element $U_i(\vec{\theta})$ by the tensor factors of its generating Pauli operator T_i (=: $R^{(n)}S$ in Eq. (5.3.1)) on each qubit. For example, the label XXX corresponds to the rotation $e^{i\theta^3_{XX,0}XXX}$. This compression may be expanded on as shown in Fig. 5.1 using the methods of [60]. For N_q qubits, QCA contains $2(2^{N_q}-1)$ gates and is proven to cover the entire Hilbert space (Theorem 9). In a practical application, QCA is to be reduced to polynomial size via a hierarchical approach outlined in Sec. 5.5. Note that the order of gate multiplication in QCA does not imply the order of gate importance in the hierarchical reduction scheme of Sec. 5.5. For instance, consider an application of the displayed QCA circuit to the open transverse-field Ising chain (Sec. 5.6). In this case, the two gates preferred in the reduction are those generated by Paulis XYI and IXY, followed by the one generated by XIY (cf. Fig. 5.5).

This sends the state $|p\rangle|s_n\rangle$ to the state

$$|p\rangle \left(e^{i\theta^n_{p,0}R^{(n)}_0} e^{i\theta^n_{p,1}R^{(n)}_1}\right)|s_n\rangle, \tag{5.3.4}$$

where the angles $\theta^n_{p,j}$ are given by the following linear transformation:

$$\theta^n_{p,j} = \sum_{S\in\mathcal{S}^{(n)}} S_p \theta^n_{S,j}, \qquad S_p = \langle p|S|p\rangle \in \{\pm 1\}. \tag{5.3.5}$$

This is the Hadamard-Walsh transformation, which is invertible, so $\theta^n_{p,j}$ can now be treated as independent parameters. On the other hand, our choice of $R^{(n)}_j$ explicitly takes the starting state $|s_n\rangle$ on qubit n to any state on the Bloch sphere. This implies that if we have the ability to create an arbitrary $n-1$-qubit state

$$|\Psi^{(n-1)}\rangle = \sum_p a_p|p\rangle, \tag{5.3.6}$$

$U^{(n)}|\Psi^{(n-1)}\rangle|s_n\rangle$ may be tuned to achieve any state of the form

$$\sum_p a_p|p\rangle \left(e^{i\theta^n_{p,0}R^{(n)}_0} e^{i\theta^n_{p,1}R^{(n)}_1}\right)|s_n\rangle, \tag{5.3.7}$$

which describes an arbitrary n-qubit state. This then completes the proof of coverage by induction, as $U^{(1)}|s_1\rangle$ covers the entire Bloch sphere. □

5.4 Children ansatzes and their construction

The cost of implementing a product VQE grows polynomially in both the number of units N_u (as this dictates the circuit size) and the number of parameters N_p (as this dictates the size of the optimization problem). Thus, an ansatz that covers the entire Hilbert space is too expensive to be of use; one must use it to construct child ansatzes of a manageable size.

Definition 10. *A product ansatz $(U', |\vec{0}'\rangle)$ is a **child ansatz** of a **parent** product ansatz $(U, |\vec{0}\rangle)$ when each unit U'_i of U' also appears in U.*

This definition is operational rather than fundamental; the variational manifold of a child ansatz is not necessarily a submanifold of the parent ansatz' variational manifold. However, one expects that these children ansatzes will still inherit some properties of the parent. In particular, we expect that a parent ansatz that spans as large a part of the Hilbert space as possible will lead to children ansatzes that are similarly large.

5.4.1 Ansatz compression and hierarchical construction

An obvious method to construct a child ansatz from a parent is to simply get rid of individual units or parameters:

Definition 11. *Given a product ansatz $(\prod_j U_j(\theta_{n_j}), |\vec{0}\rangle)$, one may **remove** a parameter θ_{n_i} to obtain the child ansatz $(\prod_{n_j \neq n_i} U_j(\theta_{n_j}), |\vec{0}\rangle)$, or **fix** a parameter $\theta_{n_i} = c\theta_{n_j}$ with $c \in \mathbb{R}$ to obtain the child ansatz $(\prod_l U'_l(\theta_{m_l})|\vec{0}\rangle)$, where $m_i = n_j$, $m_l = n_l$ for $l \neq i$, and $T'_l = cT_l$ whenever $n_l = n_i$.*

Parameter fixing may be considered strictly more general than unit removal, as fixing $\theta_{n_i} = 0\theta_{n_j}$ produces the same variational manifold as removing θ_{n_i}. However, unit removal reduces both N_u and N_p, while parameter fixing does not reduce the resulting circuit length.

Alternatively, one may construct child ansatzes using a bottom-up approach:

Definition 12. *Given a product ansatz $(\prod_j U_j(\theta_{n_j}), |\vec{0}\rangle)$, one may construct a **priority list** $(U_{j_1}, U_{j_2}, \ldots)$ of the possibly-repeated units of the*

ansatz. Such a priority list allows the construction of a ***hierarchy*** *of child ansatzes* $(U_M, |\vec{0}\rangle)$ *(for* $M > 0$*), where*

$$U_M(\vec{\theta}) = \prod_{m=1}^{M} U_{j_m}(\theta_{n_m}). \tag{5.4.1}$$

The two methods described above may be combined if desired. Subsequent generations of ansatzes will trade off a lower cost to implement against a smaller-sized variational manifold. We now focus on methods to optimize this balance. We first demonstrate how one may use unit reduction and parameter fixing to force a large VQE to respect symmetry constraints on the system. Following this, we take a rigorous perturbative approach to construct priority lists for a given target Hamiltonian.

5.4.2 Compression over symmetries

One may often restrict the ground state of a system by symmetries of the Hamiltonian; that is, operators S that commute with H. When this is true, all eigenstates $|E_0\rangle$ of H may be chosen to be eigenstates of S. This is particularly relevant in electronic systems where the particle number $\sum_i Z_i$ or parity $\prod_i Z_i$ is conserved. The symmetry is enforced on all states in a variational ansatz $(U, |\vec{0}\rangle)$ when $|\vec{0}\rangle$ is an eigenstate of S, and $[U(\vec{\theta}), S] = 0$ for all choices of the parameters θ. This in turn requires for an ordered product ansatz $U(\vec{\theta}) = \prod_i U_i(\theta_{n_i})$ that $[\prod_{i,n_i=n} U_i(\theta_{n_i}), S] = 0$ for all unique parameters n and for all choices of θ_{n_i}. If a parameter θ_{n_i} is associated to a single generator T_i, then this occurs if and only if $[T_i, S] = 0$.

When a symmetry is not respected by a variational ansatz, one may choose to either remove or fix the offending terms (see [171] for an alternative approach). Removal of generators that do not respect a given symmetry is simplest, but may be too restrictive for our desires. One may fix an ordered product ansatz to obey a symmetry that is broken by a set of commuting generators $\{T_{M_0}, T_{M_0+1}, \ldots, T_{M_1}\}$. To do this, one needs to solve the system of linear equations

$$\sum_{m=M_0}^{M_1} c_m \sum_{i,n_i=m} [S, T_i] = 0, \tag{5.4.2}$$

and fix $c_n\theta_n = c_m\theta_n$ for $N \leq n, m \leq M$. This requires fixing *all* parameters between N and M, which in turn might require rearranging the original ansatz to place specific units next to each other.

A very simple symmetry to enforce in a problem is the (antiunitary) complex conjugation operator, $\mathcal{K}i = -i\mathcal{K}$. (This symmetry is respected whenever the Hamiltonian is purely real.) As we have defined our generators T_i with an imaginary unit, $U_i = e^{i\theta_{n_i} T_i}$ commutes with $\mathcal{K}$ when T_i anti-commutes with $\mathcal{K}$. (e.g. for a single qubit, the rotation $e^{i\theta Y}$ rotates between the real eigenstates of the real X and Z Pauli operators.)

Example 13. *The YYX toy ansatz is the compression of the QCA stabilizer ansatz for two qubits over $\mathcal{K}$. It thus spans the entire Hilbert space of 2-qubit states with real coefficients (which matches the calculation of its variational area).*

In App. 5.B, we give another example of a symmetry-compressed Pauli-type ansatz - the fermionic unitary coupled cluster ansatz.

5.4.3 Size-extensivity of a variational ansatz

To show beyond-classical performance, we desire our variational quantum algorithms to be able to produce strongly entangled states, inaccessible to a classical computer. For this, we would like the VQE ansatz to represent quantum correlations in a maximally compact manner. To achieve this, we are guided by the idea of size-extensivity. The notion of size-extensivity has its origins in strongly-correlated physics, and is formalized there by the linked cluster theorem [57]. The rough notion is that: (1) if a computation treats two uncoupled systems together, it should converge to the same solution as when it treats them independently, and (2) the only complexity one should be adding to the solution of coupled systems is that which is minimally demanded. Formalizing this idea requires somewhat heavy machinery; we give a formal definition later in the text (Def. 23) and now put forward the following (weaker) statement as an informal definition.

Definition 14. *(informal) Consider variational ansatz $(U, |\vec{0}\rangle)$ for a Hamiltonian H on a system S, and an arbitrary (disjoint) partition $S = \cup_i S_i$ with a decomposition $H = \sum_i H_i + H_{\text{other}}$ where each H_i acts only on S_i (and each term in H_{other} acts on multiple S_i). In this case, the ansatz $(U, |\vec{0}\rangle)$ is size-extensive if for any such partition, the unitary $U(\vec{\theta})$ that minimizes the variational energy $E(\vec{\theta})$ (Def. 1) reduces to the form*

$$U(\vec{\theta}) = \prod_i U(\vec{\theta}_i) \tag{5.4.3}$$

if H_{other} is reduced to 0. In (5.4.3), each $U(\vec{\theta}_i)$ acts only on system S_i

(i.e. the coefficients of any part of the ansatz U that acts outside of S_i are set to 0).

In the language of Def. 14, the stronger statement of Def. 23 is needed to treat the case where $\{S_i\}$ together form a connected system, but some pairs (S_k, S_l) are mutually separated (e.g. because of spatial locality). It appears that in this case, a variational ansatz is efficient if it tends to introduce more correlations between less separated subsystem pairs (S_k, S_l). However, this heuristic needs to be re-stated more rigorously. In Def. 23, we provide such a rigorous formulation and apply it in an explicit construction of size-extensive ansatzes.

5.5 Perturbative construction for digital size-extensive ansatzes

We now propose a perturbative approach for the construction of digital size-extensive ansatzes. We formulate it in terms of a gate hierarchy list $(U_1, \ldots)$ derived from a large parent ansatz $(U, |\vec{0}\rangle)$. To decide on the hierarchy list, we split the system Hamiltonian H into the non-interacting part H_0 and the coupling JV ($\|H_0\|, \|V\| \sim 1$):

$$H = H_0 + JV \tag{5.5.1}$$

To allow for analytical treatment we consider the weak coupling limit, $J \ll 1$. In this limit, the overlap between the true ground state $|E_0\rangle$ and unperturbed excited states $|E_j^0\rangle$ is exponentially small in the number of applications of V required to couple $|E_j^0\rangle$ to the unperturbed ground state $|E_0^0\rangle$. We may rewrite the non-interacting part H_0 via a unitary transformation as

$$H_0 = \sum_{n=1}^{N_q} h_n Z_n, \tag{5.5.2}$$

which ties each $|E_j^0\rangle$ to a computational basis state $|\vec{s}\rangle$

$$H_0|\vec{s}\rangle = -\sum_{n=1}^{N_q} (-1)^{s_n} h_n |\vec{s}\rangle. \tag{5.5.3}$$

If we can further tie each state $|\vec{s}\rangle$ to one or a few variational units $U_i(\theta_i)$, we can construct a hierarchy list of these $U_i(\theta_i)$ based on the approximate magnitude of $|\langle\vec{s}|E_0\rangle|$. The resulting hierarchy list is to be used in the

VQE procedure for the original, potentially strongly coupled Hamiltonian H ($J = O(1)$).

Performing this construction in a size-extensive way runs into a challenge which we call 'back-action'. Namely, the action of any unit $U_i(\theta_i)$ on the state $\prod_{j<i} U_j(\theta_j)|\vec{0}\rangle$ may be very different to the action of $U_i(\theta_i)$ on the starting state. In particular, one could imagine this action generating an undesired term to the variational wavefunction which must be cancelled by later rotations. As we will show, one can deal with this back-action while retaining the size-extensivity. To achieve this, in the rest of this section we will expand the target equality,

$$|E_0\rangle \simeq |\psi(\vec{\theta})\rangle, \tag{5.5.4}$$

assuming that $|\psi(\vec{\theta})\rangle$ is given by a digital (i.e., Pauli-type) ansatz. We will do so in terms of a Pauli decomposition of the perturbation

$$JV = \sum_{i=1}^{N_c} J_i V_i, \qquad V_i \in \mathbb{P}^{N_q}, \tag{5.5.5}$$

and then we will equate terms based on the order of their polynomial dependence on each J_i. On the left-hand side (Sec. 5.5.1), we will use a Dyson expansion, with a convenient diagrammatic representation. On the right-hand side (Sec. 5.5.2) we will use a Taylor expansion of the exponential operators. We will show that a single condition (Def. 21) on the parent ansatz is sufficient to automatically cancel all undesired back-action. Then, we will show that an additional condition (Def. 24) causes the back-action terms to precisely cancel out any need for entangling circuits between disconnected regions (Theorem 26). This ensures the desired feature of size-extensivity, thus providing the digital quantum version of the linked-cluster theorem [57]. The QCA ansatz of Example 8 will be seen to satisfy the above conditions, and therefore gives rise to a hierarchy of size-extensive digital ansatzes.

Our perturbative approach can be thought of as a digital unitary relative of the Kirkwood-Thomas expansion [172, 173]. Also note, that as we intend to optimize the parameters $\vec{\theta}$ as part of the VQE, we will approximate these only to leading order in the interaction strength J. This makes our method potentially applicable even in the strongly correlated regime where perturbation theory breaks down.

5.5.1 Diagrammatic expansion of the ground state

To expand the left-hand side of Eq. (5.5.4), let us use vector notation $\vec{J}$ for the coupling terms J_i (and $\vec{V}$ for the operators V_i). Then, let us introduce some notation that simplifies the following expressions:

$$\vec{a}^{\cdot\vec{k}} := \prod_i a_i^{k_i} = \exp(\vec{k} \cdot \log(\vec{a})). \tag{5.5.6}$$

We wish to use this expression for both vectors of numbers (e.g. $\vec{J}$) and vectors of operators (e.g. $\vec{V}$). In the latter we must take care of ordering; as previous, we assume that the product runs right-to-left. As Pauli operators either commute or anticommute, rearranging these products simply requires one to keep track of minus signs. This may be assisted by the following definition

Definition 15. *A vector $\vec{V}$ of N_c Pauli operators defines a phase $\Gamma(\vec{k}) \in \{0,1,2,3\}$ and a state $\vec{s}(\vec{k})$ on a vector $\vec{k} \in \mathbb{N}^{N_c}$ * by*

$$\vec{V}^{\cdot\vec{k}}|\vec{0}\rangle = i^{\Gamma(\vec{k})}|\vec{s}(\vec{k})\rangle, \tag{5.5.7}$$

and a relative sign $S_{\vec{k},\vec{k}'} \in \{-1,1\}$ for $\vec{k},\vec{k}' \in \mathbb{N}^{N_c}$ by

$$\vec{V}^{\cdot\vec{k}}\vec{V}^{\cdot\vec{k}'} = S_{\vec{k},\vec{k}'}\vec{V}^{\cdot(\vec{k}+\vec{k}')}. \tag{5.5.8}$$

Then, as Pauli operators map computational basis states to computational basis states, $\vec{V}^{\vec{k}}|\vec{0}\rangle$ is an eigenstate of H_0, with energy

$$E_{\vec{s}(\vec{k})} := -\sum_{n=1}^{N_q}(-1)^{\vec{s}(k)_n}h_n. \tag{5.5.9}$$

Let us now expand the ground state as a Taylor series in $\vec{J}$:

$$|E_0\rangle = \sum_{\vec{k}\in\mathbb{N}^{N_c}} \vec{J}^{\vec{k}}|\Psi_{\vec{k}}\rangle. \tag{5.5.10}$$

Following a standard Dyson expansion (details in App. 5.C), we observe that

*We take the natural numbers $\mathbb{N}$ to include 0.

Lemma 16. *The vectors* $|\Psi_{\vec{k}}\rangle$ *take the form*

$$|\Psi_{\vec{k}}\rangle = C_{\vec{k}} \vec{V}^{\cdot \vec{k}} |\vec{0}\rangle, \tag{5.5.11}$$

where $C_{\vec{k}}$ *is a real number.*

To find the values of coefficients $C_{\vec{k}}$, we first develop a perturbative expansion for a ground state $|\tilde{E}_0\rangle$ with a special normalization condition $\langle\vec{0}|\tilde{E}_0\rangle = 1$,

$$|\tilde{E}_0\rangle = \sum_{\vec{k}\in\mathbb{N}^{N_c}} \vec{J}^{\cdot\vec{k}} |\tilde{\Psi}_{\vec{k}}\rangle. \tag{5.5.12}$$

The states $|\tilde{\Psi}_{\vec{k}}\rangle$ then satisfy (see App. 5.C):

$$|\tilde{\Psi}_{\vec{k}}\rangle = \tilde{C}_{\vec{k}} \vec{V}^{\cdot\vec{k}} |\vec{0}\rangle, \tag{5.5.13}$$

where $\tilde{C}_{\vec{k}}$ is a real number. In particular, if $\vec{\delta}_\beta$ is the unit vector with a 1 in the β index, $\tilde{C}_{\vec{k}} = \delta_{\vec{k},\vec{0}}$ if $\vec{s}(\vec{k}) = \vec{0}$, and is otherwise given by the recursive relation

$$\begin{aligned}\tilde{C}_{\vec{k}} =& (E^{(0)}_{\vec{0}} - E^{(0)}_{\vec{s}(\vec{k})})^{-1} \sum_{\beta, k_\beta>0} \Big\{ \tilde{C}_{\vec{k}-\vec{\delta}_\beta} S_{\vec{\delta}_\beta, \vec{k}-\vec{\delta}_\beta} \\ &- \sum_{\substack{\vec{k}'<\vec{k},\ k'_\beta>0 \\ \vec{s}(\vec{k}')=0}} \tilde{C}_{\vec{k}'-\vec{\delta}_\beta} \tilde{C}_{\vec{k}-\vec{k}'} S_{\vec{\delta}_\beta, \vec{k}'-\vec{\delta}_\beta} S_{\vec{k}-\vec{k}', \vec{k}'} \Big\},\end{aligned} \tag{5.5.14}$$

where $\vec{k}' < \vec{k}$ if $k'_\beta \leq k_\beta$ for all β and $\vec{k}' \neq \vec{k}$. To find the coefficients $C_{\vec{k}}$ of the normalized ground state, one may then expand the expression $|E_0\rangle = \langle\tilde{E}_0|\tilde{E}_0\rangle^{-1/2} |\tilde{E}_0\rangle$ in powers of $\vec{J}$, which allows to express $C_{\vec{k}}$ in terms of $\tilde{C}_{\vec{k}}$ obtained from (5.5.14).

We note here that we have no guarantee that the normalization constant $\mathcal{N} = \langle\tilde{E}_0|\tilde{E}_0\rangle^{-1/2}$ behaves regularly in thermodynamic limit $N_q \to \infty$. This is a standard breakdown of perturbation theory for the wavefunction, however when this occurs our approach to VQE construction is still possible, and may indeed still be practical. At the stage of estimating the variational parameters $\vec{\theta}$, we will be using the $\tilde{C}_{\vec{k}}$ coefficients, since they behave regularly and are more practical to calculate. As $\vec{\theta}$ will be optimized later on the quantum device, the estimation itself need not be exact.

The size-extensivity of our approach relies on an important relationship between $C_{\vec{k}}$ terms that are the combination of disconnected pieces. To formalize this notion of connectedness, we introduce some terminology:

Definition 17. *For a perturbative contribution $C_{\vec{k}}$, the set of couplings V_i s.t. $k_i \neq 0$, is said to be* ***activated*** *in $\vec{k}$. The set of qubits on which at least one activated coupling V_i acts non-trivially is called the* ***support*** *of $\vec{k}$.*

Then the connectedness of the contribution $C_{\vec{k}}$ is defined as follows:

Definition 18. *A perturbative contribution $C_{\vec{k}}$ is disconnected if one may write*

$$\vec{k} = \vec{k}_A + \vec{k}_B, \tag{5.5.15}$$

such that the respective supports of $\vec{k}_A$ and $\vec{k}_B$ do not share any qubits. This implies, but is not equivalent to, the following statement:

$$\vec{V}^{\cdot\vec{k}} = \vec{V}^{\cdot\vec{k}_A}\vec{V}^{\cdot\vec{k}_B} \tag{5.5.16}$$

The disconnected contributions $C_{\vec{k}}$ obey the following special property (proven in App. 5.D).

Lemma 19. *If a perturbative contribution $C_{\vec{k}}$ is disconnected w.r.t. a splitting (5.5.16) into $\vec{k}_A$ and $\vec{k}_B$,*

$$C_{\vec{k}} = C_{\vec{k}_A} C_{\vec{k}_B}. \tag{5.5.17}$$

This idea of connectedness of contributions may be described in a graphical representation of the product of operators $\vec{V}^{\cdot\vec{k}}$:

Definition 20. *Let $\vec{V}$ define the order of a decomposition of the perturbation $\vec{J}\cdot\vec{V}$ to a non-interacting Hamiltonian H_0. A perturbative diagram for a vector $\vec{k}$, is a bipartite graph with one circular vertex for each qubit, and k_β square vertices for each interaction V_β. We draw edges between each square vertex and the qubits that the corresponding V_β term acts non-trivially on, and color the edge to qubit i blue, red or black if $[V_\beta]_i = X, Y$ or Z respectively. Each circular vertex is then coloured black or white if it is connected to by an odd or even number of coloured edges respectively.*

A contribution $C_{\vec{k}}$ is connected if all square vertices in the perturbative diagram are connected *. In Fig. 5.3, we show some examples of connected and disconnected perturbative diagrams. Diagrams also allow one to read off $\vec{s}(\vec{k})$ ($s_i(\vec{k}) = 0$ when the corresponding vertex is white), and $\Gamma(\vec{k})$ mod 2 (being the number of red lines modulo 2). (The rest of $\Gamma(\vec{k})$ depends on the order in which the operations V_i are applied, which is not captured in the perturbative diagrams.)

*The circular vertices, corresponding to qubits, need not be connected, as a connected contribution need not act on all qubits.

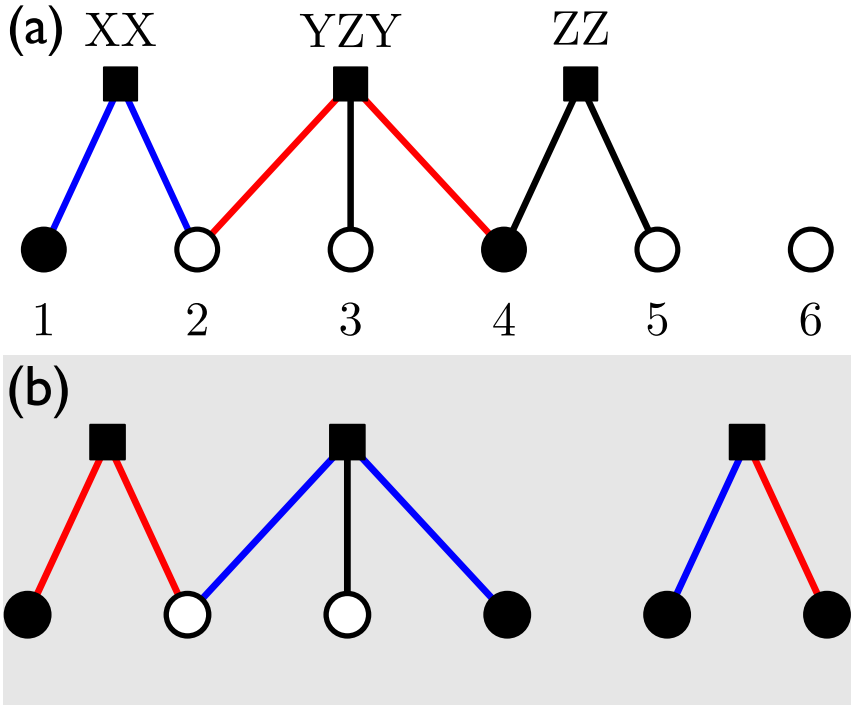

Figure 5.3: Example perturbative diagrams. (a) A connected diagram for a real contribution (even number of Y terms) to $|\vec{s}(\vec{k})\rangle = |100100\rangle$. Labels for qubits i and terms V_β are added for reference. (b) A disconnected diagram for an imaginary contribution to $|\vec{s}(\vec{k})\rangle = |100111\rangle$. Unnecessary labels here are excluded.

5.5.2 Taylor expansion of the variational ansatz

We now consider the expansion of the right hand side of Eq. (5.5.4). In keeping with the previous subsection, we wish to do this in terms of the individual perturbations J_i. Let us expand each coefficient θ_i in a power series over all interaction terms J_i

$$\theta_i = \sum_{\vec{k}} \theta_i^{(\vec{k})} \vec{J}^{\vec{k}} \left(= \sum_{\vec{k}} \theta_i^{(\vec{k})} J_1^{k_1} J_2^{k_2} \dots \right), \tag{5.5.18}$$

where the shorthand vector power notation was defined in Eq. (5.5.6). This may be substituted into the variational ansatz $(U, |\vec{0}\rangle)$

$$U(\vec{\theta}) = \prod_i \left\{ \prod_{\vec{k}} \exp\left[i\theta_i^{(\vec{k})} \vec{J}^{\vec{k}} T_i \right] \right\}, \tag{5.5.19}$$

where we added the brackets to emphasize the ordering of the product over i. Now, we take the Taylor series of the exponentials in Eq. (5.5.19), obtaining

$$U(\vec{\theta}) = \prod_i \left\{ \prod_{\vec{k}} \sum_{f=0}^{\infty} \frac{1}{f!} \left[i\theta_i^{(\vec{k})} \vec{J}^{\vec{k}} T_i \right]^f \right\}. \tag{5.5.20}$$

We will eventually wish to rearrange this product to identify all terms that share the same power of each J_i — that is, those that share the same $\vec{J}^{\vec{k}}$. This requires first expanding our product over sums to a sum over products (pulling the sum over integers g in front of the products over k and i). Each term in the resulting sum will have a unique product of powers of the different $\theta_i^{(k)}$. We can then associate this term to a function $\vec{f} : \mathbb{N}^{N_c} \to \mathbb{N}^{N_p}$; i.e. the power of $\theta_i^{(k)}$ in our term is given by $f_i(\vec{k})$. (Each such function $\vec{f}$ will correspond in Sec. 5.5.3 to a unique way to map the activations of couplings V_α from the left-hand side of Eq. (5.5.4) onto the generators T_i.) One may confirm that every unique function $\vec{f}$ corresponds to a single term in Eq. (5.5.20), and the powers of the T_i, $\vec{J}$, and the coefficient of each term may be expressed in terms of this function, allowing us to expand our unitary $U(\vec{\theta})$ as

$$U(\vec{\theta}) = \sum_{\vec{f}} \vec{J}^{\sum_{i,\vec{k}} f_i(\vec{k})\vec{k}} \left(i\vec{T} \right)^{\cdot \sum_{\vec{k}} \vec{f}(\vec{k})} \prod_{\vec{k},i} \frac{\left[\theta_i^{(\vec{k})} \right]^{f_i(\vec{k})}}{f_i(\vec{k})!}, \tag{5.5.21}$$

To put (5.5.21) in a simpler form, we define:

$$\vec{K}(\vec{f}) = \sum_{i,\vec{k}'} f_i(\vec{k}')\vec{k}' \tag{5.5.22}$$

$$\vec{N}(\vec{f}) = \sum_{\vec{k}} \vec{f}(\vec{k}), \tag{5.5.23}$$

$$\Theta(\vec{f}) = \prod_{\vec{k},i} \frac{\left[\theta_i^{(\vec{k})}\right]^{f_i(\vec{k})}}{f_i(\vec{k})!}, \tag{5.5.24}$$

which allows us to rewrite the sum as

$$U(\vec{\theta}) = \sum_{f:\mathbb{N}^{N_c}\to\mathbb{N}^{N_p}} \vec{J}^{\vec{K}(\vec{f})} \left(i\vec{T}\right)^{\cdot\vec{N}(\vec{f})} \Theta(\vec{f}). \tag{5.5.25}$$

One can give an interpretation for $\vec{K}(\vec{f})$, $\vec{N}(\vec{f})$ and $\Theta(\vec{f})$ in the expression. The vector $\vec{K}(\vec{f}) \in \mathbb{N}^{N_c}$ represents the PT order of a given term of the sum. (Note that multiple functions $\vec{f}$ will have the same PT order $\vec{K}(\vec{f})$.) $\vec{N}(\vec{f}) \in \mathbb{N}^{N_p}$ gives the activation of generators T_i in that term, and therefore tells us the computational basis state that this term produces as an operator acting on $|\vec{0}\rangle$. (Terms with $|\vec{N}(\vec{f})| \geq 2$ describe the 'back-action' of the ansatz which we will discuss in the next section.)

The information about the parameters $\vec{\theta}$ of the ansatz is now contained in the scalar coefficient $\Theta(\vec{f})$. Its values are not independent variables: $\Theta(\vec{f})$ can be fixed entirely by its action on the functions f s.t. $|\vec{N}(\vec{f})| = 1$. To see this, let us label such functions $\vec{f} = d^{\vec{k},i}$, where $d_j^{\vec{k},i}(\vec{k}') = \delta_{\vec{k},\vec{k}'}\delta_{i,j}$. These functions yield an activation of a single generator T_i from a single activation pattern $\vec{k}$ of couplings V_α. For such functions, we obtain $\Theta(d^{\vec{k},i}) = \theta_i^{(\vec{k})}$ – whose values indeed entirely determine the ansatz state. In particular, in the terms describing back-action (f s.t. $|\vec{N}(\vec{f})| \geq 2$), $\Theta(\vec{f})$ are nonlinear monomials of $\theta_i^{(\vec{k})}$, and thus are fixed by the values of $\Theta(d^{\vec{k},i})$.

5.5.3 Equating ansatz and perturbative terms

Our plan is now to solve for $\theta_i^{(\vec{k})}$, by comparing $|\psi(\vec{\theta})\rangle$ from Eq. (5.5.25) to the perturbative series for $|\Psi(\vec{J})\rangle$ from Eq. (5.5.10). We will equate the contributions coming from different PT orders, and those proportional to the same computational basis state. (The vectors $\vec{K}(\vec{f})$ and $\vec{N}(\vec{f})$ allow

us to identify which terms need be equated.) This will result in equations that are linear in the coefficients $C_{\vec{k}}$ and $\Theta(\vec{f})$. Due to the structure of $\Theta(\vec{f})$ these equations will be highly nonlinear in $\theta_i^{(\vec{k})}$. However, under certain conditions (Def. 21 and Def. 24), we find that these equations for $\theta_i^{(\vec{k})}$ may be solved iteratively, and that many coefficients will vanish exactly. This will yield a class of ansatzes which are also size-extensive, the technical definition of which we give in Def. 23. For such ansatzes, we will have a guarantee that a relatively compact circuit is capable of reproducing the perturbative series for $|\Psi(\vec{J})\rangle$ up to a given PT order $\vec{k}$. These circuits will have a relatively small (polynomial in N_q at fixed PT order $\vec{k}$) number of free parameters when used as a VQE, as this coincides with the number of leading order connected diagrams up to order $\vec{k}$.

Equating the action of the Taylor-expanded $U(\vec{\theta})$ (Eq. (5.5.25)) on the starting state $|\vec{0}\rangle$ to the expansion of the ground state $|E_0\rangle$ (Eq. (5.5.10)) and separating in orders of $\vec{J}$ obtains the form

$$C_{\vec{k}}\vec{V}^{\cdot\vec{k}}|\vec{0}\rangle - \sum_{f;\vec{K}(\vec{f})=\vec{k}} \Theta(\vec{f}) \left(i\vec{T}\right)^{\cdot\vec{N}(\vec{f})} |\vec{0}\rangle = 0. \tag{5.5.26}$$

This may be further separated by taking the inner product with different computational basis states to give the equations

$$C_{\vec{k}} - \sum_{f;\ \vec{K}(\vec{f})=\vec{k}} \Theta(\vec{f})\langle\vec{0}|\vec{V}^{\cdot\vec{k}\dagger} \left(i\vec{T}\right)^{\cdot\vec{N}(\vec{f})} |0\rangle = 0 \tag{5.5.27}$$

$$\sum_{f;\ \vec{K}(\vec{f})=\vec{k}} \Theta(\vec{f})\langle\vec{s}\neq 0|\vec{V}^{\cdot\vec{k}\dagger} \left(i\vec{T}\right)^{\cdot\vec{N}(\vec{f})} |0\rangle = 0. \tag{5.5.28}$$

Eqs. 5.5.28 contain what we call the back-action terms. These are undesirable; if one fixes the $\theta_i^{(\vec{k})}$ values one at a time, then any non-zero term appearing in Eqs. 5.5.28 will need to be cancelled out by fixing some other $\theta^J_{\vec{k}'}$ at a later point. However, these terms may be avoided for a large class of parent ansatzes:

Definition 21. *A Pauli-type ansatz* $(\prod_i e^{iT_i\theta_i}, |\vec{0}\rangle)$ *is* ***generating*** *if, for all computational basis states* $|\vec{s}\rangle \neq |\vec{0}\rangle$*, there exist generators* $T_{\vec{s},a}$ *for* $a = 0, 1$ *such that* $iT_{\vec{s},a}|\vec{0}\rangle = i^a|\vec{s}\rangle$*.*

Note that a generating ansatz requires at least sufficient parameters to span the entire Hilbert space, however it remains unclear whether a

generating ansatz does span the entire Hilbert space. Instead, we are interested in generating ansatzes here as they avoid undesired back-action

Lemma 22. *Given a generating Pauli-type ansatz* $(\prod_{\vec{s},a} e^{iT_{\vec{s},a}\theta_{\vec{s},a}}, |\vec{0}\rangle)$, *one may solve Eqs. 5.5.27 by fixing* $\theta^{(\vec{k})}_{\vec{s},a} = 0$ *unless* $\vec{s} = \vec{s}(\vec{k})$ *and* $a = a(\vec{k}) := \Gamma(\vec{k}) \mod 2$. *This solution further prevents undesired back-action by making Eqs. 5.5.28 zero term-wise.*

Proof — Eq. (5.5.27) may be rewritten as

$$\sum_a i^{a-\Gamma(\vec{k})}\theta^{(\vec{k})}_{\vec{s}(\vec{k}),a} = C_{\vec{k}} - \sum_{\substack{f;\vec{K}(\vec{f})=\vec{k}\\ |\vec{N}(\vec{f})|>1}} \Theta(\vec{f})\langle 0|\vec{V}^{\cdot\vec{k}\dagger}\left(i\vec{T}\right)^{\cdot\vec{N}(\vec{f})}|0\rangle. \tag{5.5.29}$$

We then use this equation to fix the left-hand side, being an equation of free $\Theta(\vec{f})$ terms. If this is done in ascending order in $|\vec{k}|$, one can check that all $\Theta(\vec{f})$ terms on the right-hand side at each $\vec{k}$ will have been fixed previously, implying that this fixing is well-defined. Then, one notes that

$$\langle 0|\vec{V}^{\cdot m\vec{k}\dagger}\left(i\vec{T}\right)^{\cdot\vec{N}(md^{\vec{k},i})}|0\rangle = \langle 0|\vec{V}^{\cdot\vec{k}\dagger}\left(i\vec{T}\right)^{\cdot\vec{N}(d^{\vec{k},i})}|0\rangle, \tag{5.5.30}$$

for any odd m, which implies that contributions from linear combinations of the fixed components will never appear in Eq. (5.5.28).

The above implies that the (strictly real) term $C_{\vec{k}}$ from each perturbative diagram contributes only to $\theta^{(k)}_{\vec{s}(\vec{k}),a(\vec{k})}$. Then, by definition, we have

$$iT_{\vec{s}(\vec{k}),a(\vec{k})}|\vec{0}\rangle = \pm\vec{V}^{\cdot\vec{k}}|\vec{0}\rangle, \tag{5.5.31}$$

and as Pauli operators are either entirely real or entirely imaginary, this extends to any computational basis state $|\vec{s}'\rangle$

$$iT_{\vec{s}(\vec{k}),a(\vec{k})}|\vec{s}'\rangle = \pm\vec{V}^{\cdot\vec{k}}|\vec{s}'\rangle. \tag{5.5.32}$$

This implies that for any function $\vec{f}$ such that $f_{s,a}(\vec{k}) = 0$, unless $\vec{s} = \vec{s}(\vec{k}), a = a(\vec{k})$ we have

$$\begin{aligned}\left(i\vec{T}\right)^{\cdot\vec{N}(\vec{f})}|\vec{0}\rangle &= \pm\prod_{\vec{k}}\left(iT_{\vec{s}(\vec{k}),a(\vec{k})}\right)^{f_{\vec{s}(\vec{k}),a(\vec{k})}(\vec{k})}|\vec{0}\rangle \\ &= \pm\prod_{\vec{k}}\vec{V}^{\cdot f_{\vec{s}(\vec{k}),a(\vec{k})}(\vec{k})\vec{k}}|\vec{0}\rangle = \pm\vec{V}^{\cdot\vec{K}(\vec{f})}|\vec{0}\rangle,\end{aligned} \tag{5.5.33}$$

and so the right-hand side of Eq. (5.5.29) is real, and $\theta^{(\vec{k})}_{\vec{s}(\vec{k}),1-a(\vec{k})} = 0$, by induction in $|\vec{k}|$. □

For carefully-chosen Pauli-type ansatzes, one may further cancel contributions from disconnected diagrams. This yields our formal definition of what it means for such an ansatz to be 'size-extensive' (as discussed in Sec. 5.4.3)

Definition 23. *We say that a Pauli-type ansatz $U(\vec{\theta})$ is size-extensive with respect to a perturbation JV (Eq. (5.5.1)) if, in a solution to Eqs. 5.5.27, $\theta^{(\vec{k})}_{\vec{s}(\vec{k}),a} = 0$ if $\vec{k} = \vec{k}_A + \vec{k}_B$ is disconnected (Def. 18).*

A Pauli-type ansatz satisfying this definition will satisfy Def. 14 whenever the perturbative expansion above converges. To see this, note that when the perturbative expansion converges the solution to Eqs. 5.5.27 will provide the ground state exactly. Then, consider a Hamiltonian that does not couple two systems S_i and S_j, and a term $T_{\vec{s},a}$ in our ansatz that does couple S_i and S_j. One can see that whenever $\vec{s} = \vec{s}(\vec{k}), a = a(\vec{k})$ for some $\vec{k}$ that $\vec{k}$ will be disconnected, and so $\theta_{\vec{s},a} = 0$ at all orders of k by Def. 23.

We now have the machinery to present a condition for our ansatz to be size-extensive that just relates the ansatz terms T_i to the perturbation terms V_i.

Definition 24. *A generating Pauli-type ansatz is **matched** to a perturbation JV if*

$$\langle 0|\vec{V}^{\cdot\vec{k}\dagger}\left(i\vec{T}\right)^{\cdot\vec{N}(\vec{f})}|0\rangle\langle 0|\vec{V}^{\cdot\vec{k}'\dagger}\left(i\vec{T}\right)^{\cdot\vec{N}(\vec{f}')}|0\rangle = \langle 0|\vec{V}^{\cdot(\vec{k}+\vec{k}')\dagger}\left(i\vec{T}\right)^{\cdot\vec{N}(\vec{f}+\vec{f}')}|0\rangle, \tag{5.5.34}$$

whenever $(\vec{k}, f)$ and $(\vec{k}', \vec{f}')$ act non-trivially on disconnected parts of the system.

Example 25. *Any generating variational ansatz $(\prod_{\vec{s},a} e^{T_{\vec{s},a}\theta_{\vec{s},a}}, |\vec{0}\rangle)$ for which the generators $T_{\vec{s},a}$ are compact (i.e. they only act nontrivially on qubit j if $s_j = 1$), is matched. In particular, QCA (Example 8) is both generating and matched.*

Theorem 26. *A perturbative hierarchy constructed from a Pauli-type ansatz via Eqs. 5.5.27, that is matched to a perturbation JV, is size-extensive.*

Proof — By Lemma 19, we have that $C_{\vec{k}} = C_{\vec{k}_A} C_{\vec{k}_B}$. Inserting Eq. (5.5.27), we find

$$C_{\vec{k}} = \sum_{f_A, \vec{K}(\vec{f}_A)=\vec{k}_A} \sum_{f_B, \vec{K}(\vec{f}_B)=\vec{k}_B} \Theta(\vec{f}_A)\Theta(\vec{f}_B) \times \langle\vec{0}|\vec{V}^{\cdot \vec{k}_A \dagger} \left(i\vec{T}\right)^{\cdot \vec{N}(\vec{f}_A)} |\vec{0}\rangle\langle\vec{0}|\vec{V}^{\cdot \vec{k}_B \dagger} \left(i\vec{T}\right)^{\cdot \vec{N}(\vec{f}_B)} |\vec{0}\rangle. \tag{5.5.35}$$

As disconnected parts of $\vec{k}$, either $k_{A,i} = 0$ or $k_{B,i} = 0$ for any i, implying $f_A(\vec{k}') = \vec{0}$ or $f_B(\vec{k}') = \vec{0}$ for all $\vec{k}'$ in the above sum. From this we may write

$$\begin{aligned}\Theta(\vec{f}_A + \vec{f}_B) &= \prod_{\vec{k}', i} \frac{\left[\theta_i^{(\vec{k})}\right]^{f_A(\vec{k}')+\vec{f}_B(\vec{k}')}}{(\vec{f}_A(\vec{k}') + \vec{f}_B(\vec{k}'))!} \\ &= \Theta(\vec{f}_A)\Theta(\vec{f}_B).\end{aligned} \tag{5.5.36}$$

Combining this with the definition of a matched ansatz obtains

$$C_{\vec{k}} = \sum_{\substack{f_A, \vec{K}(\vec{f}_A)=\vec{k}_A \\ f_B, \vec{K}(\vec{f}_B)=\vec{k}_B}} \Theta(\vec{f}_A + \vec{f}_B)\langle\vec{0}|\vec{V}^{\vec{k}\dagger} \left(i\vec{T}\right)^{\vec{N}(\vec{f}_A+\vec{f}_B)} |\vec{0}\rangle. \tag{5.5.37}$$

It remains to check that all $f : \mathbb{N}^{N_c} \to \mathbb{N}^{N_p}$ with $\vec{K}(\vec{f}) = \vec{k}$, $|\vec{N}(\vec{f})| > 1$, and $\Theta(\vec{f}) \neq 0$ take the form $\vec{f} = \vec{f}_A + \vec{f}_B$ with $\vec{K}(\vec{f}_A) = \vec{k}_A$ and $\vec{K}(\vec{f}_B) = \vec{k}_B$, in which case the right-hand side of Eq. (5.5.29) cancels, giving the required result. This may be seen by induction in $|\vec{K}(\vec{f})|$. Clearly it is true for $|\vec{K}(\vec{f})| = 1$. Then, fix f with $|\vec{K}(\vec{f})| > 1$, and define $f_A(\vec{k}') = f(\vec{k}')$ if $\vec{k}'_i \vec{k}_{B,i} = 0$ for all i and $f_A(\vec{k}') = 0$ otherwise, and similarly for $f_B(\vec{k}')$, and define $f_{AB} = f - f_A - f_B$. One has that $\Theta(\vec{f}) = \Theta(\vec{f}_A)\Theta(\vec{f}_B)\Theta(\vec{f}_{AB})$, but if $f_{AB} \neq 0$, it is a product of $\theta^{(\vec{k}_{AB})}_{\vec{s}(\vec{k}_{AB}),a}$ for disconnected $\vec{k}_{AB}$ with $|\vec{k}_{AB}| < K$, and thus $\Theta(\vec{f}_{AB}) = 0$. □

This result can be seen as the digital quantum cousin of the linked-cluster theorem [57].

5.5.4 The perturbative construction

Following the above, we can construct a hierarchy of the $T_{\vec{s},a}$ by estimating the corresponding value of $\theta_{\vec{s},a}$ and placing them in order. We do not

need to know the precise values of $\theta_{\vec{s},a}$, as these will be optimized as part of the VQE. Instead we plan to estimate only the largest contributions to each $\theta_{\vec{s},a}$. Under the assumption that $J_i \; J \ll h_n$ for all interaction terms i and all qubits n, we expect the largest contributions to come from those (connected) $C_{\vec{k}}$ with smallest possible $|\vec{k}|$. This may be read off immediately from the perturbative diagrams themselves

Definition 27. *A connected perturbative diagram D for a vector $\vec{k}$ is a sub-leading diagram to a diagram D' for a vector $\vec{k}'$ if:*

- *D and D' have identically coloured vertices (implying $\vec{s}(\vec{k}) = \vec{s}(\vec{k}')$).*
- *D and D' have the same number of red edges modulo 2 (implying $a(\vec{k}) = a(\vec{k}')$).*
- *D' has fewer interaction vertices than D (implying $|\vec{k}| < |\vec{k}'|$).*

A diagram D is leading if it is not a sub-leading diagram to any D'.

Note that multiple leading diagrams may exist for a single parameter $\theta^a_{\vec{k}}$.

We now wish to construct a perturbative hierarchy by drawing all leading diagrams with $|\vec{k}| < K$ interaction vertices (for some sufficiently large K), and then ordering corresponding $T^a_{\vec{s}}$ by the leading-order contributions to $\theta_{\vec{s},a}$ we obtain via Eq. (5.5.29). However, this calculation requires the normalized coefficients $C_{\vec{k}}$, which in turn require computing the perturbative series for the normalization constant $\mathcal{N}$. To avoid this cumbersome normalization procedure and also to simplify Eq. (5.5.29), we suggest to approximate $\theta^{(\vec{k})}_{\vec{s}(\vec{k}),a(\vec{k})}$ by

$$\tilde{\theta}_{\vec{s},a} = \sum_{\substack{\text{leading } \vec{k}, \\ \vec{s}(\vec{k})=\vec{s},\; a(\vec{k})=a}} \tilde{\theta}^{(\vec{k})}_{\vec{s}(\vec{k}),\vec{a}(\vec{k})} \tag{5.5.38}$$

where we define

$$\sum_a i^{a-\Gamma(\vec{k})} \tilde{\theta}^{(\vec{k})}_{\vec{s}(\vec{k}),a} = \tilde{C}_{\vec{k}} - \sum_{\substack{f;\vec{K}(\vec{f})=\vec{k} \\ |\vec{N}(\vec{f})|>1}} \tilde{\Theta}(\vec{f}), \tag{5.5.39}$$

$$\tilde{\Theta}(\vec{f}) = \prod_{\vec{k},i} \frac{\left[\tilde{\theta}^{(\vec{k})}_i\right]^{f_i(\vec{k})}}{f_i(\vec{k})!}. \tag{5.5.40}$$

We expect that typically $\tilde{\theta}_{\vec{s},a} < \tilde{\theta}_{\vec{r},b} \leftrightarrow \theta_{\vec{s},a} < \theta_{\vec{r},b}$, which implies that this approximation should preserve the perturbative hierarchy.

We now have all the machinery required to define our perturbative hierarchy.

Definition 28. *Let $\{T_{\vec{s},a}\}$ be the generators for a matched, generating variational ansatz for a Hamiltonian $H = H_0 + \vec{J} \cdot \vec{V}$. The perturbative hierachy on $\{T_{\vec{s}}, a\}$ is defined by the total order*

$$T_{\vec{s},a} < T_{\vec{r},b} \text{ if } \tilde{\theta}_{\vec{s},a} < \tilde{\theta}_{\vec{r},b}, \tag{5.5.41}$$

and if $\tilde{\theta}_{\vec{s},a} = \tilde{\theta}_{\vec{r},b}$, we choose the ordering of $T_{\vec{s},a}$ and $T_{\vec{r},b}$ at random.

The explicit calculation of the $\tilde{\theta}_{\vec{s},a}$ variables is quite time consuming. As a shortcut, we note that $\tilde{\theta}^{(\vec{k})}_{\vec{s},a}$ scales as $\vec{J}^{\vec{k}}$, which, when $J_i \ll 1$ typically dominates any combinatorial terms. To formalize this, let us define

$$J_{\vec{s},a} = \sum_{\substack{\text{leading } \vec{k}, \\ \vec{s}(\vec{k})=\vec{s},\ a(\vec{k})=a}} \vec{J}^{\vec{k}}, \tag{5.5.42}$$

and we suggest to save on calculation by assuming $\theta_{\vec{s},a} < \theta_{\vec{r},b}$ when $J_{\vec{s},a} < J_{\vec{r},b}$.

5.6 Application: transverse-field Ising model

In this section, we demonstrate the construction of a variational hierarchy and study the resulting VQE performance on a target system. As a simple target example, we take the 1-dimensional transverse-field Ising model (TFIM):

$$H_{TFIM} = -\sum_{i}^{N_q} hZ_i + \sum_{i=1}^{N_q-1} JX_iX_{i+1}. \tag{5.6.1}$$

This system is a well-known prototype for condensed matter systems, being a non-interacting set of spins at $J = 0$, an Ising chain at $h = 0$, and demonstrating a quantum phase transition at $h = J$. For our example, we consider the $h \gg J > 0$ regime, and construct a perturbative hierarchy around $J = 0$, using the QCA as a parent ansatz. The noninteracting ground state may be immediately identified as the computational basis state $|\vec{0}\rangle$ with energy $-hN_q$, which we use as the starting state of our ansatz. Non-interacting excited states $|\vec{s}\rangle$ have energy $(2|\vec{s}| - N_q)h$.

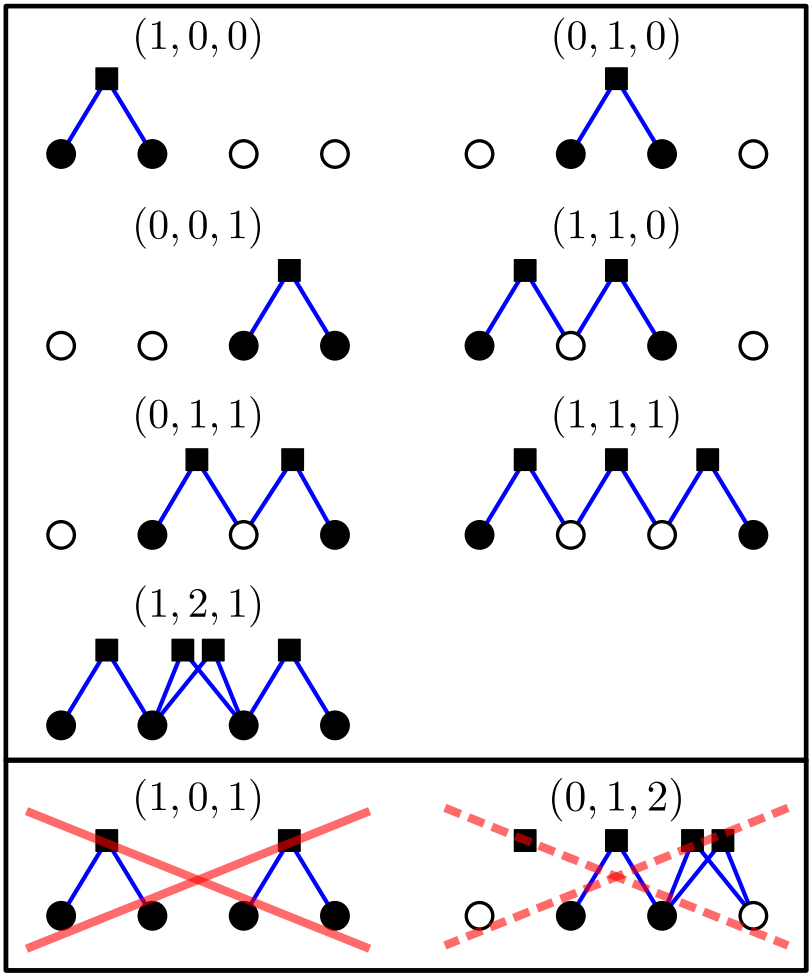

Figure 5.4: (top) The seven lowest-order connected diagrams for a four-site transverse-field Ising model, labeled by the $\vec{k}$ used in the text. (bottom) Examples of diagrams that do not need to be considered when constructing the perturbative hierarchy - (bottom left) a disconnected diagram that explicitly does not contribute to the hierarchy, and (bottom right) a diagram which will contribute to the same parameter in the hierarchy as a previous term ($\vec{k} = (1, 1, 0)$), but to lower order.

5.6.1 Example perturbative construction on four sites

To demonstrate the application of the methods developed in Sec. 5.5 in detail, we now construct the full perturbative hierarchy on a small chain ($N_q = 4$). This system has three perturbation terms, which we label $\hat{V}_i = X_i X_{i+1}$ for $i = 1, 2, 3$. These perturbations preserve the antiunitary complex conjugation symmetry $\mathcal{K}$, and the unitary global parity symmetry $Z_1 Z_2 Z_3 Z_4$. This reduces the required variational manifold dimension from $2^5 - 2 = 30$ to $2^3 - 1 = 7$ (both symmetries halve the Hilbert space dimension, but complex conjugation makes the phase equivalence redundant). In the QCA, this corresponds to removing all imaginary rotations (of the form $e^{i\theta X \ldots X}$), and all generators with an odd number of non-trivial terms. This removal will be automatic in the perturbative construction, as removed terms will never appear in the hierarchy, so we need only note the symmetries in case we 'run out' of terms to add to the variational ansatz *. The remaining generators are then

$$\begin{aligned} T_1 &= X_1 Y_2, \quad & T_2 &= X_2 Y_3, \quad & T_3 &= X_3 Y_4, \\ T_4 &= X_1 Y_3, \quad & T_5 &= X_2 Y_4, \quad & T_6 &= X_1 Y_4, \\ T_7 &= X_1 X_2 X_3 Y_4. \end{aligned}$$

For convenience in this small system, we will drop the stabilizer notation of Sec.5.3, and write the QCA as $\prod_{j=1}^{7} \exp(i\theta_j T_j)$. (For example, in the notation of Sec. 5.3 we would have written θ_6 as $\theta^4_{XII,1}$.)

To construct the perturbative hierachy, we proceed by drawing all lowest-order diagrams, and calculating the corresponding $\tilde{C}_{\vec{k}}$ contributions. In Fig. 5.4, we list the seven lowest-order connected diagrams in the system. This gives us the following:

1. 3 contributions at order J (to T_1, T_2, and T_3).
2. 2 contributions at order J^2 (to T_4 and T_5).
3. 1 contribution at order J^3 (to T_6).
4. 1 contribution at order J^4 (to T_7).

This may then be used as an initial guess for the ordering in the perturbative hierarchy. Importantly, although $\vec{k} = (1, 0, 1)$ is an order-J^2 term satisfying $\langle 0 | \vec{V}^{\vec{k}} T_7 | 0 \rangle \neq 0$, the corresponding diagram is disconnected (Fig. 5.4,

*Note that this is not always the case: if one must satisfy a symmetry by fixing parameters, both terms will appear in the hierarchy and the fixing must be done after the hierarchy is constructed.

bottom-left). This implies that its contribution to θ_7 will be cancelled out by the contributions of $(1,0,0)$ and $(0,0,1)$ (Theorem 26), and the diagram need not be considered in our construction, as we will confirm shortly. We further note that higher-order diagrams exist, e.g. that corresponding to $\vec{k} = (0,1,2)$ (Fig. 5.4, bottom-right). Although these have non-zero contribution to the actual value of the variational angles (in this case θ_2), as this contribution is at a higher-order of J we expect it to not affect the order of the hierarchy.

We now check the above ordering of the perturbative hierarchy by explicit calculation of the lowest-order contributions to $\tilde{\theta}_j$. Applying Eq. (5.5.14) recursively, the lowest-order connected contributions can be found to be (noting $S_{\vec{k},\vec{k}'} = 1$ as all V_i commute),

$$\begin{aligned}
\tilde{C}_{(1,0,0)} &= \tilde{C}_{(0,1,0)} = \tilde{C}_{(0,0,1)} = \frac{-1}{4h} \\
\tilde{C}_{(1,1,0)} &= \frac{-1}{4h}[\tilde{C}_{(1,0,0)} + \tilde{C}_{(0,1,0)}] = \frac{1}{8h^2}\left[= \tilde{C}_{(0,1,1)}\right] \\
\tilde{C}_{(1,1,1)} &= \frac{-1}{4h}[\tilde{C}_{(1,1,0)} + \tilde{C}_{(1,0,1)} + \tilde{C}_{(0,1,1)}] = -\frac{5}{64h^3} \\
\tilde{C}_{(1,2,1)} &= \frac{-1}{8h}[\tilde{C}_{(0,2,1)} + \tilde{C}_{(1,2,0)} + \tilde{C}_{(1,1,1)} \\
&\qquad - \tilde{C}_{(0,1,0)}\tilde{C}_{(1,0,1)}] = \frac{3}{256h^4}.
\end{aligned}$$

One may then calculate in turn the lowest-order approximation for the variational parameters via Eq. (5.5.39) (noting here that $\Gamma(\vec{k}) = 1$ for all $\vec{k}$ in this system).

$$\begin{aligned}
\tilde{\theta}_1 &= J\tilde{C}_{(1,0,0)} = \frac{-J}{4h}\left[= \tilde{\theta}_2 = \tilde{\theta}_3\right] \\
\tilde{\theta}_4 &= J^2\tilde{C}_{(1,1,0)} - \tilde{\theta}_1\tilde{\theta}_2 = \frac{J^2}{16h^2}\left[= \tilde{\theta}_5\right] \\
\tilde{\theta}_6 &= J^3\tilde{C}_{(1,1,1)} - \tilde{\theta}_1\tilde{\theta}_2\tilde{\theta}_3 - \tilde{\theta}_1\tilde{\theta}_5 - \tilde{\theta}_3\tilde{\theta}_4 = -\frac{J^3}{32h^3} \\
\tilde{\theta}_7 &= J^4\tilde{C}_{(1,2,1)} - \frac{1}{2}\tilde{\theta}_1\tilde{\theta}_2^2\tilde{\theta}_3 \\
&\qquad - \tilde{\theta}_4\tilde{\theta}_2\tilde{\theta}_3 - \tilde{\theta}_1\tilde{\theta}_2\tilde{\theta}_5 - \tilde{\theta}_4\tilde{\theta}_5 = -\frac{J^4}{512h^4}.
\end{aligned}$$

We see that ordering terms by $J_{\vec{s},a}$ reproduces the full perturbative hierarchy whenever $J < 2h$. We also note that the order J^2 contribution to $\tilde{\theta}_7$

from $\vec{k} = (1, 0, 1)$ is cancelled (following Theorem 26), as

$$J^2 \tilde{C}_{(1,0,1)} = \frac{J^2}{16h^2} = \tilde{\theta}_1 \tilde{\theta}_3. \tag{5.6.2}$$

We also note that the magnitudes of $\tilde{\theta}_i$ are systematically smaller than the magnitudes of corresponding perturbative terms $\vec{J}^{\vec{k}} \tilde{C}_{\vec{k}}$. This suggests that the back-action terms $\sum_{f;\vec{K}(\vec{f})=\vec{k};|\vec{N}(\vec{f})|>1} \Theta(\vec{f}) \langle 0 | \vec{V}^{\cdot \vec{k}\dagger} \left(i\vec{T}\right)^{\cdot \vec{N}(\vec{f})} |0\rangle$ in QCA may have a systematic positive effect on VQE convergence.

5.6.2 Low-order construction for a large chain

Following the analysis of the four-site example, we expect little to no deviation between parameters of the same order in a larger chain. Indeed, all first, second and third-order leading diagrams are identical up to translation along the chain (Fig. 5.5). As the on-site and interaction strengths are uniform along the chain, this implies that the coefficients for all such diagrams are likewise equal (to lowest-order). At fourth-order, two separate types of diagrams exist. One corresponds to $\vec{k} = (1, 2, 1)$ in the four-site model, and gives the same parameter estimate ($\tilde{\theta}_{\vec{s},a} = \frac{-J^4}{512h^4}$), to the QCA generators of the form $\{Y_i X_{i+1} X_{i+2} X_{i+3}\}$ The other was not present in the four-site model (as it requires 5 qubits) - it contributes a parameter estimate of $\tilde{\theta}_{\vec{s},a} = \frac{J^4}{128h^4}$ to QCA generators of the form $\{Y_i X_{i+4}\}$, placing these generators earlier in the perturbative hierarchy. The resulting ansatz thus needs only $5N_q - 13$ generators to reproduce the ground state with errors of order $(J/h)^5$. To obtain this level of accuracy with a classical calculation, one would in theory need to sum over all $(N_q - 1)^4$ combinations of individual perturbations. However, as clever grouping of terms (e.g. via tensor network contractions or similar) should reduce the time-cost of such a summation far below such numbers, this argument does not lead to an immediate guarantee of a quantum speedup for VQEs of this form.

5.6.3 Alternative hierarchies and circuit ordering

Although perturbation theory is a natural choice for developing variational hierarchies, it is not necessarily the only starting point. In the presence of strong interactions (where pertubation theory breaks down), other generator properties may provide better insight into how important they are at obtaining the ground state. In the following, we study the following

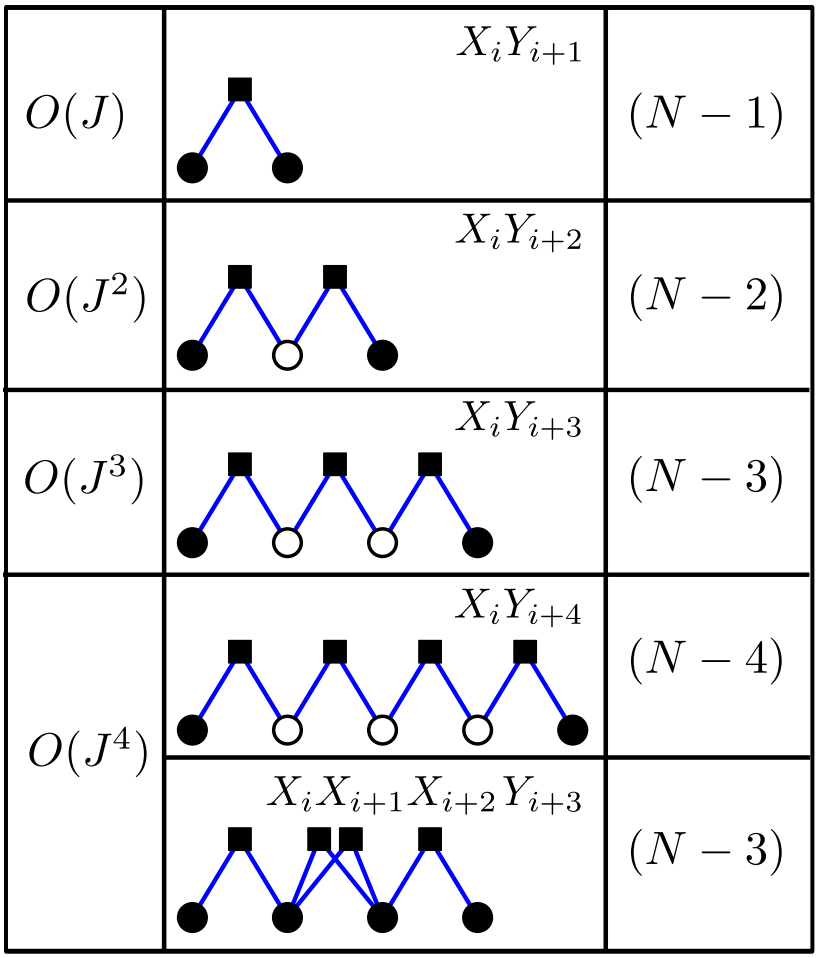

Figure 5.5: The leading connected diagrams to fourth-order on the transverse-field Ising model. Each diagram should be repeated across the entire N_q-qubit chain - the total number of copies of each diagram that will appear is written in the right-hand column. Diagrams are labelled by the generator $T_{\vec{s},a}$ that they contribute to.

natural constructions of a priority list, all of which use QCA as a parent ansatz:

- *pertQCA:* The perturbative hierarchy from Def. 28, using QCA as the parent variational ansatz.
- *revQCA:* The pertQCA hierarchy in reverse.
- *2-locQCA:* A low-weight variant of pertQCA, obtained by only allowing 2-local generators (those acting non-trivially on up to 2 qubits). When more generators are desired than in the final priority list, we loop over it repeatedly.
- *locQCA:* A geometrically local variant of pertQCA, obtained by only allowing generators acting on nearest neighbour pairs of qubits (and again looping over the priority list if required). This is equivalent to allowing only the generators which are dictated by the first-order perturbation theory, allowing for a generalization to an arbitrary Hamiltonian.

We have so far not discussed the ordering of the units within the ansatz circuit. Two natural choices present themselves: taking the order in which the gates appear in the priority list, and taking the order in which the gates appear in the parent ansatz. However, this is only well-defined when the priority list is inherited from a parent ansatz without repetition. For the above hierarchies that require looping, we only study the former choice, and denote by an asterisk results where the latter ordering is used.

5.6.4 VQE performance

We now test the performance of our variational hierarchies in different parameter regimes of the transverse-field Ising model on $N_q = 8$ sites. (Code to perform this investigation can be found at `github.com/tarrlikh/QSA`.) We take as a performance metric the relative energy error

$$\epsilon := (E_{\mathrm{VQE}} - E_0)/E_0, \tag{5.6.3}$$

where E_{VQE} is the energy of the converged VQE, and plot this as we increase the number N_p of parameters in the hierarchy. The hierarchy gives a natural strategy to perform the optimization - at each N_p, the optimized values of the previous $N_p - 1$ parameters are used as a starting guess for their new values (whilst the new parameter is initialized to 0). This approach converges much faster than re-starting each new simulation at

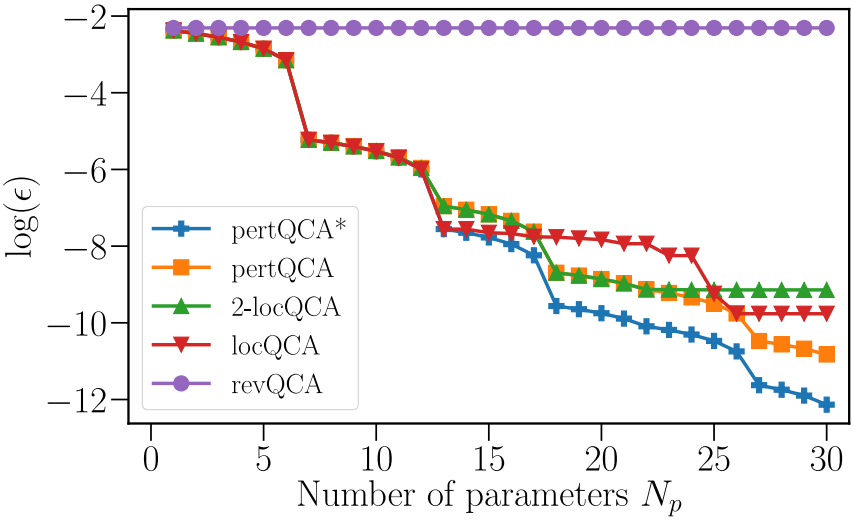

Figure 5.6: Log plot of the relative energy error ϵ (Eq. (5.6.3)) for different variational hierarchies, in a weakly-coupled transverse-field Ising model ($J/h = 0.15$). Error is plotted as a function of the number of parameters used (or equivalently the number of generators taken from the hierarchy). Description of the different hierarchies is given in the text.

the original value, as found previously in [54]. To focus on the performance of the ansatzes themselves, we do not include the effects of sampling noise or any experimental noise in our simulations.

We first investigate the weak-coupling regime where perturbation theory holds ($J/h = 0.15$). In Fig. 5.6, we plot the convergence of ϵ as the first 30 terms from all studied hierarchies are added consecutively. At each subsequent point we reoptimize all parameters using the SLSQP algorithm, starting from the local minimum found at the previous point. We observe that all hierarchies achieve good convergence, with the exception of revQCA, and that both variants of pertQCA achieve over an order of magnitude improvement over other ansatzes after 30 terms are added. We further observe that re-ordering the gates to follow the parent ansatz (pertQCA*) is preferable, leading to another order of magnitude improvement. We are unsure of the precise reason for this improvement, but suggest it may be attributed to the relatively large area of the variational manifold inherited from the parent ansatz, that may be lost under re-ordering. The discontinuities in the plot for pertQCA, pertQCA*, and

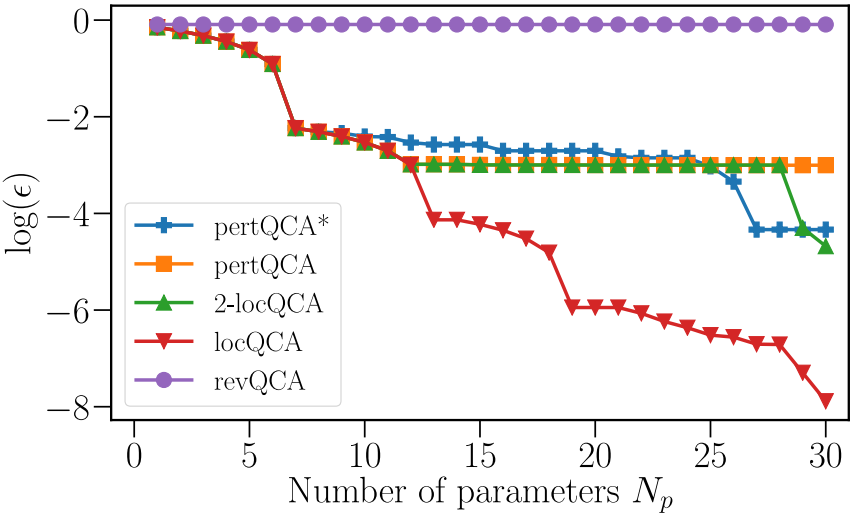

Figure 5.7: Similar convergence plot to Fig. 5.6, but in the strongly-coupled regime instead ($J/h = 6$).

2-locQCA correspond to the points where all gates up to a certain perturbation theory order have been included. This makes sense, as our theory predicts these points should correspond to the error decreasing from $O(J^n)$ to $O(J^{n+1})$.

We next investigate VQE convergence in the strongly correlated regime ($J/h = 6$). We observe that all hierarchies perform worse here than previously. We attribute this to the strongly-coupled ground state being further from the starting state than the weakly-coupled ground state. Note however, that one can obtain one of the two degenerate ground states at $h = 0$ from $|\vec{0}\rangle$ as

$$|E_0(h=0)\rangle = \prod_i e^{i\frac{\pi}{4}X_iY_{i+1}}|\vec{0}\rangle, \tag{5.6.4}$$

which is a rotation achievable after the first $N_q - 1 = 7$ terms of all considered hierarchies. This suggests that in all cases, the first order of the hierarchy is used to prepare this state, from which later orders perturb. Then, as perturbation theory around the strongly correlated ground state is significantly different to the perturbation theory around the non-interacting ground state, the generators we have chosen may not be optimal for this perturbation. This also explains the good performance of

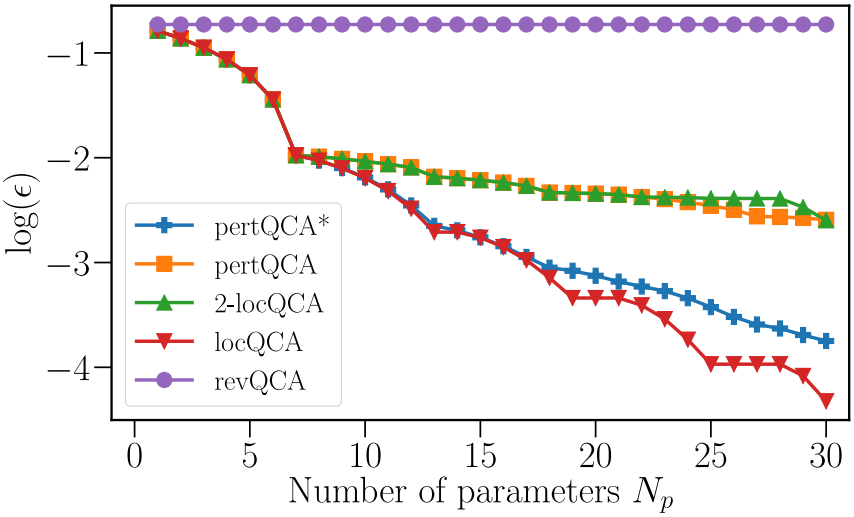

Figure 5.8: Similar convergence plot to Fig. 5.6, but in the critical regime instead ($J/h = 1$).

locQCA over the other hierarchies: by repeating local operators it ensures that it will obtain the lower orders (in h/J) of the true ground state.

We finally investigate the performance of our hierarchies in the critical regime ($J/h = 1$), where a transition between the strongly-correlated and weakly-correlated phases occurs in the thermodynamic limit. We observe that the relative error obtained by all ansatzes is the worst here, and that locQCA and pertQCA* behave similarly, obtaining up to an order of magnitude improvement over 2-locQCA and pertQCA. This loss of accuracy is not surprising, as we do not have a relatively cheap way of accessing any states perturbatively coupled to the ground state in the same manner as Eq. (5.6.4).

5.7 Conclusion

In this chapter, we have developed a diagrammatic framework for size-extensive variational quantum ansatzes, which avoids the use of Trotter-Suzuki approximation methods. We have described a large class of Pauli-generated product ansatzes demonstrably capable of spanning the entire Hilbert space with the minimum number of parameters necessary. We have demonstrated means by which one can compress ansatzes such as the

above to a practical size, by a perturbative treatment of the target system, and by taking into account any symmetries that exist. To ensure the size-extensivity of the construction, we have stated and proven the digital quantum version of the linked-cluster theorem. We have tested variants of the resulting ansatzes on the transverse-field Ising model, finding that their performance in various regimes matches our expectations based on their means of construction. We observe that ansatzes that fully match the perturbation theory give a benefit in the weak coupling regime as expected. However, in the strong-coupling regime, focusing on the locality of the ansatz at the expense of perturbation theory considerations appears to be preferred.

As is well known in the field, the performance of any VQE ansatz is system dependent. Ansatzes that are derived from perturbative physical principles can be expected to perform best when perturbation theory converges well. By contrast, those founded on adiabatic principles (e.g. the variational Hamiltonian ansatz [56]) can be expected to perform best on systems with a large gap. As these two conditions are often correlated (e.g. a gap closing often corresponds to a phase transition and a breakdown of perturbation theory), a fair comparison of ansatzes based on these two principles (and with any other ansatzes) would require an extensive numerical study. This is an obvious target for future research.

We have avoided in the above any discussion of a quantum speedup for the VQEs that we have constructed in this chapter. To the best of our knowledge this remains an open and difficult question to show for any class of VQEs. Informally, to demonstrate a quantum speedup, one requires to be able to obtain an estimate of the true ground state energy E for an N_q-qubit system, within an error ϵ, in time polynomial in N_q. This also needs to be achieved in a class of N_q-qubit systems for which no similar estimation is possible classically. The circuit length in a variational hierarchy grows polynomially in the number of parameters N_p, so it would be sufficient to show that the error $\epsilon(N_p, N)$ scales polynomially in N_p and N_q. One also needs to consider the time cost of measuring the energy (which grows polynomially in N_q) and the time cost of optimization (which grows polynomially in N_p). Our results appear to show this behavior; we observe what appears to be exponential decay in N_p for all three systems studied. (Note that the measurement and optimization requirements imply that the time cost to extract these energies from the device will still be at best polynomial.) However, 1D spin chains such as the transverse-field Ising model are well accessible by classical methods and polynomial-time algorithms are known for any weakly-coupled 2-local spin system [173], so we do not expect a quantum speedup in this case. Finding target

systems for which a speedup may be demonstrable, and further optimizing hierarchy construction to show this, are obvious targets for future research.

5.A Background

Definition 29. *The state of an N_q-qubit quantum register is represented by a norm-1 vector in the Hilbert space $\mathcal{H} = \mathbb{C}^{2^{N_q}}$, under the association $|\psi\rangle \in \mathcal{H} \equiv e^{i\phi}|\psi\rangle$ for $\phi \in \mathbb{R}$.*

Definition 30. *The Pauli basis on N_q qubits is defined as $\mathbb{P}^{N_q} := \{I, X, Y, Z\}^{\otimes N_q}$, where I, X, Y, Z are the 2×2 matrices on $\mathbb{C}^2$:*

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
$$Y = \begin{pmatrix} 0 & -1i \\ 1i & 0 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{5.A.1}$$

and $\otimes$ is the Kronecker tensor product.

$\mathbb{P}^{N_q}$ has the following nice properties:

1. $P^2 = 1$ for all $P \in \mathbb{P}^{N_q}$.
2. For $P, Q \in \mathbb{P}^{N_q}$, either $[P, Q] := PQ - QP = 0$, or $\{P, Q\} := PQ + QP = 0$, and P commutes with precisely half of $\mathbb{P}^{N_q}$.
3. $P \in \mathbb{P}^{N_q} \neq 1$ has only two eigenvalues, ± 1, and the dimension of the corresponding eigenspaces is precisely $2^{N_q - 1}$ (i.e. each P divids $\mathbb{C}^{2^{N_q}}$ in two).
4. This division by two may be further continued - given $P, Q \neq 1$ such that $[P, Q] = 0$, P and Q divide the Hilbert space into 4 eigenspaces (labeled by combinations of their eigenvalues).
5. To generalize, one can form a $[N_q, k]$ **stabilizer group** $\mathcal{S}$, generated by k Hermitian, commuting, non-generating elements of $\mathbb{P}^{N_q}$ (up to a complex phase); this diagonalizes $\mathbb{C}^{2^{N_q}}$ into 2^k unique eigensectors of dimension $2^{N_q - k}$. When $N_q = k$, these sectors contain single eigenstates, which we call **stabilizer states** [170].
6. Given such a stabilizer state $|\psi\rangle$ and Hermitian $P \in \mathbb{P}^{N_q}$, either $P|\psi\rangle = \pm|\psi\rangle$ or $\langle\psi|P|\psi\rangle = 0$.

The Pauli basis is a basis for the set of $2^{N_q} \times 2^{N_q}$ complex-valued matrices (hence the name); it is also a basis for the set of Hermitian matrices if one chooses real coefficients. However, it is not a group under matrix multiplication, as the single-qubit Pauli matrices pick up a factor of i

on multiplication - $XY = iZ \notin \mathbb{P}$. The closure of the Pauli basis is the Pauli group $\Pi^{N_q} = \{\pm i\} \times \mathbb{P}^{N_q}$; this is four times as large, and no longer has the basis properties of $\mathbb{P}^{N_q}$. The Pauli basis inherits a form of multiplication from Π^{N_q} - $P \cdot Q = R \in \mathbb{P}^{N_q}$ if $PQ = e^{i\phi}R \in \Pi^{N_q}$, at which point $\mathbb{P}^{N_q} \equiv D_2^{N_q}$. However, under this multiplication $\mathbb{P}^{N_q}$ becomes a commutative group, which sacrifices key information about its operator structure. Based on the second point in the above list, we may make the following useful definition:

Definition 31. *The **relative sign** of $P, Q \in \mathbb{P}^{N_q}$, $s_{P,Q} \in \{-1, 1\}$, is defined such that $PQ + s_{P,Q}QP = 0$. We further define the **markers** $\delta_{P,Q} = (1 + s_{P,Q})/2$, $\bar{\delta}_{P,Q} = (1 - s_{P,Q})/2 = 1 - \delta_{P,Q}$.*

This allows us to write the following useful identity:

$$e^{i\theta P}Q = Qe^{is_{P,Q}P}. \tag{5.A.2}$$

Unfortunately this does not extend to the commutation of two such exponentials; one has instead by the application of the Baker-Campbell-Hausdorff formula

$$\begin{aligned} e^{i\theta P}e^{i\phi Q} &= e^{i\phi e^{i\theta P/2}Qe^{-i\theta P/2}}e^{i\theta P}, && (5.A.3)\\ &= e^{i\phi\left[\delta_{P,Q}Q + \bar{\delta}_{P,Q}(\cos(\theta)P + \sin(\theta)PQ)\right]}e^{i\theta P} && (5.A.4)\\ &= e^{i\phi Q}e^{i\theta[\delta_{P,Q}P + \bar{\delta}_{P,Q}(\cos(\theta)P + \sin(\theta)PQ)]}. && (5.A.5) \end{aligned}$$

and the exponential expression cannot be simplified unless $\theta = n\pi/2$. In this special case, $e^{i\pi/2P}$ is a Clifford operator (being an operator that maps Pauli operators to Pauli operators); this does not define all Clifford operators, but the set $\{e^{i\pi/2P}, P \in \mathbb{P}^{N_q}\}$ does generate the Clifford group.

5.B Example of compression over symmetries: the unitary coupled cluster ansatz

As an example of symmetry-induced compression, let us construct the Trotterized unitary coupled cluster ansatz [47, 159] on a fermionic system. This can be done by taking the Pauli-type ansatz of local Majorana operators acting on an equal number of empty and filled orbitals, removing terms that do not respect $\mathcal{K}$, and fixing the remainder to respect the fermion parity. We now detail this procedure.

The UCC ansatz takes the form

$$U(\vec{\theta}) = e^{T(\vec{\theta}) - T^\dagger(\vec{\theta})}, \tag{5.B.1}$$

where the operator $T(\vec{\theta})$ is a sum of n-th order cluster operators $T^{(n)}(\vec{\theta})$ between filled states i and empty states j of the non-interacting problem.

$$T^{(n)}(\vec{\theta}) = \sum_{i_1,\ldots,i_n,j_1,\ldots,j_n} \theta^{j_1,\ldots,j_n}_{i_1,\ldots,i_n}\, \hat{c}^\dagger_{j_1} \ldots \hat{c}^\dagger_{j_n} \hat{c}_{i_1} \ldots \hat{c}_{i_n}. \tag{5.B.2}$$

The choice of $T(\vec{\theta}) - T^\dagger(\vec{\theta})$ is made to respect $\mathcal{K}$ (as creation and annihilation operators are real). One typically takes only a few $T^{(n)}$ (usually up to $n = 2$), and Trotterizes the resulting expression in terms of individual excitations to implement on a quantum computer, in which case it becomes a product ansatz. $\hat{c}^\dagger_j$ and $\hat{c}_j$ are the fermionic creation and annihilation operators for the jth orbital. These are not themselves Pauli operators, but they may be combined to make Majorana operators

$$\gamma^{(0)}_j = \hat{c}^\dagger_j + \hat{c}_j, \gamma^{(1)}_j = i(\hat{c}^\dagger_j - \hat{c}_j), \tag{5.B.3}$$

which are elements of $\mathbb{P}^{N_q}$ (up to a possible sign). (One can show this immediately upon choosing a mapping from fermions to qubits.) The fermionic number operator, $N = \sum_j \hat{c}^\dagger_j \hat{c}_j$, is equivalent to $\Gamma = \sum_j \gamma^{(0)}_j \gamma^{(1)}_j$ (for commutation purposes). To form the operator $T^{(1)} - T^{(1)\dagger}$, one may take the set of excitations $e^{i\theta^{j,a}_{i,a} \gamma^{(a)}_i \gamma^{(a)}_j}$ for $i \neq j$ (and $a = 0, 1$), and enforce the symmetry by fixing $\theta^{j,a}_{i,a} = \theta^{j,1-a}_{i,1-a}$. (Terms of the form $\gamma^0_i \gamma^1_j$ do not commute with $\mathcal{K}$.) The second-order cluster operator is slightly more complicated; one must take all terms of the form

$$\exp\left(i\theta^{j_1,j_2,b_1,b_2}_{i_1,i_2,a_1,a_2} \gamma^{a_1}_{i_1} \gamma^{a_2}_{i_2} \gamma^{b_1}_{j_1} \gamma^{b_2}_{j_2}\right), \tag{5.B.4}$$

with $i_1 \neq i_2$ ($j_1 \neq j_2$) operators for empty (filled) states, and $\sum_i a_i + b_i = 1 \mod 2$ (terms where $\sum_i a_i + b_i = 0 \mod 2$ do not commute with $\mathcal{K}$). Then, to conserve Γ, one must fix

$$\begin{aligned}\theta^{j_1,j_2,0,1}_{i_1,i_2,0,0} = \theta^{j_1,j_2,1,0}_{i_1,i_2,0,0} = -\theta^{j_1,j_2,0,0}_{i_1,i_2,1,0} = -\theta^{j_1,j_2,0,0}_{i_1,i_2,0,1} \\ = \theta^{j_1,j_2,1,1}_{i_1,i_2,0,1} = \theta^{j_1,j_2,1,1}_{i_1,i_2,1,0} = -\theta^{j_1,j_2,0,1}_{i_1,i_2,1,1} = -\theta^{j_1,j_2,1,0}_{i_1,i_2,1,1}.\end{aligned}$$

(One can confirm that all operators being fixed commute here, as required.) This procedure may be continued as needed to obtain higher-order cluster operators.

One might try to use the tools developed above and check if the Trotterized UCC ansatz tightly spans the reduced Hilbert space. On the one hand, the number of parameters in the full UCC,

$$\sum_{n=1}^{\eta} \frac{\eta!}{(\eta-n)!n!} \frac{(N_q-\eta)!}{(N_q-\eta-n)!n!} = \frac{N_q!}{(N_q-\eta)!\eta!} - 1, \tag{5.B.5}$$

does match precisely the dimension of a real Hilbert space with η particles in N_q orbitals. On the other hand, as the Trotterized UCC Jacobian is full-rank at $\vec{\theta} = \vec{0}$, we strongly suspect that it spans this Hilbert state. However, we did not find a definitive proof of this. In particular, Trotterized UCC is not a stabilizer ansatz, and we have not found an obvious construction of a stabilizer ansatz from UCC.

5.C Multivariate Dyson series

To prove the statement of Lemma 16, we need to analyze the multi-parameter expansion (5.5.10) of the ground state $|E_0\rangle$, as a perturbative solution to the corresponding eigenvalue equation

$$(H_0 + JV)|E_0\rangle = E_0|E_0\rangle. \tag{5.C.1}$$

It proves to be convenient to first find an unnormalized solution $|\tilde{E}_0\rangle$ whose expansion states $|\tilde{\Psi}_{\vec{k}}\rangle$ (cf. (5.5.12)) obey a special condition:

$$\langle\tilde{\Psi}_{\vec{0}}|\tilde{\Psi}_{\vec{k}}\rangle = \delta_{\vec{k},\vec{0}}. \tag{5.C.2}$$

The properly normalized ground state $|E_0\rangle$ is then to be obtained as $|E_0\rangle = \mathcal{N}|\tilde{E}_0\rangle$, for $\mathcal{N} = (\langle\tilde{E}_0|\tilde{E}_0\rangle)^{-1/2}$.

To find $|\tilde{\Psi}_{\vec{k}}\rangle$, one can use the Dyson series-like approach. For this, one rewrites (5.C.1) as:

$$(E_0^{(0)} - H_0)|\tilde{E}_0\rangle = (JV - \Delta)|\tilde{E}_0\rangle, \tag{5.C.3}$$

for $E_0^{(0)}$ being the unperturbed ground state energy, and quantity Δ defined as follows:

$$\Delta \equiv (E_0 - E_0^{(0)}) = \langle\tilde{\Psi}_{\vec{0}}|JV|\tilde{E}_0\rangle. \tag{5.C.4}$$

Eq. (5.C.3) can be rewritten as:

$$|\tilde{E}_0\rangle = |\tilde{\Psi}_{\vec{0}}\rangle + (E_0^{(0)} - H_0)^{-1}(JV - \Delta)|\tilde{E}_0\rangle, \tag{5.C.5}$$

where the action of the inverse operator $(E_0^{(0)} - H_0)^{-1}$ is well-defined since the state $(JV - \Delta)|\tilde{E}_0\rangle$ has no overlap with $|\tilde{\Psi}_{\vec{0}}\rangle$ (cf. (5.C.4) and (5.C.2)). Using expansion (5.5.12) and the form of perturbation $JV = \vec{J}\cdot\vec{V}$, one recovers from (5.C.5) a set of equations on $|\tilde{\Psi}_{\vec{k}}\rangle$ for all $\vec{k} \neq \vec{0}$:

$$|\tilde{\Psi}_{\vec{k}}\rangle = G_0 \left(\sum_\beta V_\beta |\tilde{\Psi}_{\vec{k}-\vec{\delta}_\beta}\rangle - \sum_{\vec{k}'+\vec{k}''=\vec{k}} \Delta_{\vec{k}'} |\tilde{\Psi}_{\vec{k}''}\rangle \right), \tag{5.C.6}$$

$$G_0 \equiv (E_0^{(0)} - H_0)^{-1}, \Delta_{\vec{k}} \equiv \sum_\beta \langle \tilde{\Psi}_{\vec{0}} | V_\beta | \tilde{\Psi}_{\vec{k}-\vec{\delta}_\beta}\rangle, \tag{5.C.7}$$

for $\vec{\delta}_\beta$ the unit vector with the β component equal to 1. Note, that the action of G_0 here is again well-defined, since it acts on a state which has a zero overlap with $|\tilde{\Psi}_{\vec{0}}\rangle$ (cf. (5.C.7) and (5.C.2)). Now, with (5.C.6), we expressed each state $|\tilde{\Psi}_{\vec{k}}\rangle$ in terms of states $|\tilde{\Psi}_{\vec{k}'}\rangle$ which belong to lower PT orders: $|\vec{k}'| < |\vec{k}|$. Using (5.C.6) and the unperturbed ground state $|\tilde{\Psi}_{\vec{0}}\rangle = |\vec{0}\rangle$, one can obtain all the states $|\tilde{\Psi}_{\vec{k}}\rangle$ up to any desired order.

Given the states $|\tilde{\Psi}_{\vec{k}}\rangle$, one can also find the expression for the normalization $\mathcal{N}$, as a multi-parameter series:

$$\mathcal{N} = \sum_{\vec{k}} \mathcal{N}_{\vec{k}} \vec{J}^{\vec{k}} \tag{5.C.8}$$

The expansion states $|\Psi_{\vec{k}}\rangle$ of the normalised ground state $|E_0\rangle$ are then given by:

$$|\Psi_{\vec{k}}\rangle = \sum_{\vec{k}'+\vec{k}''=\vec{k}} \mathcal{N}_{\vec{k}''} |\tilde{\Psi}_{\vec{k}'}\rangle \tag{5.C.9}$$

With this scheme for finding the expansion states $|\Psi_{\vec{k}}\rangle$, we're ready to prove Lemma 16. To do so, first we will use (5.C.6) and prove the validity of the expression (5.5.13), together with the recursive relation (5.5.14). Then, using (5.C.9), we will extend our proof also to the states $|\Psi_{\vec{k}}\rangle$, recovering the statement of Lemma 16.

Proof — We start with a proof of the relation (5.5.13) for the states $|\tilde{\Psi}_{\vec{k}}\rangle$, by induction in PT order $|\vec{k}|$. We first note that for $|\vec{k}| = 0$, we have a single state $|\tilde{\Psi}_{\vec{k}=\vec{0}}\rangle = |\vec{0}\rangle$ that clearly satisfies (5.5.13) - this will be the base of our induction. Next, we have to prove (5.5.13) for $|\tilde{\Psi}_{\vec{k}}\rangle$ with an arbitrary $\vec{k}$, assuming the validity of (5.5.13) for all $|\tilde{\Psi}_{\vec{k}'}\rangle$ s.t. $|\vec{k}'| < |\vec{k}|$. To do so, let us express $|\tilde{\Psi}_{\vec{k}}\rangle$ using (5.C.6) and show that the different

terms that are present on the r.h.s. are proportional to the state $\vec{V}^{\cdot\vec{k}}|\vec{0}\rangle$ with a real coefficient. The terms of the type $G_0 V_\beta |\tilde{\Psi}_{\vec{k}-\vec{\delta}_\beta}\rangle$, assuming expression (5.5.13) for $|\tilde{\Psi}_{\vec{k}-\vec{\delta}_\beta}\rangle$, can be rewritten as:

$$G_0 V_\beta |\tilde{\Psi}_{\vec{k}-\vec{\delta}_\beta}\rangle = G_0 \tilde{C}_{\vec{k}-\vec{\delta}_\beta} V_\beta \vec{V}^{\cdot(\vec{k}-\vec{\delta}_\beta)}|\vec{0}\rangle \tag{5.C.10}$$

$$= \frac{S_{\vec{\delta}_\beta,\vec{k}-\delta_\beta}\tilde{C}_{\vec{k}-\vec{\delta}_\beta}}{E^{(0)}_{\vec{0}} - E^{(0)}_{\vec{s}(\vec{k})}} \vec{V}^{\cdot\vec{k}}|\vec{0}\rangle. \tag{5.C.11}$$

The other contributions to the r.h.s. of (5.C.6) are of the form $G_0 \Delta^{(\vec{k}')}|\tilde{\Psi}_{\vec{k}''}\rangle$, such that $\vec{k}' + \vec{k}'' = \vec{k}$. The factor $\Delta^{(\vec{k}')}$ here can be rewritten using the assumption of induction:

$$\Delta_{\vec{k}'} = \sum_\beta \langle\vec{0}|\tilde{C}^{(\vec{k}'-\vec{\delta}_\beta)} V_\beta \vec{V}^{\cdot(\vec{k}'-\vec{\delta}_\beta)}|\vec{0}\rangle \tag{5.C.12}$$

$$= \left(\sum_\beta S_{\vec{\delta}_\beta,\vec{k}-\delta_\beta}\tilde{C}_{\vec{k}'-\vec{\delta}_\beta}\right)\langle\vec{0}|\vec{V}^{\cdot\vec{k}'}|\vec{0}\rangle \tag{5.C.13}$$

$$= \Delta^{\mathrm{Re}}_{\vec{k}'}\langle\vec{0}|\vec{V}^{\cdot\vec{k}'}|\vec{0}\rangle, \tag{5.C.14}$$

where we introduced the shorthand notation $\Delta^{\mathrm{Re}}_{\vec{k}'}$ for the real coefficient $\left(\sum_\beta S_{\vec{\delta}_\beta,\vec{k}-\delta_\beta}\tilde{C}^{(\vec{k}'-\vec{\delta}_\beta)}\right)$. With this observation about $\Delta_{\vec{k}'}$ and the assumption of induction at hand, the following manipulation can be performed:

$$G_0\Delta_{\vec{k}'}|\tilde{\Psi}_{\vec{k}''}\rangle = \Delta^{\mathrm{Re}}_{\vec{k}'}\tilde{C}_{\vec{k}''}G_0\vec{V}^{\cdot\vec{k}''}|\vec{0}\rangle\langle\vec{0}|\vec{V}^{\cdot\vec{k}'}|\vec{0}\rangle \tag{5.C.15}$$

$$= \frac{\Delta^{\mathrm{Re}}_{\vec{k}'}\tilde{C}_{\vec{k}''}S_{\vec{k}'',\vec{k}'}}{E^{(0)}_{\vec{0}} - E^{(0)}_{\vec{s}(\vec{k})}}\delta_{\vec{s}(\vec{k}'),\vec{0}}\vec{V}^{\cdot\vec{k}}|\vec{0}\rangle, \tag{5.C.16}$$

where we used the condition $\vec{k}'+\vec{k}'' = \vec{k}$. Combining (5.C.11) and (5.C.16), we see that the expression (5.C.6) indeed implies the form (5.5.13) of $|\tilde{\Psi}_{\vec{k}}\rangle$, with a real coefficient $\tilde{C}_{\vec{k}}$ which is given by the formula (5.5.14).

Before extending this result to the coefficient states $|\Psi_{\vec{k}}\rangle$ of the normalized ground state $|E_0\rangle = \mathcal{N}|\tilde{E}_0\rangle$, we will need to make an aside and prove the following property of the coefficients $\mathcal{N}_{\vec{k}}$:

$$\mathcal{N}_{\vec{k}} = \mathcal{N}^{\mathrm{Re}}_{\vec{k}}\langle\vec{0}|\vec{V}^{\cdot\vec{k}}|\vec{0}\rangle, \tag{5.C.17}$$

for a real coefficient $\mathcal{N}^{\mathrm{Re}}_{\vec{k}}$. First, one can observe that an analogous property holds for the coefficients $Z_{\vec{k}}$ of $Z \equiv \langle \tilde{E}_0 | \tilde{E}_0 \rangle = \mathcal{N}^{-2}$:

$$Z = \sum_{\vec{k}} \vec{J}^{\vec{k}} Z_{\vec{k}}, \tag{5.C.18}$$

$$Z_{\vec{k}} = \sum_{\vec{k}'+\vec{k''}=\vec{k}} \tilde{C}_{\vec{k}'} \tilde{C}_{\vec{k}''} \langle \vec{0} | \left(\vec{V}^{\cdot \vec{k}''} \right)^{\dagger} \vec{V}^{\cdot \vec{k}'} | \vec{0} \rangle \tag{5.C.19}$$

$$= \sum_{\vec{k}'+\vec{k''}=\vec{k}} i^{2\Gamma_{\vec{k}''}} S_{\vec{k}'',\vec{k}'} \tilde{C}_{\vec{k}'} \tilde{C}_{\vec{k}''} \langle \vec{0} | \vec{V}^{\cdot \vec{k}} | \vec{0} \rangle \tag{5.C.20}$$

$$= Z^{\mathrm{Re}}_{\vec{k}} \langle \vec{0} | \vec{V}^{\cdot \vec{k}} | 0 \rangle, \tag{5.C.21}$$

with a real coefficient $Z^{\mathrm{Re}}_{\vec{k}}$ defined as $\sum_{\vec{k}'+\vec{k''}=\vec{k}} i^{2\Gamma_{\vec{k}''}} S_{\vec{k}'',\vec{k}'} \tilde{C}_{\vec{k}'} \tilde{C}_{\vec{k}''}$; in this derivation, we used (5.5.11) for states $|\tilde{\Psi}_{\vec{k}}\rangle$. Now, observe that $Z_{\vec{0}} = 1$, which means that the norm $\mathcal{N} = Z^{-1/2} = (1+\epsilon)^{-1/2}$ can be expressed as a Taylor series in $\epsilon = \sum_{\vec{k} \neq 0} \vec{J}^{\vec{k}} Z_{\vec{k}}$, which is a quantity of order $O(J)$. Expanding the terms of such Taylor series, one observes that the coefficients $\mathcal{N}_{\vec{k}}$ are given in terms of products of coefficients $Z_{\vec{k}}$ such that the combined perturbation theory order $\vec{k}$ is conserved - for example, a product $Z_{\vec{k}_1} Z_{\vec{k}_2}$ will contribute to $\mathcal{N}_{\vec{k}_1+\vec{k}_2}$. This allows to obtain the property (5.C.17) from (5.C.21) term by term. For instance, $Z_{\vec{k}_1} Z_{\vec{k}_2}$ is proportional to $\langle \vec{0} | \vec{V}^{\cdot(\vec{k}_1+\vec{k}_2)} | \vec{0} \rangle$ with a real coefficient:

$$Z_{\vec{k}_1} Z_{\vec{k}_2} = Z^{\mathrm{Re}}_{\vec{k}_1} Z^{\mathrm{Re}}_{\vec{k}_2} \langle \vec{0} | \vec{V}^{\cdot \vec{k}_1} | \vec{0} \rangle \langle \vec{0} | \vec{V}^{\cdot \vec{k}_2} | \vec{0} \rangle \tag{5.C.22}$$

$$= \delta_{\vec{s}(\vec{k_1}),\vec{0}} \delta_{\vec{s}(\vec{k_2}),\vec{0}} S_{\vec{k}_1,\vec{k}_2} Z^{\mathrm{Re}}_{\vec{k}_1} Z^{\mathrm{Re}}_{\vec{k}_2} \langle \vec{0} | \vec{V}^{\cdot(\vec{k}_1+\vec{k}_2)} | \vec{0} \rangle. \tag{5.C.23}$$

This statement can be directly extended to any product of multiple $Z_{\vec{k}}$'s, recovering (5.C.17), as desired.

To prove expression (5.5.11), we simply use the property (5.C.17) and (5.5.13) for $|\tilde{\Psi}_{\vec{k}}\rangle$, in the formula (5.C.9):

$$|\Psi_{\vec{k}}\rangle = \sum_{\vec{k}'+\vec{k}''=\vec{k}} \mathcal{N}^{(\vec{k}'')}|\tilde{\Psi}_{\vec{k}'}\rangle \tag{5.C.24}$$

$$= \sum_{\vec{k}'+\vec{k}''=\vec{k}} \mathcal{N}^{\mathrm{Re}}_{\vec{k}''}\tilde{C}_{\vec{k}'}\vec{V}^{\cdot\vec{k}'}|\vec{0}\rangle\langle\vec{0}|\vec{V}^{\cdot\vec{k}''}|\vec{0}\rangle \tag{5.C.25}$$

$$= \sum_{\vec{k}'+\vec{k}''=\vec{k}} \delta_{\vec{s}(\vec{k}''),\vec{0}}S_{\vec{k}',\vec{k}''}\mathcal{N}^{\mathrm{Re}}_{\vec{k}''}\tilde{C}_{\vec{k}'}\vec{V}^{\cdot\vec{k}}|\vec{0}\rangle \tag{5.C.26}$$

$$= C_{\vec{k}}\vec{V}^{\cdot\vec{k}}|\vec{0}\rangle, \tag{5.C.27}$$

$$C_{\vec{k}} \equiv \sum_{\vec{k}'+\vec{k}''=\vec{k}} \delta_{\vec{s}(\vec{k}''),\vec{0}}S_{\vec{k}',\vec{k}''}\mathcal{N}^{\mathrm{Re}}_{\vec{k}''}\tilde{C}_{\vec{k}'}. \tag{5.C.28}$$

This concludes our proof of Lemma 16. □

5.D Separability of disconnected contributions

In what follows, we prove Lemma 19.

Proof — Consider a disconnected contribution $|\Psi_{\vec{k}}\rangle = C_{\vec{k}}\vec{V}^{\cdot\vec{k}}|\vec{0}\rangle$ to the ground state $|E_0\rangle$ of the Hamiltonian $H = H_0 + \vec{J}\cdot\vec{V}$, with a corresponding splitting $\vec{k} = \vec{k}_A + \vec{k}_B$. The two sets of couplings that are activated, respectively, in $\vec{k}_A$ and $\vec{k}_B$, we will denote A and B. We also introduce two non-intersecting sets of qubits, Q_A and Q_B, such that they include, respectively, the supports of $\vec{k}_A$ and $\vec{k}_B$, and their union $Q_A \cup Q_B$ constitutes the whole set of qubits.

Let us consider an auxilliary Hamiltonian H', which is equal to H with a constraint $J_i = 0$ for all couplings V_i which are not in $A \cup B$. In the PT series for the ground state $|E_0\rangle'$ of such an auxilliary Hamiltonian,

$$|E_0\rangle' = \sum_{\vec{k}'} \vec{J}^{\vec{k}'}C'_{\vec{k}'}\vec{V}^{\cdot\vec{k}'}|\vec{0}\rangle, \tag{5.D.1}$$

the terms $C'_{\vec{k}'}$ are equal to the corresponding terms $C_{\vec{k}'}$ in the full series (5.5.10) – namely those, where no couplings V_i are activated besides those in $A \cup B$. In particular, (5.D.1) still contains the disconnected contribution of interest, $C'_{\vec{k}'=\vec{k}} = C_{\vec{k}}$.

On the other hand, H' is a sum of two independent Hamiltonians, defined on subsystems Q_A and Q_B:

$$H' = H'_A \otimes \mathbb{I}_{Q_B} + \mathbb{I}_{Q_A} \otimes H'_B, \tag{5.D.2}$$

$$H'_A \equiv -\sum_{i \in Q_A} h_i Z_i + \sum_{i \in A} J_i V_i, \tag{5.D.3}$$

$$H'_B \equiv -\sum_{i \in Q_B} h_i Z_i + \sum_{i \in B} J_i V_i. \tag{5.D.4}$$

This implies that the ground state $|E_0\rangle'$, will be a tensor product of the ground states of H'_A and H'_B,

$$|E_0\rangle' = |E_0\rangle'_A |E_0\rangle'_B. \tag{5.D.5}$$

In turn, the subsystem ground states $|E_0\rangle'_A$ and $|E_0\rangle'_B$ can themselves be written as PT series in couplings restricted on A and B, separately:

$$|E_0\rangle'_A = \sum_{\vec{k}'_A} \vec{J}^{\vec{k}'_A} C'_{\vec{k}'_A} \vec{V}^{\cdot \vec{k}'_A} |\vec{0}\rangle_{Q_A}, \tag{5.D.6}$$

$$|E_0\rangle'_B = \sum_{\vec{k}'_B} \vec{J}^{\vec{k}'_B} C'_{\vec{k}'_B} \vec{V}^{\cdot \vec{k}'_B} |\vec{0}\rangle_{Q_B}, \tag{5.D.7}$$

whose terms, again, are identical to those in the full series (5.5.10), with only couplings from A (B) activated: $C'_{\vec{k}'_A} = C_{\vec{k}'_A}$ ($C'_{\vec{k}'_B} = C_{\vec{k}'_B}$). Combining (5.D.1), (5.D.5), (5.D.6) and (5.D.7), for our term of interest $C_{\vec{k}}$ we obtain the desired relation:

$$C_{\vec{k}} = C_{\vec{k}_A} C_{\vec{k}_B}. \tag{5.D.8}$$

□

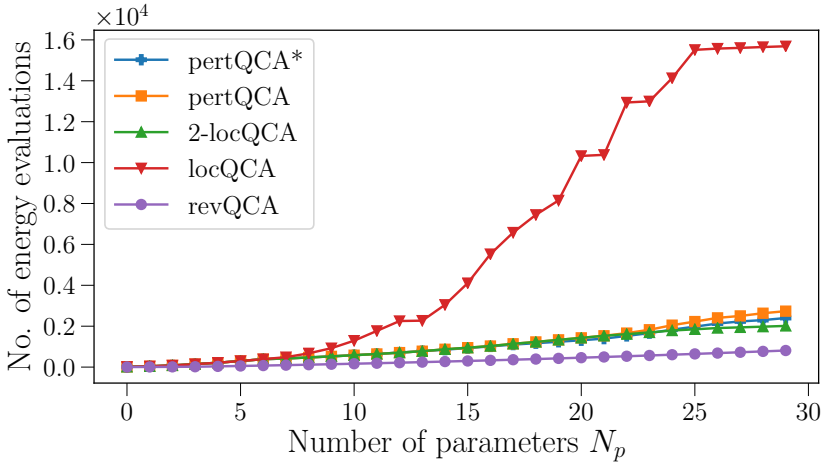

Figure 5.9: Plot of the optimization convergence speed (Eq. (5.6.3)) for different variational hierarchies in a weakly-coupled transverse-field Ising model ($J/h = 0.15$). Convergence is represented by a total number of energy function evaluations n_{ev} and plotted as a function of the number of parameters used. Note that the optimization of N_p ansatz parameters always uses the optimized value of $N_p - 1$ parameters for initialization (see Sec. 5.6.4). Because of this, in $n_{ev}(N_p)$ we always include $n_{ev}(N_p - 1)$ and the resulting plots are by definition monotonic.

5.E Convergence speed of classical optimization of QCA

In this appendix we show the convergence rate of our classical optimization of QCA in terms of the number of function evaluations for Fig. 5.6, Fig. 5.7 and Fig. 5.8 (Fig. 5.9, Fig. 5.10 and Fig. 5.11 respectively). We have not performed any metaparameter tuning for this optimization, which would likely improve these numbers significantly. The optimization here was performed in the absence of realistic conditions on quantum hardware (in particular in the absence of sampling noise); any further optimization of convergence times would need to take this into account in order to make a realistic comparison to other ansatzes.

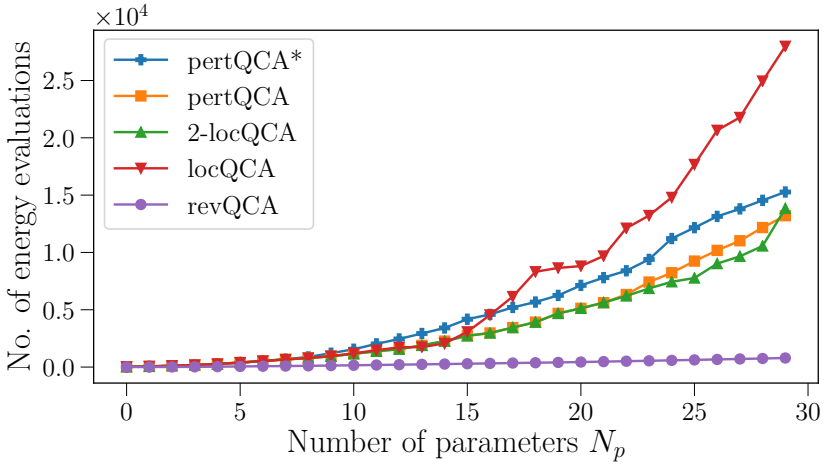

Figure 5.10: Plot of convergence speed similar to Fig. 5.9, but in the strongly-coupled regime instead ($J/h = 6$).

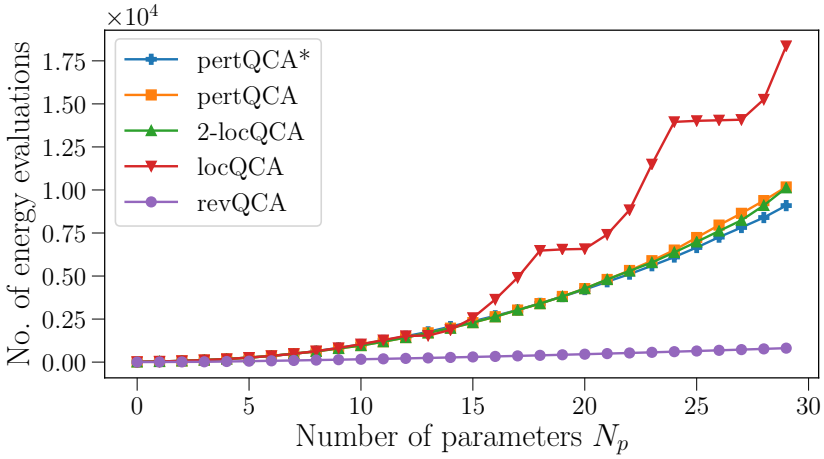

Figure 5.11: Plot of convergence speed similar to Fig. 5.9, but in the critical regime instead ($J/h = 1$).

6 Quantum digital cooling

6.1 Introduction

Ground state preparation is an essential algorithm in the quantum computing toolbox. Any polynomial-time quantum algorithm can be mapped to the problem of estimating the ground state energy of an artificial Hamiltonian given an approximation to its ground state [44], and without such additional input this problem is known to be QMA-hard for even 2-local Hamiltonians [43]. Digital quantum simulation of problems in materials science and chemistry, one of the 'killer apps' of a quantum computer, is most often concerned with properties of ground states of the simulated systems [12, 174], and many problems in optimization may be mapped to ground state finding problems [45, 162]. This has led to a wide range of schemes for digital ground state approximation, via adiabatic evolution [45], variational methods [47, 159, 162], phase estimation [175], and approximate imaginary time evolution [176–178]. However, these algorithms suffer from large computational costs or approximation errors, making designing better schemes an active area of interest.

In nature, ground states are achieved by coupling to a large cold reservoir, which takes energy from the system in keeping with the second law of thermodynamics. Simulating an entire bath would require an impractically large quantum register, however it has long been suggested that this may be mimicked by coupling to a single qubit which may be reset to its ground state with sufficient frequency [12]. This idea has been since studied in digital quantum computing for the initialization of quantum devices [62, 63], and in analogue settings for the preparation of physical [64] and artificial [65, 66] ground states. However, cooling an artificial system in the digital quantum setting provides a set of unique challenges — the system being studied may be completely different from the physical quantum hardware, and the digitized Hamiltonian may be only an approximation to the target of interest. Furthermore, the periodic non-unitary reset may break the unitary evolution in short time-scale chunks which do not conserve energy, implying that one may artificially reheat the system without clever protocol design. This is of critical importance in near-term

devices, where limited coherence times compete against the desire for slower cooling cycles.

In this chapter, we detail how one may prepare ground states of an artificial Hamiltonian on a digital quantum computer via quantum digital cooling (QDC). We first present an analytic study of the cooling of a two-level system, from which two different approaches may be outlined to de-excite to the ground state whilst preventing reheating. We investigate the behaviour of both methods in the digitized setting, and find they continue to be robust. The protocols deriving from these two principles are tested in the one-qubit black-box Hamiltonian setting, where the energy gap and matrix elements are unknown. We extend these protocols to N-qubit systems, and investigate their ability to cool small-scale simulations of the transverse-field Ising model numerically. Our LogSweep protocol, based on the weak-coupling approach, is demonstrated to converge to the ground state of all three phases of the transverse-field Ising model. It further shows a relative energy error constant in the system size at a fixed protocol length for the weakly-coupled and critical phases of this model, which corresponds to a requirement to simulate time evolution for $O(N^2)$ and $O(N^3)$ Trotter steps respectively. By contrast, the stong-coupling BangBang protocol shows the ability to prepare low-cost ground-state approximations of the same model in the paramagnetic and ferromagnetic regime, but seems to perform much worse close to the critical point, where the system spectrum shows a less-ordered structure. The small number of calls to the system evolution operator needed to realize this protocol makes it attractive for near term application.

6.2 Cooling a system with a single fridge qubit

In nature, gapped physical systems cool down to a state with high overlap to the ground state when interacting with a bath that is cold and large, under the condition of ergodicity. By *cold*, we mean that temperature T_B of the bath is small compared to the ground state gap Δ_S of the system to be cooled: $k_\mathrm{B}T_\mathrm{B} \ll \Delta_\mathrm{S}$ (with k_B Boltzmann's constant). By *large*, we mean that the bath has a sufficiently large Hilbert space that the above condition is unchanged by the addition of the energy from the system. By *ergodic*, we mean the system must not have symmetries that prevent excitations to be transferred from the system to the bath, or that reduce the effective size of the accessible bath Hilbert space. Given a system with Hamiltonian

H_S and eigenstates $H_S|E_j\rangle = E_j|E_j\rangle$, energy conservation implies that the bath must have states at energies $E_j - E_0$ to allow de-excitation of the eigenstates E_j. This is typically achieved by considering a bath with a continuous or near-continuous low-energy spectrum [Fig. 6.1(a)]. However, the bath need not cool all states immediately to the ground state. Instead, a bath typically absorbs single quanta of energy $\epsilon = E_i - E_f$ that correspond to local excitations of the system $|E_i\rangle \to |E_f\rangle$, at a rate given by Fermi's golden rule:

$$\frac{dP_{\mathrm{i}\to\mathrm{f}}}{dt} = \frac{2}{\hbar}\int_0^\infty d\epsilon\, |\langle E_f, \epsilon|H_C|E_i, 0\rangle|^2\, \rho_B(\epsilon) \times \lim_{t\to\infty} \frac{\sin[(E_i - E_f - \epsilon)\, t]}{E_i - E_f - \epsilon} \tag{6.2.1}$$

$$= \frac{2\pi}{\hbar}\, |\langle E_f, \epsilon|H_C|E_i, 0\rangle|^2\, \rho_B(E_i - E_f), \tag{6.2.2}$$

where H_C is the coupling between the system and the bath, and ρ_B is the density of states of the bath *. This requires the bath to be large enough to prevent reexcitation of these states as the system continues cooling. In other words, the bath must have a large Hilbert space compared to the one of the system. This ensures that, at equilibrium, most of the entropy is distributed in the bath.

To represent such a large bath with an ancillary register on a quantum device in order to cool a system register would be impractically costly. In this work, we approximate the presence of a much larger bath B with a single ancilla qubit F [Fig. 6.1(b)], with bath Hamiltonian $H_B \to H_F = \epsilon\, Z_F/2$. This can be considered a simplified form of a collisional model [179] that does not allow for non-Markovian effects (that would be in our case unwanted). The coupling between the bath and the system takes the form $H_C = \gamma X_F \otimes V_S/2$, where γ is the coupling strength, and V_S a *coupling term* that acts on the system alone. A key advantage of the digital approach is that we are free to choose V_S as desired to optimize the cooling protocols. The Hamiltonian of the entire system and bath then takes the form

$$H = H_S + H_F + H_C. \tag{6.2.3}$$

This has an immediate problem, as the bath can only absorb a single quantum of energy ϵ, but we may circumvent this by periodically resetting the ancilla qubit to $|0\rangle$. The non-unitary reset in effect extracts energy and entropy from the ancilla to a much larger external bath (the experimenter's

*In the rest of the chapter we assume $\hbar = 1$

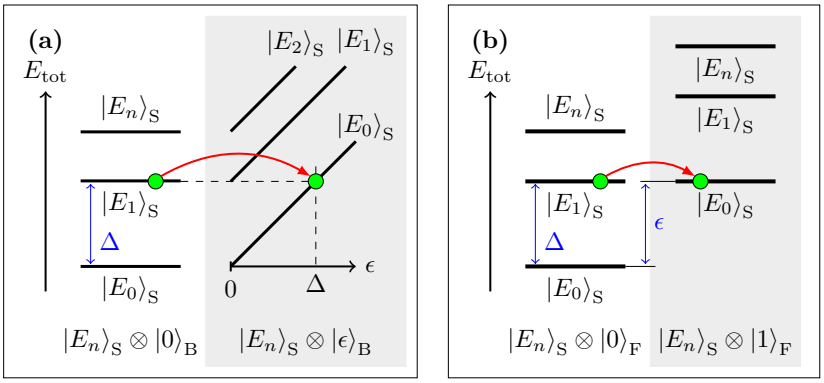

Figure 6.1: The de-excitation of the system transition $|E_1\rangle_S \to |E_0\rangle_S$ mediated by: **(a)** a continuous-spectrum natural bath B, where an excitation $|\epsilon\rangle_B$ at energy ϵ is produced, and **(b)** a single-qubit digital fridge F, which can be excited if $\epsilon = \Delta$.

environment). For this reason we call the ancilla qubit a 'fridge' qubit (hence F). The non-unitarity introduced in the process is necessary to dissipate entropy, allowing to prepare the ground state from an arbitrary starting state. As the time between resets is finite, the $t \to \infty$ limit in Eq. (6.2.1) is no longer justified and energy is no longer conserved. This is both a blessing and a curse — we need not precisely guess the energy gap $\Delta = E_i - E_f$ of the transition that we need to de-excite, but we run the risk of reheating the system at each cooling round. Minimizing re-heating while maximizing the range of targeted de-excitations is key to the successful design of QDC protocols. In a realistic experiment, qubit re-heating would be effectively increased by reset infidelity on the ancilla qubit, making the protocol less effective.

6.3 De-exciting a single transition: the 1+1 model

In order to design some basic protocols for QDC, we turn to a toy '1+1' model. We take a single-qubit system with Hamiltonian $H_S = \Delta\, Z_S/2$, and couple it to a single fridge qubit with coupling term $V_S = X_S$. Although this model is simple, it can for instance represent a pair of levels being

targetted for cooling in a much larger quantum system. We will make use of this interpretation when extending these cooling protocols in section 6.4.

6.3.1 Elementary approaches to digital cooling: strong and weak-coupling

Let us first assume Δ is known, in which case resonant cooling ($\epsilon = \Delta$) can be seen to be the most effective choice of ϵ. With this fixed, the transition probabilities after time t may be calculated exactly to be

$$P_{1\to 0} = \sin^2\left(\frac{\gamma}{2}t\right), \; P_{0\to 1} = \frac{\gamma^2 \sin^2(t\Omega)}{4\Omega^2}, \tag{6.3.1}$$

where $\Omega = \sqrt{\gamma^2/4 + \epsilon^2}$. We wish to maximise the cooling probability $P_{1\to 0}$ while minimizing the reheating probability $P_{0\to 1}$ by optimizing the remaining free parameters: the coupling strength γ and the cooling time t. To maximize the cooling rate $P_{1\to 0} = 1$, we must set

$$t = \pi\gamma^{-1}. \tag{6.3.2}$$

We assume this constraint throughout this chapter. This goes beyond the perturbative regime $\gamma t \ll 1$ in which Eq. (6.2.1) is formulated. However, we can take two very different approaches to minimize reheating, based on strong or weak coupling. The weak-coupling approach is based on the observation that the off-resonant transition $P_{0\to 1}$ is bounded by $\gamma^2/4\Omega^2$. As such, we may suppress reheating to an arbitrary level by choosing sufficiently small γ. The time-cost for Hamiltonian simulation of e^{iHt} scales at best linearly in t [180], so this implies one may obtain the ground state with failure probability p in time $O(p^{-1})$, regardless of the initial state of the qubit. The strong-coupling approach consists of tuning γ so that $\Omega t = \pi$, which is achieved when

$$\gamma = \frac{2}{\sqrt{3}}\epsilon. \tag{6.3.3}$$

This fixes the reheating exactly to 0, guaranteeing the qubit to be in the ground state in the shortest possible time, but at the cost of requiring fine-tuning.

Unlike in analog quantum simulation, digital devices cannot exactly implement the dynamics of the Hamiltonian in Eq. (6.2.3), and must approximate it digitally instead. A common approach to such digitization is that of the Suzuki-Trotter expansion [58, 181], which we now explore

for the two cooling paradigms. We apply the (second-order) expansion of the coupled system-bath evolution with Trotter number M,

$$e^{-i(H_S+H_F+H_C)\,t} \sim \left[e^{-iH_C\frac{t}{2M}}e^{-i(H_S+H_F)\frac{t}{M}}e^{-iH_C\frac{t}{2M}}\right]^M. \tag{6.3.4}$$

Note that, when we later approach larger systems, we will practically realize $e^{-iH_S t/M}$ as a single second-order Trotter step, effectively implementing a second-order Trotter simulation of the coupled Hamiltonian $H_S + H_F + H_C$ with trotter number M. If we restrict to the subspace containing the states involved in the cooling transition $|10\rangle_{SF} \to |01\rangle_{SF}$, at resonant cooling we have $H_S + H_F \propto \mathbb{1}$ (specifically, in this model $|01\rangle$ and $|10\rangle$ generate a zero-eigenvalue subspace of $H_S + H_F$). Thus, the Trotterized evolution behaves exactly like the continuous one with regards to the cooling transition. We study reheating probabilities as a function of t for different values of M in the weak-coupling regime. We observe (Fig. 6.2) that the digitized evolution approximates well the behavior of the continuum limit whenever $t\Omega/\pi \lesssim M$ (i.e. for the first M Rabi oscillations with pulse Ω). For longer times $t\Omega\pi \gtrsim M$, the second-order Trotter approximation fails, leading to reheating rates far larger than in the continuum limit. This allows us to define a practical choice of M to avoid reheating due to digitization — we require

$$M \gg \sqrt{1+\epsilon^2/\gamma^2}, \tag{6.3.5}$$

which sets the working point $t = \pi\gamma^{-1}$ before the $M/2$ Rabi oscillation. However, in the strong-coupling case $t\,\Omega/\pi = \sqrt{3}$, which implies that a single step is sufficient. Indeed, digitized cooling with probability 1 and no reheating can be realized by a *bang-bang* approach (inspired by similar approach in variational methods [182, 183]). This consists in defining the evolution as in Eq. (6.3.4) with $M = 1$, as long as the coupling strength is adjusted to $\gamma = 2\epsilon$. With this choice, the digitized evolution implements resonant Ramsey interference on the cooling transition $|10\rangle_{SF} \to |01\rangle_{SF}$ and anti-resonant Ramsey interference on the reheating transition $|00\rangle_{SF} \to |11\rangle_{SF}$.

A key difference between the two approaches to digital cooling is in their behavior off-resonance, i.e. when the energy gap is mistargetted or not precisely known. For the bang-bang approach, detuning reduces the cooling efficiency while symmetrically boosting reheating [Fig. 6.3(a)]. The wide resonance peak around zero detuning makes this approach ideal to quickly cool transitions which energy is known up to a small error. In the weak-coupling approach the resonance peak becomes sharper and the reheating gets more suppressed as the coupling is made smaller

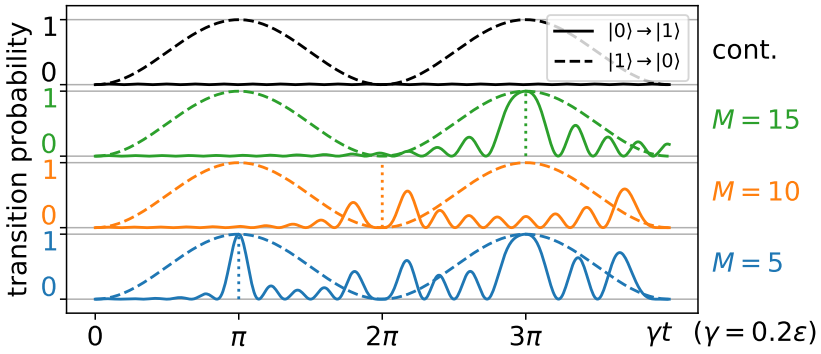

Figure 6.2: Effects of Trotterization on cooling and reheating probabilities as a function of the coupling time t, for different numbers of Trotter steps M per cooling cycle. Vertical dotted lines indicate the M-th reheating oscillation, at which point the Trotter approximation fails.

[Fig. 6.3(b)], approaching the energy conservation limit. Detuning makes cooling inefficient, but thanks to the low reheating probability this weak-coupling cooling can be iterated while changing ϵ to try to match the transition energy, without destroying the cooling effect.

6.3.2 Common symmetries and the coupling alternation method

In large systems of interest, we do not know the Hamiltonian's eigenstates, making it more challenging to couple between them. This is required for cooling, as can be seen by the direct dependence of the cooling rate $\frac{dP_{i\to f}}{dt}$ on the overlap $|\langle E_f, \epsilon|H_C|E_i, 0\rangle|^2$ (Eq.6.2.2). This overlap dependence will prohibit cooling if the system Hamiltonian H_S and the coupling operator V_S share a common symmetry S (i.e., $[S, H_S] = [S, V_S] = 0$). When this is the case, the Hamiltonian may be simultaneously diagonalized with H_S, and a state with some overlap to any eigenspace of S that does not contain the ground state cannot be cooled to the ground state by coupling with V_S. Note that this is a strictly stronger condition than just requiring $[H_S, V_S] \neq 0$. A simple solution is to alternate over a set of couplings $\{V_S^i\}$ as we cool. Then, any symmetry S of H need commute with *each* V_S^i in order to guarantee that a state starting with overlap in a high-energy eigenspace will remain there. Sets of coupling terms $\{V_S^i\}$ on N

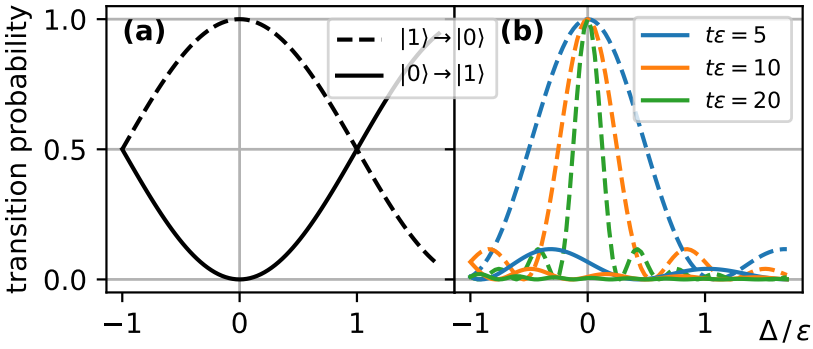

Figure 6.3: Effect of fridge-system detuning $\delta = \Delta - \epsilon$ on the cooling (dashed lines) and reheating (solid lines) probabilities for **(a)** the bang-bang cooling approach, and **(b)** the weak-coupling cooling approach, where colors indicate different couping strengths.

qubits need only be size $O(N)$ to break commutation with all non-trivial operators (for example, the set of all single-qubit Pauli operators), so although symmetries need to be taken into account, they will not destroy the feasibility of QDC protocols.

This issue may be demonstrated on the protoype $1+1$ qubit model by considering the system Hamiltonian $H_S = h\,\vec{n}\cdot\vec{\sigma}$, where $\vec{n}$ is a 3-dimensional unit vector (so H_S points in a random direction on the Bloch sphere), $2h$ is a fixed energy splitting, and σ is the vector of Pauli-matrices. For a fixed coupling operator V_S, there is a risk that $[H_S, V_S] \approx 0$. When this is the case, the off-diagonal elements of V_S in the system eigenbasis are zero, preventing cooling. We may prevent this by alternating between different coupling terms during the cooling protocol, such that no non-trivial Hamiltonian can commute with all such coupling terms. This may be achieved for the $1+1$ model by iterating over $V_S^i \in \{X_S, Y_S, Z_S\}$, or alternatively over $V_S^i \in \{X_S, Z_S\}$. The effectiveness of this scheme is studied in Fig. 6.4 for resonant coupling. We see the probability $P_{1\to0}$ of successful cooling of the weak coupling approach ($t\,\epsilon = 10$) increased to $\min(P_{1\to0}) = 97\%$ for all choices of $\vec{n}$ when iterating over $V_S^i = X_S, Y_S, Z_S$, and above 95% when iterating $V_S^i = X_S, Z_S, X_S$, compared to the possibility for complete cooling failure [$\min(P_{1\to0}) = 0$] when V_S^i is held constant. Similar results are seen for off-resonant coupling.

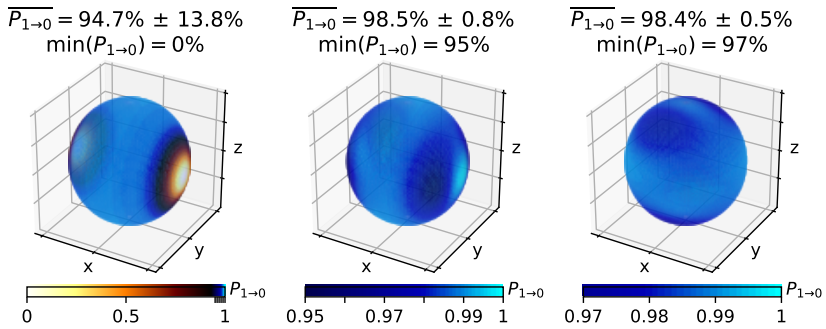

Figure 6.4: Probabilities $P_{1\to0}$ of transitioning from $|1\rangle$ to $|0\rangle$ after three iterations of the weak-coupling ($t\,\epsilon = 10$) cooling procedure, with coupling potentials $V_{\mathrm{S}}^{i} = X, X, X$ (left), $V_{\mathrm{S}}^{i} = X, Y, X$ (center), $V_{\mathrm{S}}^{i} = X, Y, Z$ (right), on a system qubit with Hamiltonian $H_{\mathrm{S}} = h\,\vec{n}\cdot\vec{\sigma}$ and known energy splitting h. The orientation of the unit vector $\vec{n}$ is represented on spherical surfaces. The average, standard deviation and minimum of $P_{1\to0}$ are shown above each panel.

6.4 Scalable QDC protocols

We now look to use the insight obtained for cooling in the 1+1 toy model to develop QDC schemes for larger systems. The sub-additivity of entropy places a rough lower bound on the number of cooling steps required to cool an N-qubit system. This limits the entropy ΔS_{S} that the system can transfer to the fridge qubit before the non-unitary reset to $\Delta S_{\mathrm{S}} \geq -\Delta S_B \geq -1$ bit. If we demand the ability to cool an arbitrary state, a QDC protocol must also be able to cool the maximally-mixed state, which has entropy $S_{\mathrm{S}} = N$. We then require N repetitions of an optimal coupling-and-reset step to reach the pure ground state (which has entropy $S_{\mathrm{S}} = 0$). This can be obtained in the simple example of cooling N non-interacting qubits with known energies, by simply repeating the protocols of the $1+1$ model. However, this cannot be generalised to arbitrary strongly-correlated systems, as cooling is complicated by irregular and unknown energy gaps and coupling terms between eigenstates. This is to be expected, as preparing ground states of arbitrary Hamiltonians is a known QMA-hard problem [43]. However, as cooling is a physically-motivated process, we hope QDC to be able to achieve polynomial scalings for systems of physical interest, i.e. models of systems that are found in low-temperature equilibrium states in nature.

In the rest of this chapter, we introduce two scalable QDC protocols for

N-qubit systems: the strong-coupling-based BangBang protocol and the weak-coupling-based LogSweep protocol. These extend and generalize the two approaches we established for the 1+1 toy model of section 6.3.1. Each protocol iterates over a sequence of *cooling steps*, each of which consists of coupling the fridge qubit to part of the system for a short time evolution, and then resetting the fridge qubit to its ground state. The protocols differ in the choice of coupling strengths γ_i, coupling terms V_S^i and fridge energies ϵ_i at each i-th cooling step. [The coupling time for each cooling step is fixed by Eq. (6.3.2)].

6.4.1 The BangBang protocol

We now develop a protocol to extend the strong-coupling approach (Eq. 6.3.3) to a larger system. This motiviation is in line with recently proposed bang-bang approaches to adiabatic state preparation [182, 183], which are known to outperform initial theoretical expectations stemming from a naive Trotter error estimate. We are thus optimistic that this 'BangBang' protocol may provide a low-cost, near-term method for QDC. However, such a protocol needs to associate each fridge-system coupling with a single fridge energy, that should match the transitions that this coupling triggers. An appropriate choice of fridge energy can be estimated via a perturbation theory approximation. To derive this approximation, we note that the rate of a transition between eigenstates $|E_i\rangle \to |E_j\rangle$ depends on the matrix element of the coupling V_S:

$$V_{(ij)} := \langle E_i|V_S|E_j\rangle = \frac{\langle E_i|[H_\mathrm{S}, V_S]|E_j\rangle}{E_i - E_j}. \tag{6.4.1}$$

If V_S is local and bounded, $[H_\mathrm{S}, V_S]$ is as well, so the matrix element $V_{(ij)}$ will be bounded proportionally to $(E_i - E_j)^{-1}$. The matrix element is additionally bounded by $\|V\|$; this second bound becomes dominant when $E_i - E_j/\|V\|$ falls below the maximum off-diagonal element of $[H, V]$ in any basis, which we define with the notation $\|[H, V]\|_\perp$:

$$\|O\|_\perp = \max_{\langle\phi|\psi\rangle=0} |\langle\phi|O|\psi\rangle| = \max_{|\Phi\rangle,|\Psi\rangle} \frac{\langle\Phi|O|\Phi\rangle - \langle\Psi|O|\Psi\rangle}{2}, \tag{6.4.2}$$

where O is Hermitian and the maxima are taken over all possible states $|\psi\rangle, |\phi\rangle$ and $|\Psi\rangle, |\Phi\rangle$. A simple proof is given in Appendix 6.A. We use this energy scale to set the fridge energy:

$$\epsilon_i = \|[V_\mathrm{S}^i, H_\mathrm{S}]\|_\perp \tag{6.4.3}$$

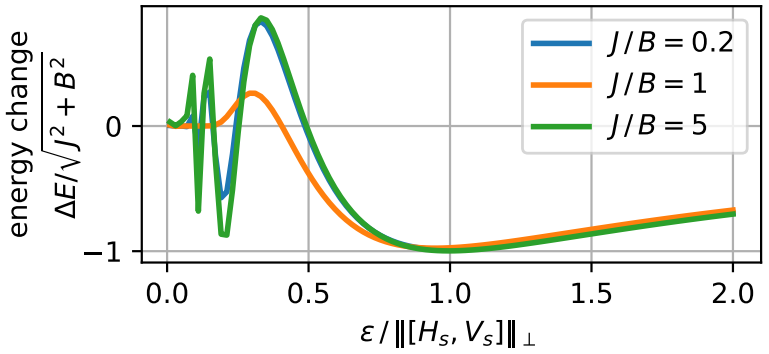

Figure 6.5: Change in energy expectation value for the application of a single cooling step to the maximally mixed state of a $N = 8$ qubit transverse field Ising chain Eq. (6.4.5), depending on the fridge energy ϵ. The coupling potential is $V_S = Y_3$, the Pauli Y on the third qubit. The relation $B^2 + J^2 = 1$ fixes the energy scale.

for any coupling potential V_S^i. This way, the maximum-energy transitions accessible by V_S are on resonance, while smaller energy ones (which are less important for cooling) still have a higher probability of cooling than of reheating [see Fig. 6.3(a)]. This defines the BangBang protocol: we iterate over coupling to each qubit, with ϵ_i fixed by Eq. (6.4.3). As this protocol does not attempt to suppress reheating, we choose a single coupling $V_S = Y_n$ for the n-th qubit, instead of iterating over $V_S = X_n, Y_n, Z_n$ (as was suggested in Sec.6.3.2). In general, the best choice of V_S will depend on the system that we want to cool, and the couplings should be picked to enable as many transitions as possible. We repeat the coupling to each qubit R times, resulting in a total of RN cooling steps. Each cooling step contains two first-order Trotter steps simulating $e^{-iH_C t/2}$ (of depth $d^{(1)}_{H_C}$), a single second-order Trotter step for $e^{-iH_S t}$ (of depth $d^{(2)}_{H_S}$), and a reset gate, resulting in a total circuit depth

$$d_{\text{BangBang}} = RN(2d^{(1)}_{H_C} + d^{(2)}_{H_S} + 1). \tag{6.4.4}$$

To test the BangBang protocol, we study the cooling of a N-qubit transverse-field Ising chain

$$H_S = \sum_{i=0}^{N} BX_i + \sum_{i=0}^{N-1} JZ_iZ_{i+1}, \tag{6.4.5}$$

where B represents the transverse magnetic field Zeeman splitting and J is the Ising coupling strength. The relative coupling strength J/B dictates whether the system is in the paramagnetic ($J/B \ll 1$), ferromagnetic ($J/B \gg 1$), or critical ($J/B \sim 1$) phases. This ability to simply tune between three phases of matter with significantly different physical properties make the TFIM a good benchmark model to investigate the ability of different QDC schemes in various scenarios.

We first demonstrate that our choice for the fridge energy Eq. (6.4.3) is appropriate. In Fig. 6.5, we plot the effect of a single cooling step on the maximally-mixed state. We observe that cooling is maximized for fridge energies around the point defined by Eq. (6.4.3), for all phases of the TFIM. We find this behaviour to hold for all other (local) choices of coupling potential V_S used in this chapter, as predicted.

We next turn to the ability of the BangBang protocol to prepare an approximation ρ of the ground state, starting from a maximally-mixed (i.e. infinite temperature) initial state. We benchmark by the final state with two figures of merit: the ground state fidelity

$$F = \mathrm{Tr}\big[\, |E_0\rangle\langle E_0|\, \rho \,\big], \tag{6.4.6}$$

and the energy relative to the ground state energy $Tr[H_S\rho]/|E_{GS}|$. This last property is local in local system, and represents an energy density in TFIM. To verify convergence, we compare cooling results to a *reheating limit*, obtained by running the protocol with the ground state as initial state. We observe that all TFIM phases converge after $R \approx N$ repetitions (with the weakly-coupled phase system converging already at the first repetition). In Fig. 6.6 we plot the energy density of the cooled state, as well as the reheating limit, as a function of the number of sites in the system. This shows that convergence is indeed achieved for $R = N$ independently on the system phase and size, and that the final energy density stays approximately constant, without showing any other trend. The BangBang protocol achieves a final energy density close to 90% and 95% of $\|H_S\|_\perp$ for the ferromagnetic and paramagnetic regime respectively, while performing significantly worse in the critical regime. This is to be expected, as in this regime the spectrum is no longer banded, and excitation energies are not as uniform as in the paramagnetic or ferromagnetic regimes. Following Eq. (6.4.4), the protocol's circuit depth is $7NR$ for a gate-set containing single- and double-qubit rotations (and the reset gate). Given the low cost of the protocol, we suggest that this is of particular interest for near-term experiments, and may be further refined by other cooling protocols, or methods such as quantum phase estimation, in the long term.

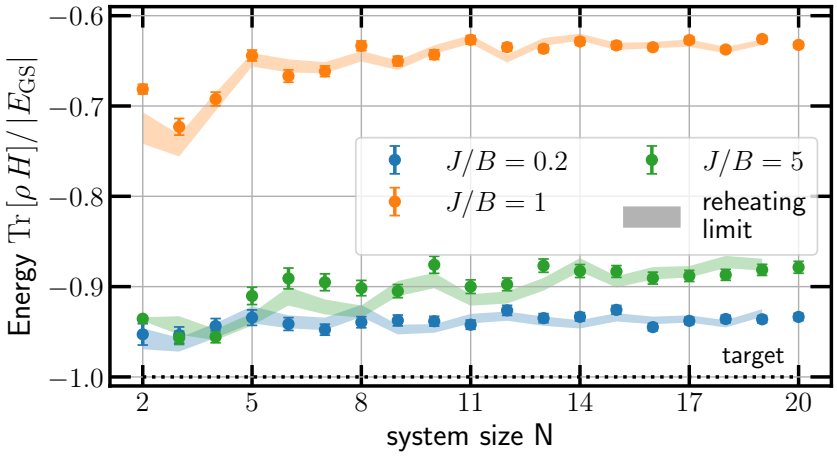

Figure 6.6: Performance of the BangBang protocol as a function of the system size N for the three different phases of the transverse-field Ising model (detailed in legend). The coupling potentials are $V_S^i = Y_i$. Dots correspond to result when the protocol is applied to the maximally-mixed state, shaded regions corresponds to result when protocol is applied to the true ground state (which gives a bound on protocol re-heating). Data generated by Trotterized wave-function simulations of the protocol, and random sampling of the initial mixed state and of nonunitary operations. All points are run with 200 samples, and average results are plotted with the sampling error.

6.4.2 The LogSweep protocol

Refrigeration at weak-coupling suppresses reheating, but only allows for the cooling of transitions within a narrow energy band [as shown in Fig. 6.3(a)]. We may take advantage of this in a larger system, where a wide range of energy gaps are present, by sweeping the fridge energy ϵ_k from high to low as we iterate over cooling steps. (As low-energy transitions are more susceptible to re-heating than high-energy transitions, this will in general be more efficient than sweeping from low to high.)

To construct a full protocol, we further need to fix the set of fridge energies ϵ_k and linewidths $\delta_k = t_k^{-1} = \pi\,\gamma_k$ that we plan to use for each cooling step. We will be guided by two principles. First, the target band of fridge energies $(E_{\min}, E_{\max})$ should be tightly covered by the cooling lines $\epsilon_k \pm \delta_k$. Second, the reheating should be suppressed to a certain degree throughout the protocol. As by Eq. (6.3.1) the reheating suppression

depends on γ_k/ϵ_k, we fix this value to a small constant throughout the protocol (i.e. we choose $\gamma_k \propto \epsilon_k$). Thus we define the LogSweep protocol, where the fridge energy ϵ_k is sweeped over $(E_{\min}, E_{\max})$ in a logarithmic gradation. Specifically, given the *gradation number* K, we index each cooling step $k = 1, \ldots, K$, and we define

$$\epsilon_k = E_{\min}^{\frac{k-1}{K-1}} E_{\max}^{1-\frac{k-1}{K-1}}, \tag{6.4.7}$$

and choose δ_k to fix $\epsilon_{k+1}+\delta_{k+1}/\zeta = \epsilon_k - \delta_k/\zeta$, with ζ a constant (potentially dependent on K). In App. 6.B, we prove that such a scheme will cool a single transition in the range $(E_{\min}, E_{\max})$ with probability 1 as $K \to \infty$, and in App. 6.C we demonstrate that the logarithmic gradation is optimal for such a scheme for a choice of $\zeta(K) \sim \log(K)$. To make sure all system excitations have a chance of being dissipated, we further iterate the couplings V_S over a set of local couplings $\{V_S^i\}$ throughout the system: for the considered spin systems we choose $\{V_S^i\} \equiv \{X_n, Y_n, Z_n\}$ for each qubit n (see Sec.6.3.2), for a total of $3NK$ cooling steps. The number of Trotter steps M_k for each cooling step k is chosen to prevent re-heating. This follows Eq. 6.3.5, but as transition energies between system eigenstates may be as large as the Hamiltonian spread $2\|H_S\|_\perp$, we set

$$M_k = 2\sqrt{1 + \frac{2\|H_S\|_\perp^2}{\gamma_k^2}}. \tag{6.4.8}$$

The choice of the fridge energy range $[E_{\min}, E_{\max}]$ will generally depend on heuristics on the system. $E_{\max}$ should be greater or equal than the largest energy of the transitions that we are able to de-excite with the chosen couplings V_S (for local Hamiltonians we can estimate this with the techniques described in 6.4.1). For ground state cooling, $E_{\min}$ should be close to the system ground state gap Δ_{GS}, as no transition with an energy lower than Δ_{GS} can lead from an excited state to the ground state.

We first test the LogSweep protocol as applied to the 1+1 model defined in Sec. 6.3.1, with the system gap Δ now taking an unknown value between $E_{\min}$ and $E_{\max}$ (Fig. 6.7). At each step $k = 1, \ldots, K$ we want to maximise cooling of transitions $\Delta \sim \epsilon_k$, while minimizing reheating of previously-cooled transitions $\Delta \sim \epsilon_{k'}$, $k' < k$. As demonstrated by the black dashed curve in Fig. 6.7, when $E_{\max}/E_{\min} = 5$ this can be achieved well with only $K \approx E_{\max}/E_{\min}$ steps. Note that, to maintain a constant relative linewidth (and thus constant maximum reheating per step), we should scale $K \sim E_{\max}/E_{\min}$. This implies $K \to \infty$ as $E_{\min} \to 0$, in line with the third law of thermodynamics.

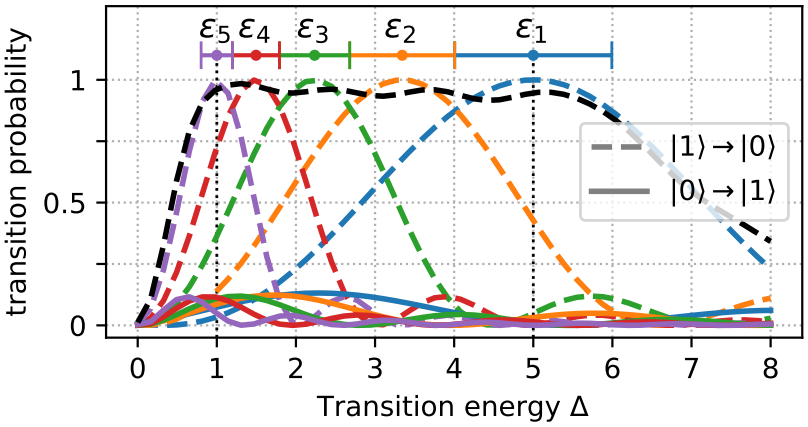

Figure 6.7: Choices of energies ϵ_k and linewidths δ_k (bars at the top of the graph showing $\epsilon_k \pm \delta_k$) for a $K = 5$ LogSweep protocol applied to the model introduced in section 6.3.1 with an unknown $\Delta \in (E_{\min} = 1, E_{\max} = 5)$. Colored lines show cooling (dashed) and reheating (solid lines) probabilities for each j-th step alone, the dashed black line shows the cooling probability after sequential application of the 5 steps.

In a larger system, the situation is more complex than in the model above. Instead of a single transition from the excited state $|E_1\rangle \to |E_0\rangle$ which occurs with unit probability when $\epsilon = \Delta = E_1 - E_0$, our system may transition to one of many eigenstates $|E_j\rangle$, to each with a transition probability $A_{i,j}$ (assuming a initial state $|E_i\rangle$). As there are many possible target states, the maximum transition probability might be very small ($\max_j A_{i,j} \ll 1$). If we restrict to a single transition $|E_i\rangle \to |E_j\rangle$ with the above reduced cooling rate, one may show that the LogSweep protocol still cools that transition with unit probability as $K \to \infty$, albeit at a rate that scales exponentially in $A_{i,j}$. Luckily, we do not need to ensure any specific transition occurs, instead we may cool sequentially

$$|E_i\rangle \to |E_{j_0}\rangle \to |E_{j_1}\rangle \to \ldots \to |E_0\rangle, \tag{6.4.9}$$

with a growing number of possible cooling paths as the system grows and the transition probabilities spread over more eigenstates. A good choice of the fridge energy interval $[E_{\min}, E_{\max}]$ and of the coupling potentials $\{V_S^i\}$ allows all eigenstates to be connected to the ground state by sequences of transitions $|E_{j_l}\rangle \to |E_{j_{l+1}}\rangle$ that have unit probability of being de-

excited for $K \to \infty$. However, a single transition probability approaches 1 only over the entire LogSweep protocol. In particular, if the transition $|E_{j_l}\rangle \to |E_{j_{l+1}}\rangle$ during step k^* of the protocol corresponds to an energy loss $E_{j_l} - E_{j_{l+1}} \gg \epsilon_{k^*}$, this transition will be off-resonance for the entire remaining duration of the protocol (as $\epsilon_k < \epsilon_{k^*}$ for $k > k^*$), making it unlikely to occur. This can cause convergence issues especially when cooling systems with banded spectra. For such systems , as we set $E_{\min} \approx \Delta_{\mathrm{GS}}$ as detailed above, there may be a point k^* in the protocol after which ϵ_k will become smaller than the average interband gap, but never as small as the spread of a single band. After this point, states at the bottom of a band might transition to states in the lower band, but states at the top of each band never have any resonant transitions to lower energy states, thus becoming absorbing states. This effect is clearly shown in Fig. 6.8, representing the LogSweep-cooled states of the transverse-field Ising model in different regimes. We start with the maximally-mixed state, and plot the resultant distribution over the eigenstate energies. In the banded regimes (side panels), we observe that the distribution of energies in any given band is tilted towards the higher-energy states in that band (i.e. the aforementioned absorbing states), by some orders of magnitude. This dead-ends ultimately hinder sequential cooling, and prevent the LogSweep cooling from converging to the same state independently on the initial state. The effect worsens as K is increased, as transition linewidths δ_k become smaller making off-resonant transitions less and less probable. This issue can be fixed in practice by using an initial state with fewer high-energy excitation (e.g. a classical approximation of a low-energy state). We solve the issue in principle, by constructing an *iterative LogSweep* protocols, where the LogSweep cooling is repeated with growing K. The early, lower-cost iterations cool the highest energy excitations, while the larger K iterations grant vanishing reheating, and probabilities approaching unity for the cooling transitions allowed by symmetries. Thus, adding iterations with larger and larger K, will make the whole protocol converge to the system ground state (unless symmetries forbid all paths from some states to the ground state). Note that this adjustment is not required for systems with a continuous spectra (i.e. critical systems), as in such a system there will be on-resonance transitions for any state with an energy $E_{\min}$ or more above the ground state.

We now investigate the performance of the (iterative) LogSweep protocol on different phases of the transverse-field Ising model. In Fig. 6.9, we plot the ground state infidelity of the prepared state ρ [$1 - F$ with F as in Eq. (6.4.6)], as a function of K. The protocol consists in $K - 1$ sweeps of a LogSweep QDC protocol, each sweep having gradation number

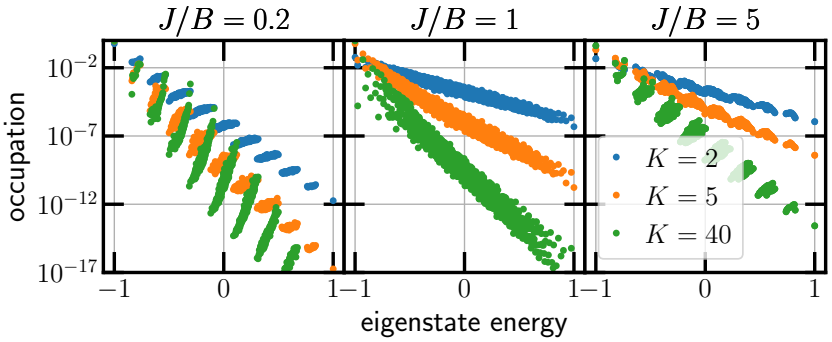

Figure 6.8: Effect of banding on single LogSweep iterations. A maximally mixed state in the three different phases of the 7-qubit TFIM spin chain is evolved by the LogSweep protocol for three different values of K. We plot the distribution of the result here over the system's eigenstates (indexed by energy), at three different values of K. We see that while the critical system demonstrates an approximate thermal or exponential distribution, the weak and strongly-coupled systems demonstrate an inversion in the population of the system within each band, which increases with K. Data generated by continuous-evolution density-matrix simulation.

$g_l = 2, \ldots, K$. The Hamiltonian simulation is performed by second-order Trotter approximation. We investigate the protocol effect on two initial states ρ_0: the maximally-mixed state $\rho_0 = \mathbb{1}/2^n$ to check for cooling capabilities (dots), and the ground state $\rho_0 = |E_0\rangle\langle E_0|$ (crosses) to show the lower bound originated by reheating. We observe polynomial convergence to the ground state in all three phases of the model, attaining an infidelity of $\varepsilon = 1 - F$ in approximately $K \sim O(\varepsilon^{-1/\beta})$ energy gradation steps for $\beta \approx 0.4$ - 0.8. Additionally, we verify that the protocol converges to the reheating limit for the critical and strongly-coupled regimes. In the weakly-coupled regime instead, although the cooling is far more efficient because of the local nature of the system excitations, the reheating bound is not saturated. We attribute this to the very small linewidths $\{\delta_k\}$, consequence of the well-defined transition energies, together with the strong banding of the system spectrum.

The number of Trotter steps for a single iteration of the LogSweep protocol with gradation number g_l on a system of N spins with Hamiltonian H_S scales as $O\big(\|H_S\|_\perp \Delta_{GS}^{-1} N g_l^2 \log(g_l)^{-1}\big)$. Thus, the iterative implementation required to deal with the banded cases needs a total number of Trotter

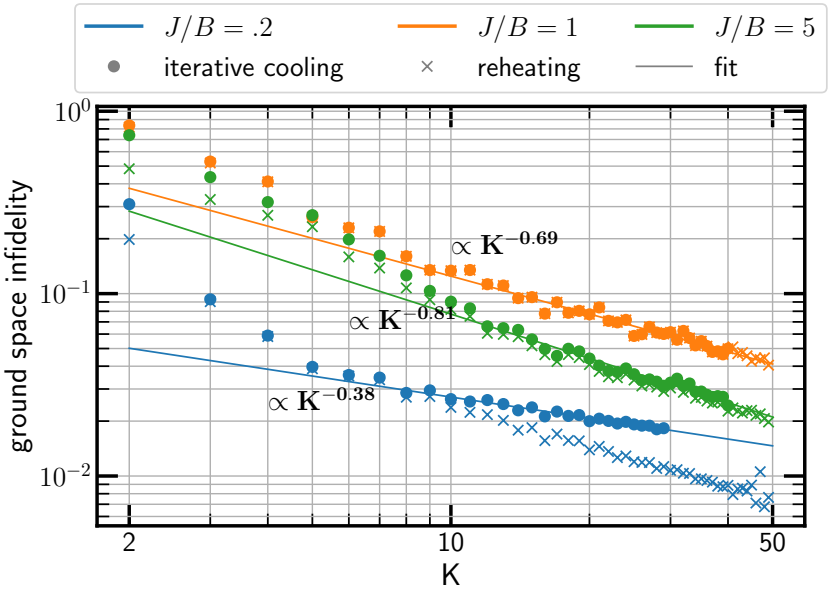

Figure 6.9: Convergence of the LogSweep protocol to the ground state as a function of the gradation number K, starting from the maximally mixed state (dots) and the ground state (crosses), for three phases of the transverse-field Ising model (detailed in legend). Data was generated by deterministic density-matrix simulations of the iterative LogSweep protocol, with second-order Trotter Hamiltonian simulation.

steps

$$M_{\text{tot}} \sim O\big(\|H_S\|_\perp \Delta_{\text{GS}}^{-1} N K^3 \log(K)^{-1}\big) \tag{6.4.10}$$

The gate complexity required to attain an error (infidelity) ε for the models studied scales thus as $O(\varepsilon^{-3})$ - $O(\varepsilon^{-8})$.

We next investigate the scaling of the LogSweep protocol as a function of the system size. In Fig. 6.10 we plot the relative error in the ground state energy as a function of the system size for a single (not iterated) LogSweep with gradation number $K = 5$. We see a constant error in the ground state energy as a function of the system size for the weakly-coupled and critical systems. Thus, here we expect no need to scale K with N for the protocol to be accurate. Let us also note that the gap in these two cases shrinks as $\Delta_{\text{GS}}/\|H\| \sim N^{-1}$ and $\Delta_{\text{GS}}/\|H\| \sim N^{-2}$ respectively. Using the above arguments and the estimate (6.4.10), one can find how the circuit

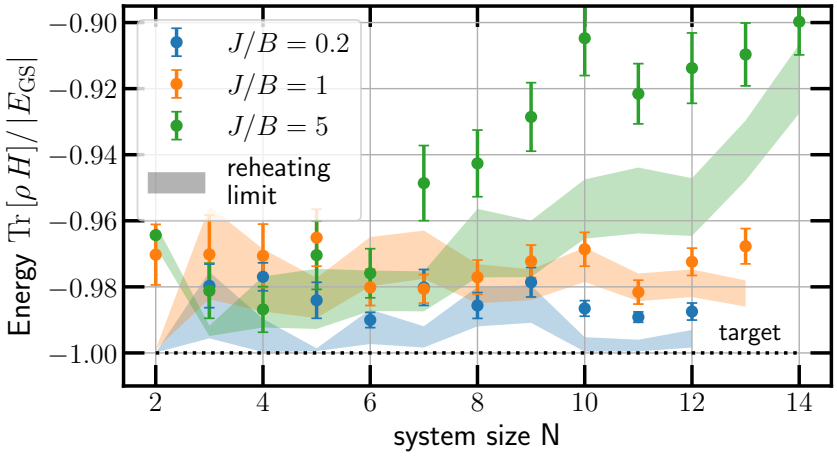

Figure 6.10: Performance of the LogSweep protocol as a function of the system size for the three different phases of the transverse-field Ising model (detailed in legend), with fixed $K = 5$. Dots correspond to result when protocol is applied to the maximally-mixed state, shaded region corresponds to result when protocol is applied to the true ground state (which gives a bound on protocol re-heating). Data generated by Trotterized wave-function simulations of the protocol, and random sampling of the initial mixed state and of nonunitary operations. All points are run with 100 samples, and average results are plotted with the sampling error.

length (in terms of time evolution steps), required to obtain a constant energy error, scales with N. We obtain $O(N^2)$ for the weakly-coupled and $O(N^3)$ for the critical case. From this analysis, we expect that the QDC protocol may be asymptotically competitive with methods such as adiabatic state preparation, whose runtime naively scales as $O(1/\Delta_{\mathrm{GS}}^2)$ [45, 174]. In the strongly-correlated phase, we do not see such positive results; the energy error increases with the system size, though the relative error remains beneath 10% for up to 14 spins. This may be explained by the relative growth of the extension of excitations within the strongly correlated phase, while cooling is performed with local couplings. Due to the error in the simulation, we are unable to reliably extract an estimate of the computational cost in the same way as for the critical and weakly-coupled systems. Future work may explore whether this error may be improved on by adjusting the form of the coupling terms $\{V_{\mathrm{S}}^i\}$ based on heuristics on the considered system.

6.5 Conclusion

In this chapter, we investigated how cooling can be simulated on a digital quantum computer, and demonstrated that this can be exploited for the design of scalable algorithms for preparing ground states of N-qubit systems. We identified how one can meet many of the fundamental challenges that the digital approach to cooling raises and use the leverage offered exclusively by digital quantum hardware, namely the freedom of choice in the coupling strength and fridge energy. We laid out a general approach of simulating a cold bath with a single ancilla qubit, which is iteratively coupled to various locations in the system and reset periodically to extract entropy and energy. We studied how to digitize the system-fridge coupling simulation without causing additional reheating, and how to avoid symmetries which produce non-ergodic behavior that hinders cooling. By tuning coupling parameters beyond the perturbative regime described by Fermi's golden rule, efficient cooling of targeted transitions can be realized. Following these principles we proposed two protocols for preparing approximate ground states of N-qubit systems — the BangBang protocol and the LogSweep protocol. We studied numerically how these protocols perform on the three phases of the 1D transverse-field Ising model. We found that the BangBang protocol quickly cools the system near to the ground state in the paramagnetic and in the ferromagnetic regime, but has difficulty in the critical regime. The LogSweep protocol is observed to cool all three phases to the ground state at a polynomial cost in the overlap error. In the weakly-coupled and critical phases, the LogSweep protocol further demonstrates a constant energy error as a function of the system size (for fixed gradation number), making it a competitive state preparation method.

The introduction of quantum digital cooling opens future research directions related to the characterization of proposed protocols, their optimization, and their extension beyond ground state preparation. A study of the effect of noise on currently proposed QDC protocols, and the optimization of such protocols for noise resilience, are in order to establish their applicability on near-term devices. Applying QDC to more complex physical systems, in areas such as quantum spin liquids, many-body localization and quantum chemistry, would bring new challenges to the protocol construction. A thorough study of the role in the cooling process played by the symmetries and locality of coupling could lead to the design of more optimized protocols. Furthermore, various extensions to the QDC protocols proposed in this chapter can be suggested. In a parallelized version of QDC, the use of multiple fridge qubits coupled to various locations in

the system might allow to trade space complexity for time complexity. A variationally-optimized QDC protocol might be devised, that can efficiently prepare a state in the ground state manifold of some Hamiltonian starting from an arbitrary initial state — differently from the variational quantum eigensolver [159] which requires the preparation of a fiducial state at every iteration. The principles of QDC might inspire a new class of efficient non-unitary quantum algorithms, where non-unitary operations are mediated by a single ancillary qubit, with possible application e.g. in the simulation of open quantum system dynamics.

One application of particular future interest for QDC protocols is in the preparation of Gibbs thermal states, which are useful e.g. for semi-definite programming [184]. This seems especially promising given the near-thermal distribution in Fig. 6.8 of the critical system under the evolution of the LogSweep scheme. However, it is as of yet unclear how to overcome the finite width of the distribution, and how well these protocols behave in the banded case (or for more general systems). Adjustment of the LogSweep protocol to produce robust thermal state preparation schemes is an obvious target for future research.

6.A Proof of Eq. (6.4.2)

To prove Eq. (6.4.2) we first show that

$$|\langle\phi|O|\psi\rangle| \leq \max_{|\Phi\rangle,|\Psi\rangle} \frac{\langle\Phi|O|\Phi\rangle - \langle\Psi|O|\Psi\rangle}{2}, \tag{6.A.1}$$

for all $|\psi\rangle, |\phi\rangle : \langle\psi|\phi\rangle = 0$. We can assume without loss of generality $\langle\phi|O|\psi\rangle$ is real and nonnegative (if it's not, we can multiply one state by an irrelevant global phase), and drop the absolute value. As $\langle\psi|\phi\rangle = 0$ we can define the states $|\pm\rangle = \frac{|\phi\rangle \pm |\psi\rangle}{\sqrt{2}}$ we can then write

$$\begin{aligned}\langle\phi|O|\psi\rangle =& \frac{1}{2}(\langle\phi|O|\psi\rangle + \langle\psi|O|\phi\rangle) \\ =& \frac{\langle+|O|+\rangle - \langle-|O|-\rangle}{2}\end{aligned}$$

immediately proving Eq. (6.A.1). The opposite inequality is proven by noticing that the $|\Psi\rangle$ and $|\Phi\rangle$ that maximize the right of Eq. (6.4.2) have to be eigenvalues (by the variational principle). With these, we can redefine the states $|\pm\rangle = \frac{|\Phi\rangle \pm |\Psi\rangle}{\sqrt{2}}$ which are also granted to be orthogonal, thus

$$\begin{aligned}\frac{\langle\Phi|O|\Phi\rangle - \langle\Psi|O|\Psi\rangle}{2} =& \Re[\langle+|O|-\rangle] \\ \leq& |\langle+|O|-\rangle| \\ \leq& \max_{\langle\phi|\psi\rangle=0} |\langle\phi|O|\psi\rangle|\end{aligned}$$

which combined with Eq. (6.A.1) proves Eq. (6.4.2).

6.B Asymptotic reheating and cooling probabilities for QDC protocols

Let us consider a two-state subsystem of a larger Hilbert space with a gap energy E, evolving under a QDC protocol on the kth step via a coupling term that does not mix the $\{|01\rangle, |10\rangle\}$ and $\{|00\rangle, |11\rangle\}$ subspaces (where the second index denotes the fridge). Under this assumption, the evolution of the system within this space is a Markov process. Following the main text, let the fridge energy on the kth step be ϵ_k, the coupling strength be γ_k, and the time evolved for in the cooling protocol t_k. Additionally, let

the spacing of the fridge energies to be

$$(\epsilon_k - \epsilon_{k+1}) = \zeta(\delta_k + \delta_{k+1}) = \frac{\alpha}{2}(\gamma_k + \gamma_{k_+1}),$$

for some K-dependent $\alpha = \alpha(K) = \frac{2}{\pi\zeta(K)}$. We may calculate the transition matrix for the Markov process, $p^{(k)}(E)$ (defined by $p^{(k)}_{a,b}(E) = P(|a\rangle \to |b\rangle)$ in a single cooling step) as

$$p^{(k)}(E) = \begin{pmatrix} 1 - \sin^2(\Omega_k t_k/2)\frac{\gamma_k^2}{\Omega_k^2} & \sin^2(\omega_k t_k/2)\frac{\gamma_k^2}{\omega_k^2} \\ \sin^2(\Omega_k t_k/2)\frac{\gamma_k^2}{\Omega_k^2} & 1 - \sin^2(\omega_k t_k/2)\frac{\gamma_k^2}{\omega_k^2} \end{pmatrix}, \tag{6.B.1}$$

where

$$\omega_k = \sqrt{(E - \epsilon_k)^2 + \gamma_k^2} \tag{6.B.2}$$

$$\Omega_k = \sqrt{(E + \epsilon_k)^2 + \gamma_k^2}. \tag{6.B.3}$$

Assuming no additional cooling or heating to the rest of the system during the protocol, the transition matrix for the $k_0 \to k_1$ block takes the form

$$P_{k_0,k_1}(E) = \prod_{k=k_0}^{k_1} p^{(k)}(E), \tag{6.B.4}$$

and the transition matrix for the entire process may be written $P(E) = P_{1,K}(E)$.

Exact analytic evaluation of this expression in the large K limit is quite difficult. Instead, we aim for a conservative estimate, bounding the final cooling probability $p_c = [P(E)]_{01}$ from below. For this, given the energy E, we first lower bound the 'initial' cooling around the resonant step k_c, i.e. such k_c that $|\epsilon_{k_c} - E|$ is minimal. Then we give an upper bound on reheating during the following protocol steps $k = k_c, ..K$. Given the estimated cooling probability $p_c^{(k_c)}$ and reheating probability $p_{rh}^{(k_c;K)}$, we can obtain a lower bound for p_c:

$$p_c > (1 - p_{rh}^{(k_c;K)})p_c^{(k_c)} \tag{6.B.5}$$

The value of $p_c^{(k_c)}$ can be conservatively estimated from the formula:

$$1 - p_c^{(k_c)} < \prod_{k=1}^{K}(1 - \sin^2(\omega_k t_k/2)\frac{\gamma_k^2}{\omega_k^2}) \tag{6.B.6}$$

$$< \prod_{k, \frac{|E-\epsilon_k|}{\gamma_k} < 1} ((E - \epsilon_k)^2/\gamma_k^2), \tag{6.B.7}$$

where the second line follows from the inequality $\sin(\frac{\pi\sqrt{1+x^2}}{2})/(1+x^2) \geq \min(0, 1-x^2)$ applied to each term in the product. In the perfect resonance scenario, $|E - \epsilon_{k_c}| = 0$ and the cooling probability is exactly 1. The worst case scenario is when E is right between the two neighbouring ϵ_k's, thus $|E - \epsilon_{k_c}| = \frac{\alpha}{2}\gamma_k$. In this case, we can calculate the logarithm of (6.B.7) in the leading order of K^{-1}, α:

$$2 \sum_{k=k_c^{(-)}}^{k_c^{(+)}} \log \left| \frac{\epsilon_k - E}{\gamma_k} \right| = 2 \int_{\epsilon^{(-)}}^{\epsilon^{(+)}} \log \left| \frac{\epsilon - E}{\gamma(\epsilon)} \right| d\epsilon \frac{dk}{d\epsilon} \tag{6.B.8}$$

$$= \frac{2}{\alpha} \int_{\epsilon^{(-)}}^{\epsilon^{(+)}} \log \left| \frac{\epsilon - E}{\gamma(\epsilon)} \right| \frac{d\epsilon}{\gamma(\epsilon)}. \tag{6.B.9}$$

Here, we used the fact that $\gamma\alpha$ defines energy spacing (and so $\frac{d\epsilon}{dk} = \alpha\gamma(\epsilon)$), and introduced summation limits $k_c^{(\pm)}$, $\epsilon^{(\pm)}$ as the points where $\frac{\epsilon - E}{\gamma} = \pm 1$. As this implies scaling $\epsilon^{(\pm)} = E + O(\gamma)$, (6.B.9) should scale as $O(1/\alpha)$. The calculation can be completed for the LogSweep gradation ϵ_k, γ_k, which implies $\epsilon_k' \propto \gamma(\epsilon) \propto \epsilon$. In particular, if $x = \frac{\epsilon - E}{\gamma}$ then $dx = \frac{E d\epsilon}{\epsilon\gamma} = \frac{d\epsilon}{\gamma}(1 + O(1/K))$, and we have:

$$\frac{2}{\alpha} \int_{\epsilon^{(-)}}^{\epsilon^{(+)}} \log \left| \frac{\epsilon - E}{\gamma(\epsilon)} \right| \frac{d\epsilon}{\gamma(\epsilon)} = \frac{4}{\alpha} \int_0^1 \log x \; dx = -\frac{4}{\alpha}. \tag{6.B.10}$$

Substituting into Eq. 6.B.7, we find the initial cooling probability bounded by

$$p_c^{(k_c)} \gtrsim 1 - \exp(-4/\alpha(K)). \tag{6.B.11}$$

The reheating accumulated between steps k_c and K, $p_{rh}^{(k_c;K)}$, can be upper bounded as:

$$p_{rh}^{(k_c;K)} \leq 1 - \prod_{k=k_c}^{K} \left(1 - \sin^2(\Omega_k t_k/2) \frac{\gamma_k^2}{\Omega_k^2} \right) \tag{6.B.12}$$

The product in Eq. (6.B.12) can be further estimated as:

$$\prod_{k=k_c}^{K}\left(1-\sin^2(\Omega_k t_k/2)\frac{\gamma_k^2}{\Omega_k^2}\right) \geq \prod_{k=k_c}^{K}\left(1-\frac{\gamma_k^2}{\Omega_k^2}\right) \tag{6.B.13}$$

$$\geq \prod_{k=k_c}^{K}\left(1-\frac{\gamma_k^2}{(E+\epsilon_k)^2}\right) \tag{6.B.14}$$

$$\simeq \exp\left(-\sum_{k=k_c}^{K}\frac{\gamma_k^2}{(E+\epsilon_k)^2}\right), \tag{6.B.15}$$

where in the last line we assumed that $\gamma_k \ll E+\epsilon_k$ for all k. As we are most concerned about the large K asymptotics of the total cooling probability, let us now analyze how the expression (6.B.15) behaves in this limit. Since γ_k^2 scales as $O(1/K^2)$ and we have K terms in the sum, we generally expect $O(1/K)$ scaling for the sum. Such scaling would imply a rapidly vanishing reheating for a large-K protocol. In the specific case of the LogSweep protocol, to the leading order in $1/K$ one indeed obtains:

$$p_{rh}^{(k_c;K)} \lesssim \sum_{k=k_c}^{K}\frac{\gamma_k^2}{(E+\epsilon_k)^2} \approx \frac{1}{\alpha(K)}\int_{E_{\min}}^{E}\frac{\gamma(\epsilon)}{(E+\epsilon)^2}d\epsilon \tag{6.B.16}$$

$$\approx \frac{\log\frac{E_{\max}}{E_{\min}}}{\alpha^2(K)K}\left(\frac{1}{2}-\frac{E}{E+E_{\min}}+\log\left(\frac{2E}{E+E_{\min}}\right)\right) \tag{6.B.17}$$

$$\equiv \frac{\mathrm{R}(E_{\min},E_{\max},E)}{\alpha^2(K)K}. \tag{6.B.18}$$

Here, we used Eq. (6.4.7) and the fact that $\alpha(K)\gamma_k$ defines energy spacing $|\epsilon_{k+1}-\epsilon_k|$. Finally, combining Eqq. (6.B.5) - (6.B.18), we obtain an asymptotic lower bound to the final cooling probability:

$$p_c = \left(1-\exp\left(-\frac{4}{\alpha(K)}\right)\right)\cdot\left(1-\frac{\mathrm{R}(E_{\min},E_{\max},E)}{\alpha^2(K)K}\right). \tag{6.B.19}$$

This estimate implies $p_c \to 1$ for large K, provided that both $e^{-4\alpha^{-1}(K)} \to 0$ and $\frac{1}{K\alpha^2(K)} \to 0$.

To ensure that the infidelity is minimized and thus $\alpha(K)$ is optimal, we solve the extremum condition $\partial_\alpha(e^{-4\alpha^{-1}}+\frac{R}{\alpha^2 K})=0$ for α. The solution can be expressed in terms of the product logarithm function W, $\alpha(K) = 4\,W^{-1}(8K/R)$. For large K, at the leading order we obtain simply: $\alpha(K) = 4\ \log^{-1}(8K/R)$. The infidelity then scales down almost linearly

with K: $1 - p_c = \frac{\log^2(8K/R)}{16K}$. This asymptotically optimal $\alpha(K)$ yields the choice $\zeta(K) = \frac{1}{2\pi}\log(8K/R)$, which we use in all our simulations.

6.C Optimizing energy spacing in LogSweep protocol

In Sec. 6.4.2, we argued that the energy spacing of the LogSweep protocol is optimal for the protocol precision for a K-step protocol. This was based on the reheating estimate taken from the cooling step k_c only. One may ask, if this persists when one includes the total reheating into account. In the large K limit, we can use the estimate (6.B.15) for this check. Fixing the constraint $\gamma_k = \frac{|\epsilon_{k+1}-\epsilon_k|}{\alpha}$, we proceed by means of variational calculus:

$$\frac{\delta}{\delta\epsilon_k}\sum_{k=k_c}^{K}\frac{\gamma_k^2}{(E+\epsilon_k)^2} = 0 \tag{6.C.1}$$

$$\Rightarrow \frac{\delta}{\delta\epsilon(k)}\int_{k_c}^{K}\frac{(\epsilon'(k))^2}{(E+\epsilon(k))^2}dk = 0 \tag{6.C.2}$$

$$\Rightarrow \epsilon''(k)\cdot(E+\epsilon(k)) = (\epsilon'(k))^2. \tag{6.C.3}$$

The solution to Eq. (6.C.3) that satisfies boundary conditions $\epsilon(k_c) = E$, $\epsilon(K) = E_{\min}$, is as follows:

$$\epsilon_k = (2E)^{\frac{K-k}{K-k_c}}(E+E_{\min})^{\frac{k-K}{K-k_c}+1} - E. \tag{6.C.4}$$

This shows that the logarithmic character of the optimal spacing persists when we consider total reheating (cf. Eq. (6.4.7)). However, we cannot directly use the embelished result (6.C.4) for our cooling protocol. That is because this formula uses the targeted energy E as a reference, whereas we are targetting a continuum of energies. Therefore, we keep using the simpler and more practical formula Eq. (6.4.7) for the LogSweep protocol.

6.D Cooling rate for LogSweep protocol in a large system

In a large system, the above analysis is complicated by the presence of multiple transitions from every energy level. We now give a simplified analysis that focuses on a pair of states $|E_i\rangle$, $|E_j\rangle$, in a spirit similar

to Appendix 6.B. This means we formulate the protocol as a Markov process equivalent to (Eq. 6.B.1), where the transitions to levels other than i and j are ignored. Note that in the perturbative limit, this is a good approximation of the actual Markov process as restricted onto the subspace $|E_i\rangle, |E_j\rangle$. Specifically, even though we ignore the indirect transitions between i and j via other levels, this is justified at the first order of pertubation theory. Unlike in the $1+1$ model however, the transitions here are imperfect. If our total coupling has strength γ (i.e. $\|H_C\| = 2^N\gamma$), following the analysis in Sec. 6.4.1 the coupling between states $|E_i\rangle$ and $|E_j\rangle$ will take the form $\gamma\sqrt{A_{i,j}}$ with $\sqrt{A_{i,j}}$ scaling down as $O((E_i - E_j)^{-2})$. This has the effect of scaling both the cooling and re-heating rates by $A_{i,j}$, recasting the Markov process (Eq. 6.B.1) as

$$p_{i,j}^{(k)} = \begin{pmatrix} 1 - A_{i,j}\sin^2(\frac{\Omega_k t_k}{2})\frac{\gamma_k^2}{\Omega_k^2} & A_{i,j}\sin^2(\frac{\omega_k t_k}{2})\frac{\gamma_k^2}{\omega_k^2} \\ A_{i,j}\sin^2(\frac{\Omega_k t_k}{2})\frac{\gamma_k^2}{\Omega_k^2} & 1 - A_{i,j}\sin^2(\frac{\omega_k t_k}{2})\frac{\gamma_k^2}{\omega_k^2} \end{pmatrix}.$$

As this only reduces both the heating and cooling rates, our claim that reheating in the LogSweep protocol tends to 0 as $K \to \infty$ still holds. However, we need to repeat the analysis of App. 6.B to bound the cooling rate $p_c^{(k_c)}$ below and check that it continues to tend to 1. For the sake of generality, we drop the i, j indices, and consider a cooling probability restricted by a k-independent factor A.

With this adjustment, we may recast Eq.6.B.7 when $A << 1$ as

$$1 - p_c^{k_c} < \prod_{k, \frac{|E-\epsilon_k|}{\gamma_k}<1} \left[\left(1 - \frac{A\pi^2}{4}\right) + \frac{A\pi^4}{48}\frac{(E-\epsilon_k)^2}{\gamma_k^2}\right]. \tag{6.D.1}$$

Then, taking the log and converting again to an integral, we obtain

$$\log(1 - p_c^{k_c}) < \frac{1}{\alpha}\int_{\epsilon^-}^{\epsilon^+} \log\left[B + A'\frac{(E-\epsilon)^2}{\gamma(E)^2}\right]\frac{d\epsilon}{\gamma(\epsilon)}, \tag{6.D.2}$$

where $A' = \frac{A\pi^4}{48} \sim 2A$, and $B = 1 - \frac{A\pi^2}{4} < 1$. Next, setting $x = \frac{E-\epsilon}{\gamma(\epsilon)}$, and using the fact that for the LogSweep protocol $\gamma(\epsilon) \sim \epsilon$, we find

$$\log(1 - p_c^{k_c}) < \frac{2}{\alpha}\int_{-1}^{+1} \log\left(B + A'x^2\right) dx. \tag{6.D.3}$$

This may be evaluated by integrating by parts, giving

$$\log(1-p_c^{k_c}) < -\frac{2}{\alpha}\int_{-1}^{+1}\frac{x^2}{BA'^{-1}+x^2} \tag{6.D.4}$$

$$= \frac{-4}{\alpha}\left[1 - BA'^{-1}\tan^{-1}\left(A'B^{-1}\right)\right] \tag{6.D.5}$$

$$\sim -\frac{4}{3\alpha}A'^2B^{-2} + O(A^4). \tag{6.D.6}$$

Using the optimal scaling $\alpha(K) = 4\log^{-1}(K)$ we identified in Appendix 6.B, this adjusts our bound in the cooling rate to

$$p_c^{k_c} \gtrsim 1 - K^{-\frac{1}{3}A'^2B^{-2}}, \tag{6.D.7}$$

which continues to tend to 1 as $K \to \infty$, albiet at a rate reduced proportional to A.

This result requires some consideration in a large system — if our coupling Γ from a state $|E_i\rangle$ is spread over transitions to J states $|E_j\rangle$, we have $A_{i,j} \sim J^{-1}$, and the probability of any transition being cooled can be found to be

$$\prod_j (1-p_{c,j}^{k_c}) \sim e^{-\sum_{j=1}^{J}\frac{1}{3\alpha}A_j^2(1-A_j^2)} \sim e^{-\frac{1}{3\alpha J}}. \tag{6.D.8}$$

This implies that we require $\alpha \sim J^{-1}$ in order to maintain a constant cooling rate, which in turn may require adjustments to the optimal scaling identified in Appendix 6.B. As such adjustments are highly system-dependent, we do not investigate them further here.

6.E Effect of banding on QDC protocols

In this appendix we demonstrate the effect of banding on single sweeps of the LogSweep protocol. In Fig. 6.11, we plot the infidelity of a single shot of the LogSweep protocol with gradation number K acting on the maximally-mixed state, as a function of K (triangular markers). We see that in the critical case, the system continues to tend to the ground state polynomially in K. However, for the TFIM chain in the weakly- and strongly-coupled phases, we find that the protocol fails to converge as a function of K, due to the banding issue described above. This lack of convergence is rectified in the series marked by dots (same data as in Fig. 6.9) by repeating the LogSweep protocol as a function of K. We note

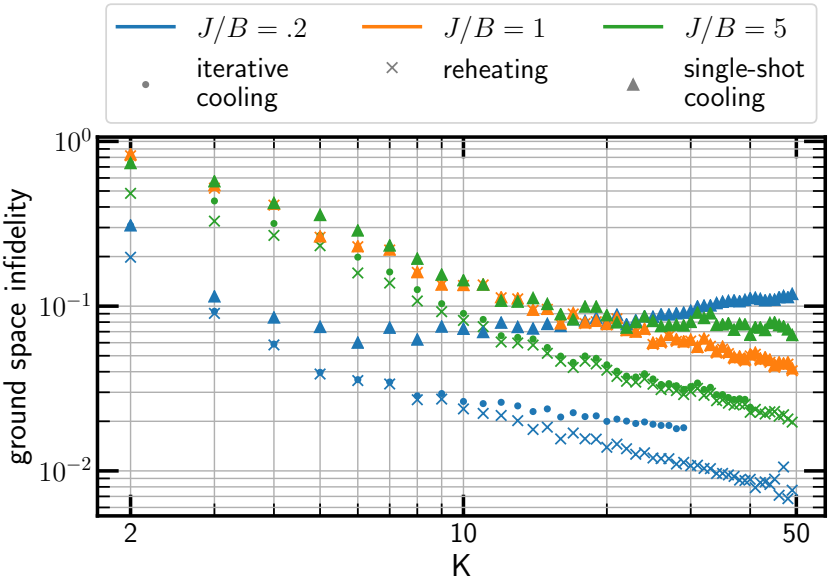

Figure 6.11: Difference between cooling by applying the a single LogSweep protocol with gradation number K (round markers), and iterating LogSweep for all $g_l = 2, ..., K$ (triangular markers). The iterative and reheating data are the same as in Fig. 6.9, the same context and simulation techniques apply.

that the failure in the strong-coupling case is not of the same degree as in the weak-coupling case, which we ascribe to the fact that the banding is not as strongly pronounced in Fig. 6.8, and so the result has not yet presented itself.

7 Measurement-driven navigation in many-body Hilbert space

7.1 Introduction

Quantum state preparation is a prominent routine in quantum information processing toolbox [185–199]. Such procedure often implies steering a quantum system from a "simple" towards a more complex, pre-designated resourceful state, e.g. a many-body entangled state. A steering protocol is characterized by an as short as possible runtime and high resulting overlap with the target state. Constructing such protocols can be done in multiple distinct ways. One is to design the Hamiltonian of the system, such that its unitary evolution leads to a designated state. This paradigm is represented by methods like digital computation or analog simulation [187, 189–192]. Such protocols require exact knowledge of the starting state, as well as the precise timing of the unitary evolution, to be accurate. Another strategy is to make use of the environment, adding a dissipative element to the evolution. Combined with the Hamiltonian evolution, this results in methods such as drive-and-dissipation [195, 196]. Finally, one can design a sequence of generalized measurements, which brings the system towards the target state via measurement back-action alone [200–202]. The relevant part of the evolution is then completely governed by the system-detector coupling (see also Ref. [203]). Unlike protocols involving pre-defined unitary evolution, such measurement-driven state preparation may not require knowledge of the starting state and fine-tuning of the system Hamiltonian [202].

The above types of state-preparation strategies can be referred to as *passive*, meaning that these protocols are pre-determined and pursued regardless of how the system evolves. Given this perspective, it appears beneficial to go beyond the forms of control described above, and introduce the concept of *active* decision making. This type of steering exploits information extracted during the system's evolution to decide on the

operations that follow. This is also referred to as closed-loop quantum control and is typically used to improve the Hamiltonian-based state preparation [69–73, 205]. In that case, extracting the necessary information requires the introduction of measurements into the protocol, which may result in an undesired back-action. Nevertheless, in certain cases, closed-loop control of Hamiltonian evolution does yield an improvement to the speed and the fidelity of the protocol.

Another possibility, which is a subject of increasing interest, is to employ active decision in measurement-driven protocols (implying no Hamiltonian drive) [74, 75, 206]. In such protocols, the necessary information about the system is naturally available from the employed measurements. The active decision is then being made about possible changes in the subsequent generalized measurements, such that the target state is prepared as rapidly and accurately as possible [74, 75, 201, 206]. Some general theorems have been stated concerning such active measurement-driven state preparation [206], along with some specific protocols designed to reach single-qubit target states [74, 75]. However, it remains unclear how an active measurement-driven protocol can be effectively harnessed to engineer resourceful many-body states. In this case, the large size of the Hilbert space makes it challenging to actively steer the system evolution in the desired direction.

In this chapter, we establish a general framework for measurement-driven active navigation in Hilbert space and construct active-decision protocols for measurement-only steering of many-body states. In particular, we focus on states manifesting genuine multipartite entanglement. When attempting to address the problem, one is naturally constrained by a few factors. One is that only reasonably local system-detector couplings are to be used in the protocol. Moreover, we require that the number of distinct system-detectors couplings available for steering does not scale up faster than the system's size (this number should not be super extensive). This natural requirement restricts the capabilities of the protocol and results in the need of correlating different system-detector couplings. Another prerequisite is that applying one type of coupling generally leads to an update in the expected benefits from other couplings. This phenomenon, which we refer to as "coupling frustration", calls for nontrivial coordination between different coupling applications. Finally, there is a problem of orienteering: it is relatively easy to "get lost" in the many-body Hilbert space when exploring it with the set of tools limited by locality and extensivity (cf. Ref. [205]).

We note that upon the availability of indefinite computational power, one can always find an optimum sequence of measurements through dynamic

programming techniques (cf. Ref. [206]). Roughly speaking, this can be done by considering all possible future quantum trajectories of the system. However, in a large Hilbert space, it is practically impossible to realize the theoretically optimal feedback policy. This is because the extensive consideration of outcome scenarios is too complex for a many-body Hilbert space of an already not very large system (it increases at least exponentially with the system's size and the duration of the protocol). Instead, we aim at designing heuristic strategies for active decision-making, which would allow for a significant – but not necessarily optimal – speedup of the protocol.

To meet these challenges, we introduce *Hilbert-space navigation techniques*. The first technique, which we term *greedy orienteering policy*, is based on the notion of a cost function. A simple example of such a cost function could be the target state infidelity. Minimizing it in a greedy protocol may already yield a reasonable advantage compared to the passive policy. To test this approach, we study numerically the preparation of a ground state of Affleck-Lieb-Kennedy-Tasaki (AKLT) spin-1 model [207]. The numerical study shows the speedup factor that increases with system size, reaching factor 9.5 for $N = 6$. Looking ahead, we discuss the fundamental challenge of landscape flatness that may arise for some target states when using simple infidelity as the cost function. Although this issue did not arise in the example we considered, we propose a possible modification to the cost function which should remedy this problem if it occurs.

The second technique is to map the Hilbert space onto a colored multigraph, referred to as the Quantum State Machine. The vertices of such a graph correspond to the basis states, and the edges represent the actions of generalized measurements. Upon an appropriate choice of basis states, such Quantum State Machine representations allow for improved navigation in Hilbert space. This can be done by heuristically representing it as quantum wayfinding on the graph. To substantiate this heuristic, we introduce the notion of semiclassical coarse-graining of a Quantum State Machine graph. Optimizing the exploration of these graphs by choosing the most appropriate system-detector couplings results in advantageous active-decision protocols. To exemplify this navigation paradigm we consider the preparation of the 3-qubit W-state, with a numerical study demonstrating a 12.5-fold improvement in protocol runtime.

Throughout the chapter, we assume that we know the initial state of the system. This can be a "cheap" (say, product) and robust quantum state that does not require many resources for its preparation. However, one can directly generalize the above approaches to the case where the initial state is unknown and is therefore represented by a density matrix. In the

more intricate case of a Quantum State Machine-based policy, one would then need to take a weighted combination of graph navigation protocols with different initial states.

The remainder of the chapter is organized as follows. In Sec. 7.2, we introduce the basics of measurement-induced steering. Specifically, in Sec. 7.2.1, we define the steering protocols and their elements, as well as the quantitative measure of the protocol's success. Then, in Sec. 7.2.2, we illustrate these definitions as applied to passive steering of a single qubit. The general selection criteria, including locality and extensivity, for the system-detector couplings, which are to be used for the active steering, are addressed in Sec. 7.2.3. In Sec. 7.3, we introduce the notion of frustration of steering and discuss the possibilities of protocols' speed-up for mutually commuting (Sec. 7.3.1) and non-commuting (Sec. 7.3.2) couplings. In the latter case, we develop a "parent-Hamiltonian" approach. A "quantum-compass" approach to active-decision steering, based on the greedy cost-function accumulation policy, is developed in Sec. 7.4, where we also employ this scheme to the preparation of the AKLT state. In Sec. 7.5, we develop an alternative active-steering framework – a "Quantum State Machine." In Sec. 7.5.1, we introduce the generalities of this approach based on the underlying representation of the steering protocol in terms of a quantum graph. Next, we discuss the quantum parts of this graph (Sec. 7.5.2), as well as the coarse-graining procedure, with the resulting coarse-grained graph being semiclassical (Sec. 7.5.3). This type of Hilbert-space orienteering is illustrated in Sec. 7.5.4, where an active-decision steering protocol for preparation of a three-qubit W-state is presented. Our findings are summarized and discussed in Sec. 7.6.

7.2 Measurement-driven state preparation

7.2.1 Generalities

Measurement-driven state steering is a specific class of state-preparation protocols. Its basic building blocks are coupling the quantum system (s) to quantum detectors (ancillary systems) utilizing engineered interactions, followed by strong measurement on the detectors (d). The goal of designing a measurement-based steering protocol is to generate a process that prepares the desired system state by utilizing a sequence of measurement back-actions.

Here, we will additionally assume that the internal evolution of the system and the detector are trivial (their Hamiltonians are kept null:

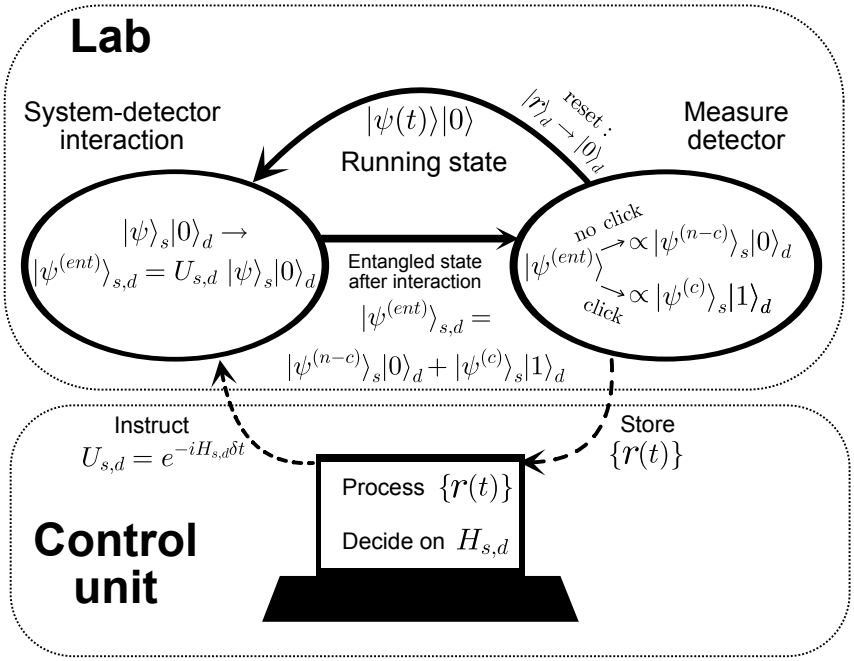

Figure 7.1: Basic design of the measurement-driven state preparation. The procedure starts with a given initial state $\rho_s^{(in)}$ and proceeds with a protocol, as described in Def. 32, until a good accuracy of the target state $|\psi_0\rangle$ is achieved. The control unit decides on the system-detector interaction unitary $U_{s,d}$ based on the stored record of detector readouts. We focus on constructing an optimized policy for decision-making, such that the target state is simulated as efficiently as possible.

$H_s = 0$, $H_d = 0$), as in Refs. [74, 75, 206], so that the unitary dynamics in the problem is governed solely by the coupling between the system and detectors determined by Hamiltonian $H_{s,d}$. For concreteness of analysis, we also constrain the detector to be a qubit initialized in a trivial state $|0\rangle$, and the system to be represented by N spins. Although a general spin S can be considered, we focus on the cases $S = 1/2$ and $S = 1$. We assume certain knowledge about the initial state of the system, which is described by the initial density matrix $\rho_s^{(\mathrm{in})}$. For the sake of simplicity, we further address the target state which is a pure state $|\psi_0\rangle$.

Although the ensuing protocol can be further generalized (see Section 7.6), we now formally fix its structure as given below:

Definition 32. *A measurement-driven state steering is a protocol that is performed to prepare a state $|\psi_0\rangle$, starting from the state $\rho_s^{(in)}$. It runs by repeating iterative cycles of the following form (see Fig. 7.1):*

1. *Prepare the detector qubit in the state $|0\rangle$.*
2. *Based on the available information, select the system-detector coupling Hamiltonian $H_{s,d}$ to be used in the next step.*
3. *Perform a system-detector evolution governed by a Hamiltonian $H_{s,d}$ for a short time interval δt: $U_{s,d} = e^{-iH_{s,d}\delta t}$.*
4. *Once the system-detector evolution is over, projectively measure the detector qubit in the Z-basis. Store the readout r for further processing.*
5. *Decide whether the protocol is to be continued or terminated. In the former case, return to step 1.*

Now, in the vast space of protocols that have such structure, we would like to emphasize the distinction between two classes of protocols: passive and active.

In a passive protocol, the stored readouts $\{r(t)\}$ from step 4 may influence the decision for protocol termination or continuation at step 5, but not the choice for the interaction Hamiltonian $H_{s,d}$ made at step 2 in the next protocol cycles. Hamiltonians $H_{s,d}$ can still be chosen differently for different iterations: e.g. for a large system, the detector qubit can be coupled to different subsystems thereof. However, $H_{s,d}$ used at each cycle in the passive protocol has to be pre-determined from the outset. If a passive protocol also has a pre-determined duration (and thus doesn't use readouts $\{r(t)\}$ at the termination step), we land in a subclass of passive protocols where the readouts don't have any influence on the protocol. We would refer to such protocols as "blind steering". For blind steering, the readouts of the detector at any given step can be averaged, i.e., following the measurement, the detector's density matrix is traced out. In this chapter, however, we will focus on the non-blind version of passive steering, where readouts are indeed employed for an informed protocol termination.

In contrast to passive protocols, in an active protocol one uses the readouts $\{r(t)\}$ to make an informed decision for the interaction Hamiltonians $H_{s,d}$ as well as for termination/continuation of the protocol. Active decision-making has to follow a certain policy, which becomes the crucial part of the protocol. For a good active policy, its adoption should result in

a significant speedup of the protocol compared to its passive counterpart. Alternatively, one can also fix the protocol runtime and aim to improve the precision of the state preparation. We focus on the former target: minimizing protocol runtime for a fixed target precision. The major challenge in this chapter is to construct such advantageous active decision-making policies. By comparing active steering with the (non-blind) passive steering as defined above, we investigate the advantage offered specifically by the directed evolution, i.e. active decision-making for $H_{s,d}$.

Before we move on to the issue of active policy constructions, let us discuss the criteria for termination of a running protocol. In general, one cannot guarantee "perfect steering," i.e., obtaining the desired target state with the fidelity of 1 in a finite number of protocol cycles. Instead, one may consider preparing the target state with *infidelity* R:

$$R\left(\rho_s^{(\mathrm{fin})}, |\psi_0\rangle\right) \equiv 1 - \langle\psi_0|\rho_s^{(\mathrm{fin})}|\psi_0\rangle, \tag{7.2.1}$$

where the state $\rho_s^{(\mathrm{fin})}$ is the final state of the system once the protocol is terminated. It is worth emphasizing that the system evolution during the protocol is probabilistic and depends on the stochastic readouts $\{r(t)\}$. It follows that different runs of the same protocol may yield different $\rho_s^{(\mathrm{fin})}$ and, thus, the infidelity R. Therefore, to characterize the protocol as a whole, we introduce the following accuracy measure:

Definition 33. *We refer to a measurement-driven state-preparation protocol as ϵ-precise, if the infidelity between the final state and the target state is bounded by ϵ for any run of the protocol:*

$$R\left(\rho_s^{(fin)}, |\psi_0\rangle\right) < \epsilon. \tag{7.2.2}$$

Given the knowledge of the readout sequence, we may simulate the quantum system state (the quantum trajectory) on a computer in parallel to the measurement run. Thus one can infer the running system state exactly (referred to as filtering in the literature [69]), and test inequality (7.2.2). This sets a trivial criterion for protocol termination, which we will apply by default to all passive and active protocols considered in this chapter. Namely, a protocol can be terminated right after the cycle when the target state infidelity becomes smaller than ϵ, thus making it an ϵ-precise protocol. Apart from controlling the precision, we are interested in the number of cycles N_c, after which the protocol has been terminated. As N_c may also differ greatly, depending on a specific run, we will characterize the protocol by $\langle N_c\rangle_{\mathrm{run}}$, where the averaging is performed over many runs.

Note that here averaging is taken over stochastic readout sequences. In reality, steering errors as well as external noise may be further contributing factors to the stochasticity. For the given target state and target precision ϵ (cf. Definition 33), our goal is to find an ϵ-precise protocol such that $\langle N_c \rangle_{\text{run}}$ is as small as possible. We will be considering this minimization as the key goal of our constructions.

7.2.2 Passive steering: Single qubit

As a simple example of a measurement-driven protocol, we consider single-qubit steering (for a more general consideration, the reader is referred to Sec. 7.2.3). For simplicity, we will assume the target state to be $|0\rangle$, and the starting state to be a perfectly mixed state: $\rho_{\text{start}} = \text{diag}(1/2, 1/2)$. A single coupling suffices to guarantee the preparation of the target state (in fact, from an arbitrary starting state) with an arbitrary precision [202]:

$$H_{s,d} = \gamma \sigma_s^- \sigma_d^+ + \text{H.c.} \tag{7.2.3}$$

Here, σ_s and σ_d are Pauli matrices acting in the system and detector spaces, respectively. By construction, a protocol that operates with only a single coupling Hamiltonian $H_{s,d}$, i.e., without a readout-based option of choosing different couplings, is considered passive. Nevertheless, even for passive protocols, one can introduce a policy based on the measurement outcomes, which would accelerate quantum-state steering.

Let us first address a protocol that runs for $N_c^{(\text{pass})}$ cycles using the coupling (7.2.3), regardless of the measurement outcomes. Under the definition given in Sec. 7.2, this would be an example of *blind* steering. In this case, the probability of obtaining a readout $r = 0$ decreases exponentially with the total number of cycles N_c. Tracing out detector outcomes (since we are blind to measurement outcomes), this results in a density matrix:

$$\rho(N_c) = \begin{pmatrix} 1 - e^{-N_c \gamma^2 \delta t^2}/2 & 0 \\ 0 & e^{-N_c \gamma^2 \delta t^2}/2 \end{pmatrix}. \tag{7.2.4}$$

Given the threshold infidelity ϵ, we need to run the protocol for $N_c^{(\text{pass})}(\epsilon)$ cycles:

$$N_c^{(\text{pass})}(\epsilon) = \frac{1}{\gamma^2 \delta t^2} \log\left(\frac{1}{2\epsilon}\right) \tag{7.2.5}$$

This characterizes the efficiency of the completely blind passive protocol [202] for the single-qubit setup.

Next, we consider passive protocols where the sequence of readouts is recorded. One then needs to interpret the measurement outcomes, which for this setup is straightforward. We note that when the readout is $r = 1$ (click event), the target state is instantly prepared [cf. Eq. (7.2.9)]. Therefore, one can terminate the protocol directly after the detector clicks for the first time: in this case, all further steps are simply redundant and do not result in any evolution of the system. This will constitute a termination-policy improvement of the passive blind protocol for this single-qubit case. If $r = 0$, i.e. no click is measured (such a null-measurement [208] event still gives the system a nudge towards the target state by measurement back-action), the protocol simply continues until a certain maximal number of cycles, $N_c^{(\max)}$. The target infidelity ϵ would be directly related to $N_c^{(\max)}$ in a way equivalent to the blind protocol runtime (7.2.5). The average runtime of the non-blind passive protocol, $\tilde{N}_c^{(\mathrm{pass})} \equiv \langle N_c \rangle_{\mathrm{run}}$, is then given by

$$\tilde{N}_c^{(\mathrm{pass})} = \frac{1}{2\gamma^2 \delta t^2} \left(1 - e^{-\gamma^2 \delta t^2 N_c^{\max}}\right) + \frac{N_c^{\max}}{2}. \tag{7.2.6}$$

This runtime is strictly smaller than the runtime for the passive blind protocol, Eq. (7.2.5), and yields a twofold speedup in the $\epsilon \to 0$ limit. It is worth emphasizing, however, that the termination policy can realistically be applied only to the case of few-body quantum states. For such a policy to be useful, a single detector click should signify that the system is in the target state. This can only be realized when the detector is coupled to all elements of the system. For a many-body system (many qubits), a natural assumption of locality rules out such a coupling: a click of the detector coupled to a subsystem of the system does not guarantee that the whole system is steered to the desired state. Nevertheless, the above simple example shows that detector readouts can be used for accelerating the state preparation. In what follows, we will focus on active feedback strategies. There, instead of protocol termination, the local-measurement outcomes are employed for choosing the most efficient sequence of further measurement cycles.

7.2.3 Selection criteria for system-detector couplings

Families of system-detector couplings

Both for the active and passive protocols, a key feature is the choice of coupling Hamiltonians $H_{s,d}$. Given the target state $|\psi_0\rangle$, it is natural to constrain this choice to a certain family $\{H_{s,d}(\mathbf{p})\}$, for a set of (discrete or

continuous) parameters $\mathbf{p}$. Deciding on the choice of $H_{s,d}$ in each protocol cycle translates into selecting the value of $\mathbf{p}$. Before discussing the policies for doing so, we address a different question: How to effectively preselect this family $\{H_{s,d}(\mathbf{p})\}$? To answer this question, it is useful to consider the following general decomposition of $H_{s,d}$:

$$H_{s,d} = V_s \sigma_d^+ + V_s^\dagger \sigma_d^- + \tilde{V}_s \sigma_d^z, \tag{7.2.7}$$

where V_s and $\tilde{V}_s$ are arbitrary system operators and matrices $\sigma_d^\pm = \frac{1}{2}(\sigma_d^x \pm i\sigma_d^y)$ act on the detector. In Eq. (7.2.7), we discard any terms of the form $\sim \mathbb{I}_d$, as those represent the internal system evolution. Furthermore, for our purposes, it is also sufficient to consider a special case of the decomposition, where $\tilde{V}_s = 0$.

With Eq. (7.2.7) in mind, let us consider the transformation of the system state ρ_s under a single cycle of the steering protocol. First, let us consider the system state transformation that is performed when the measurement outcomes are averaged over (blind measurement). In the weak measurement limit, $\delta t \to 0$, this is represented by the map:

$$\begin{aligned} &\rho_s \to \Lambda_{V_s}(\rho_s) \\ &\quad \equiv \left(1 - \frac{\delta t^2}{2} V_s^\dagger V_s\right) \rho_s \left(1 - \frac{\delta t^2}{2} V_s^\dagger V_s\right) + V_s \rho_s V_s^\dagger \delta t^2. \end{aligned} \tag{7.2.8}$$

We note that the terms of order $\mathcal{O}(\delta t^2)$ in this expression represent the standard Lindbladian jump operator. Based on the map (7.2.8) by tracing out the detector readouts after each step, one derives a Lindblad equation describing the system evolution for the blind steering [202].

Let us now turn back to our protocol, where the different measurement outcomes are discriminated. During step 4 of the protocol cycle (cf. Definition 32), there is a probability

$$p^{(\mathrm{cl})}(\rho_s, V_s) = \delta t^2 \operatorname{tr}(V_s \rho_s V_s^\dagger)$$

that a qubit flip is measured in the detector (click probability). The resulting state in the limit of small δt is then:

$$\rho_s \to \Lambda_{V_s}^{(\mathrm{cl})}(\rho_s) \equiv \frac{V_s \rho_s V_s^\dagger}{\operatorname{tr}(V_s \rho_s V_s^\dagger)}. \tag{7.2.9}$$

A "no-click" scenario occurs with probability

$$p^{(\mathrm{ncl})}(\rho_s, V_s) = 1 - p^{(\mathrm{cl})}(\rho_s, V_s),$$

and results in a state:

$$\rho_s \to \Lambda_{V_s}^{(\mathrm{ncl})}(\rho_s) \equiv \frac{\left(1-\frac{\delta t^2}{2}V_s^\dagger V_s\right)\rho_s\left(1-\frac{\delta t^2}{2}V_s^\dagger V_s\right)}{1-\delta t^2 \mathrm{tr}(V_s^\dagger V_s \rho_s)}. \tag{7.2.10}$$

Note that for the weak-measurement limit considered here ($||V_s\delta t|| \ll 1$), the click probability is parametrically smaller than that for the no-click event: a qubit flip can be recorded in the detector only rarely.

Extensivity and locality

We are now in a position to expound our considerations for the family $\{H_{s,d}(\mathbf{p})\}$ in terms of the operators $\{V_s(\mathbf{p})\}$. For a meaningful comparison between active and passive protocols, we first require that there exists a passive protocol that employs Hamiltonians $\{H_{s,d}(\mathbf{p})\}$ to reach the target state $|\psi_0\rangle$. For concreteness, we assume that the passive protocol involving all the family members is a cyclic one: each of the couplings $\{H_{s,d}(\mathbf{p})\}$ is employed one after another in a predefined manner, and the cycle is repeated once all the couplings are employed. The size of the family is restricted by the number of available system-detector couplings, which is assumed to scale up with increasing systems' size not faster than extensively.

It is then natural to demand that none of $\{H_{s,d}(\mathbf{p})\}$ can move the system state away from the target state. Given Eqs. (7.2.9) and (7.2.10), this yields a dark-state condition $V_s(\mathbf{p})|\psi_0\rangle = 0$ for every $\mathbf{p}$. This is equivalent to every operator $V_s(\mathbf{p})$ taking the following form:

$$V_s = \sum_{\alpha=1}^{D-1} v_\alpha |\psi_0\rangle\langle\psi_\alpha| + \sum_{\alpha,\beta=1}^{D-1} w_{\alpha\beta}|\psi_\beta\rangle\langle\psi_\alpha|, \tag{7.2.11}$$

where D is the Hilbert-space dimensionality of the system, and $\{|\psi_\alpha\rangle\}$ is any (many-body) basis for the system that includes $|\psi_0\rangle$ as a basis state. This general form of the system-detector coupling is, however, not realistic for many-body systems, as $D = 2^N$ grows exponentially with the number of qubits for an N-qubit system. Thus, having in mind steering of many-body states, we should further restrict the family of available steering operators.

The second condition for $\{V_s(\mathbf{p})\}$ is that these couplings can realistically be engineered in an experimental realization of the system. In this chapter, we focus on the most basic aspect of this condition: locality. One may consider two types of locality: geometric and operator (k-locality [16]).

Geometric locality of the operator V_s implies that such interaction only requires coupling the system spins that are in geometrical proximity during the experiment. A k-local operator V_s implies that only k system spins are coupled at a time. It is natural to impose the locality constraint not on the full operator V_s, but its individual terms. For example, if V_s involves all system spins, but its individual terms only couple 2 spins at a time, we will consider V_s a 2-local coupling (in line with [16]). Note that a k-local operator V_s implies an interaction Hamiltonian $H_{s,d}$ that is $(k+1)$-local.

Sufficient conditions for the coupling operators. Room for active decision-making

It is worth stressing at this point that $V_s(\mathbf{p})$ following the form given by Eq. (7.2.11) for all $\mathbf{p}$ is necessary but not sufficient for $|\psi_0\rangle$ to be the only dark state of the passive protocol. For some choices of such a family $\{V_s(\mathbf{p})\}$, a spurious final state $|\psi_0'\rangle \neq |\psi_0\rangle$ might be reached. However, this would imply a dark-state condition $V_s(\mathbf{p})|\psi_0'\rangle = 0$ (for every $\mathbf{p}$), and this should not hold for a generic (say, random-matrix-type) operator V_s, which satisfies $V_s|\psi_0\rangle = 0$. For generic coefficients v_α and $\omega_{\alpha\beta}$ in (7.2.11), one does not expect an existence of a spurious final state (for that, an extra constraint is needed, such as vanishing of certain v_α, $\omega_{\alpha\beta}$, or a specific relation between the coefficients).

One concludes that a family consisting of a single Eq. (7.2.11)-type coupling V_s is sufficient to prepare $|\psi_0\rangle$ in a passive protocol without generating a dark space. Notably, reducing the family to a single member would leave no room for active decision-making in a protocol defined by this family (an active protocol necessitates at least two operators to choose from). On the other hand, such an ultimate V_s would not generically satisfy the crucial locality conditions and, thus, would be unrealistic to implement. Natural counterexample couplings V_s' that have multiple dark states arise in the important case when V_s' acts only on a part of the system.

To construct such a counterexample, one may start from an arbitrary operator V_s that satisfies the dark-state condition $V_s|\psi_0\rangle = 0$ for a single state $|\psi_0\rangle$ in a given system. Now, consider a larger system embedding the original one and construct a different target state which is a tensor product of $|\psi_0\rangle$ and a certain auxiliary state: $|\Psi_0\rangle \equiv |\psi_0\rangle \otimes |\tilde{\psi}_0\rangle$. In this case, one may take $V_s \to V_s'$, where $V_s' = V_s \otimes \mathbb{I}_{\tilde{s}}$ still satisfies condition $V_s'|\Psi_0\rangle = 0$ relative to this new target state in the extended Hilbert space. Yet for a general starting state of the total system, the operator V_s' is obviously not sufficient to prepare the extended target state $|\Psi_0\rangle$ – also

implying the existence of spurious dark states (in fact, all states of the form $|\psi_0\rangle \otimes |\tilde{\psi}_0\rangle$ turn out to be dark, for arbitrary $|\tilde{\psi}_0\rangle$).

We see that in this "self-evident" fashion of selecting a steering operator, the condition of $|\Psi_0\rangle$ being a dark state for a single operator V_s' was not sufficient for V_s' to be able to guarantee preparation $|\Psi_0\rangle$. As discussed above, an operator V_s capable of steering a unique dark state is typically highly nonlocal, in contrast to the limited capacity of a localized operator $V_s \otimes \mathbb{I}_{\tilde{s}}$. We conclude that a family of multiple operators $\{V_s(\mathbf{p})\}$ is needed to realistically prepare a target state, once that state is sufficiently complicated. This, in turn, opens the door for gaining advantage through active decision-making.

7.3 Types of system-detector couplings

The preselected family of coupling operators $\{V_s(\mathbf{p})\}$ determines both the performance of the ensuing passive protocols and the possibilities for active policy construction. In the present section, we identify the crucial role of the commutation properties of $\{V_s(\mathbf{p})\}$. We first consider N-qubit steering protocols which employ coupling operators $\{V_s(\mathbf{p})\}$ that are mutually commuting. As a shorthand, we denote this as non-frustrated steering. We show that a realistic passive protocol of this type can be designed for product states and certain graph states. Commuting couplings also allow for a simple feedback strategy, which results in a significant speedup of the respective passive protocol. Next, we move on to passive steering protocols that are frustrated. Such frustration of local couplings naturally arises for many-body target states related to local parent Hamiltonians. We propose an explicit method of constructing a family of non-commuting operators $\{V_s(\mathbf{p})\}$ that allows to prepare such a many-body target state in a passive protocol. This forms the basis for Secs. 7.4 and 7.5, where we move on to the active versions of frustrated steering protocols.

7.3.1 Mutually commuting couplings

Here we focus on N-qubit steering protocols implemented with mutually commuting couplings $\{V_s(\mathbf{p})\}$. As will be demonstrated, a passive protocol of this type can be constructed for an arbitrary target state, yielding an asymptotically precise passive preparation. However, we find that this construction would, in general, require non-local system-detector couplings $H_{s,d}$, deeming the implementation of the protocol for many-body states impractical. We then identify an exception to this rule – a subclass of

graph states that can be obtained using local commuting couplings. For this, we discuss the constraints coming from both geometric locality, as well as k-locality. Finally, we extend the discussion from such passive protocols to their active counterparts. To achieve this, we propose a simple feedback strategy that speeds up non-frustrated steering in a substantial way.

As a trivial example of non-frustrated steering, consider an N-qubit product state as a target state, e.g., $|00..0\rangle$. The starting state will be assumed to be the perfectly mixed state. To prepare it with a steering protocol, one can use a set of couplings parameterized by the qubit number $i = 1, ..N$:

$$V_s^{(\mathrm{prod})}(i) = \gamma \sigma_i^-. \tag{7.3.1}$$

Passively alternating between the steering cycles employing $V_s^{(\mathrm{prod})}(i)$ with different i guarantees preparation of the target state with any given accuracy. This directly follows from the analysis of Secs. 7.2.2 and 7.2.3. For an active version of the protocol, partial protocol termination can be applied: if a click is registered when measuring any qubit i, the coupling $V_s^{(\mathrm{prod})}(i)$ is dropped out from the sequence of couplings that will be applied in further cycles. In other words, the steering with this "fired" coupling is terminated at this point, whereas other couplings remain active – hence the term "partial termination". Since this implies a readout-based decision on the set of steering couplings that are used at a given step, we classify this as an active steering protocol. In the $\epsilon \to 0$ limit, this strategy results in the following relation between active and passive runtimes:

$$N_c^{(\mathrm{act})}(\epsilon) = \frac{N_c^{(\mathrm{pass})}(\epsilon)}{2} + \frac{N}{2\gamma^2 \delta t^2}, \tag{7.3.2}$$

which leads to up to a substantial 2-fold speedup for the active version, similarly to Eq. (7.2.6).

Non-frustrated steering towards any target state $|\psi_0\rangle$ can in principle be designed if we allow for an arbitrary coupling set. Indeed, given a many-body unitary transformation to $|\psi_0\rangle$ from a product state $|00..0\rangle$, i.e., $|\psi_0\rangle = U_\psi |00..0\rangle$, one may formally construct a family of couplings:

$$V_s^{(U_\psi)}(i) = \gamma U_\psi \sigma_i^- U_\psi^\dagger. \tag{7.3.3}$$

Clearly, any protocol for $|\psi_0\rangle$ preparation using couplings of the form of Eq. (7.3.3) would be a unitary equivalent of the same protocol which uses couplings of Eq. (7.3.1) to prepare $|00..0\rangle$. Therefore, a passive protocol iterating over $V_s^{(U_\psi)}(i)$ for different i would successfully prepare the target

state $|\psi_0\rangle$. We also conclude that a partial-termination strategy can be applied to this coupling set with the same effect as for the product state target. Note, however, that in most cases employing $V_s^{(U_\psi)}(i)$ would not be practically feasible. Indeed, since U_ψ is a general many-body operation, the couplings $V_s^{(U_\psi)}(i)$ would involve arbitrarily non-local terms. For most N-spin states $|\psi_0\rangle$ with large N, one thus expects that the resulting $V_s^{(U_\psi)}(i)$ would break any requirement of geometric or k-locality.

This locality-violation rule can be circumvented for U_ψ which is given by a shallow circuit and thus $|\psi_0\rangle$ which is weakly entangled. As a resourceful example of such $|\psi_0\rangle$, consider a graph state defined on a generic graph G [209]:

$$|\psi_G\rangle = \left(\prod_{\substack{(j,k)\in \\ \text{edges}(G)}} U^{(\text{gr})}_{(j,k)} \right) \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right)^{\otimes N}, \tag{7.3.4}$$

$$U^{(\text{gr})}_{(a,b)} = \exp\left(\mathrm{i}\pi |00\rangle\langle 00|_{a,b}\right), \tag{7.3.5}$$

in which case

$$U_\psi = \left(\prod_{\substack{(j,k)\in \\ \text{edges}(G)}} U^{(\text{gr})}_{(j,k)} \right) \left(\prod_{j\in \text{qubits}} \exp\left(\mathrm{i}\frac{\pi}{4}\sigma_j^y\right) \right).$$

Since two-qubit rotations $U^{(gr)}_{(j,k)}$ all mutually commute, the coupling $V_s^{(U_\psi)}(i)$ acts only on spin i and on the spins j whose vertices share an edge with i in the graph G. Therefore, this coupling is $(k+1)$-local if there are k edges coming out of vertex i. Moreover, $V_s^{(U_\psi)}(i)$ is also geometrically local, if the graph G only connects the qubits which are in geometric proximity. We conclude that for the graphs satisfying the above conditions, a realistic preparation of graph states with local non-frustrated steering is possible. Such a protocol can be sped up in the same way it was possible for the product states – using active feedback via the partial-termination strategy.

For the perfectly mixed starting state, the partial-termination policy gives an optimal speed-up of a protocol driven by non-frustrated couplings $V_s^{(U_\psi)}(i)$. Indeed, the protocols in question are then equivalent to an independent set of N 1-qubit steering protocols (under the unitary transformation U_ψ). This picture, however, breaks down for a more general

starting state. Let us first consider the trivial target $U_\psi = \mathbb{1}$, $|\psi_0\rangle = |00..0\rangle$, while the starting state is itself entangled (e.g. $\frac{1}{\sqrt{2}}(|00..0\rangle + |11..1\rangle)$). In this case, the click received from a single coupling $V_s(i)$ may imply that multiple couplings can be dropped from the applied sequence, and not just $V_s(i)$ itself. This would be more optimal than the partial termination strategy outlined above. The same picture extends to the more interesting case when the target state $|\psi_0\rangle$ is entangled itself, e.g. a graph state, while the starting state is a product state. Indeed, under the unitary mapping U_ψ which takes an entangled state $|\psi_0\rangle$ to $|00..0\rangle$, the product starting state in turn becomes entangled. Hence, the previous reasoning applies and partial-termination would generally not be an optimal active policy in this situation. Instead, one may accelerate it further by applying one of the frustrated-coupling strategies outlined in the following sections.

7.3.2 Frustrated system-detector couplings

From now on, let us focus on accelerating steering protocols which employ couplings constrained by at least one of the two notions of locality. Under this premise, for target states other than the product states and states prepared by a shallow circuit, we would generally need to go beyond the non-frustrated protocols outlined above. The first question to tackle is how to design the local couplings V_s that are suitable for a passive protocol. In principle, this can be addressed on a case-by-case basis, tailoring some coupling set with a specific target state in mind. (This approach will be demonstrated for the W-state preparation in Sec. 7.5.4.) However, this is not always a straightforward task. Therefore, it is interesting to know whether one can devise a general scheme to this end. For this, we propose an approach based on a parent-Hamiltonian construction.

The parent Hamiltonian H_ψ of $|\psi_0\rangle$ has $|\psi_0\rangle$ as a non-degenerate ground state, and is constructed from $|\psi_0\rangle$ in the form:

$$H_\psi = \sum_j H_\psi^{(j)}. \tag{7.3.6}$$

Here, all the terms $H_\psi^{(j)}$, while defined as acting on the entire Hilbert space, are local in the real space given that the state $|\psi_0\rangle$ hosts a limited amount of entanglement [210] (implying an area-law dependence of the entanglement entropy accommodated in $|\psi_0\rangle$). Note that, by the construction of the parent Hamiltonian, terms $H_\psi^{(j)}$ have $|\psi_0\rangle$ as their common ground state, although they generally do not commute with each other: this is possible because of their respective ground state degeneracy. These degenerate

ground spaces of $H_\psi^{(j)}$ will be the central element of our construction of the coupling family. For the term $H_\psi^{(j)}$, which nontrivially acts on some m qubits, let us denote its m-qubit ground states as $|\phi_a^{(j)}\rangle$ and the excited states $|\theta_a^{(j)}\rangle$. Given these, we can construct the coupling operators of the following form:

$$V_s^j(\mathbf{w},\mathbf{v},\mathbf{u}) = \sum_{ab} w_{ab}|\phi_a^{(j)}\rangle\langle\theta_b^{(j)}| + \sum_{ab} v_{ab}|\theta_a^{(j)}\rangle\langle\theta_b^{(j)}| + \sum_{ab} u_{ab}|\phi_a^{(j)}\rangle\langle\phi_b^{(j)}|. \tag{7.3.7}$$

A particular example of this construction will be addressed in detail in the context of the AKLT model in Sec. 7.4 (see also Refs. [195, 202]).

For a generic (fixed) value of the parameters involved, running a passive protocol with the coupling given by Eq. (7.3.7) allows one to steer the system into the ground state of $H_\psi^{(j)}$. Alternating the coupling operators by selecting terms with different j at different measurement steps allows for steering the system into the joint ground space of all couplings $H_\psi^{(j)}$. This space is given by the target state $|\psi_0\rangle$ only, as it is the non-degenerate ground state of H_ψ. Thereby, as long as the parent Hamiltonian H_ψ is local, we have managed to construct an appropriate coupling set for a passive protocol (also see [195] for a related statement proven in more detail).

Now, let us consider an active-protocol construction. First, we note that unlike in the "non-frustrated" protocol construction, the operators V_s^j for different choices of parameter sets are, in general, not mutually commuting. This also applies to the couplings with different values of j, as $H_\psi^{(j)}$ generally do not commute. Therefore, the measurement outcome of steering by V_s^j (where, because of the locality of $H_\psi^{(j)}$, j corresponds to a certain region in real space) impacts the outcomes of steering at other locations. As a result, the partial-termination strategy cannot be applied to this coupling set, as it assumes that the respective cycles of the protocol can be considered separately. Instead, the feedback strategy for the frustrated steering should continuously coordinate the application of different couplings in the protocol. In a many-body context, this becomes a complicated navigation-type problem (cf. Ref. [205]). We devote the following two sections to the study of such possible coordination policies.

7.4 Quantum compass: Cost-function policies

One way to enable the Hilbert-space navigation is to introduce a cost function $C(\rho_s)$, which is to be minimized in the protocol. The basic example would be the infidelity $C(\rho_s) = R(\rho_s, |\psi_0\rangle)$ of the system state ρ_s to the target state $|\psi_0\rangle$, defined in Eq. (7.2.1). Achieving the global minimum $R(\rho_s, |\psi_0\rangle) = 0$ of this cost function would be equivalent to preparing the target state. In general, to calculate $R(\rho_s, |\psi_0\rangle)$, one needs to know the state of the system ρ_s. This is, in principle, feasible, as we control the system evolution given all measurements outcomes and therefore can numerically simulate it in parallel to the experiment. However, the requirement of such a simulation being done in parallel to the experiment puts a restriction on the size of the system that one can work with. For now, we will accept this limitation; finding ways to mitigate it is among the worthwhile potential extensions of the work presented in this chapter.

With a given cost function $C(\rho_s)$ at hand, we can use it to form the active decision for the coupling operator $V_s(\mathbf{p})$. The ultimate strategy is to pick $V_s(\mathbf{p})$ which brings the system to the global minimum $|\psi_0\rangle$ in the fastest expected time. For $C(\rho_s) = R(\rho_s, |\psi_0\rangle)$ this is equivalent to the ultimate strategy defined by dynamic programming [206], requiring unrealistic computation power. Therefore, the notion of the global cost function does not give any additional advantage in constructing such a strategy. Instead, one can use its cheaper version – the "greedy strategy." Specifically, one can use $V_s(\mathbf{p})$ that yields the fastest expected reduction of the cost function in a single step of the evolution:

$$V_s^{(\mathrm{greed})}(\mathbf{p}) = \mathrm{argmin}_{V_s(\mathbf{p})} R[\Lambda_{V_s(\mathbf{p})}(\rho_s)], \tag{7.4.1}$$

where $\Lambda_{V_s(\mathbf{p})}(\rho_s)$ is defined in Eq. (7.2.8) If there are multiple minima, we will assume that argmin returns a random representative among those. With only a small amount of computations needed to decide for the optimal next coupling $V_s^{(\mathrm{greed})}(\mathbf{p})$, this greedy procedure allows us to avoid the complex long-term analysis of the protocol.

As one can see from a direct implementation, the greedy minimization of the cost function can accelerate the state preparation by a large factor. To demonstrate this, we consider the example of the ground state in an AKLT spin chain as the target state. This is an entangled state of N spin-1 particles governed by the Hamiltonian H_{AKLT}:

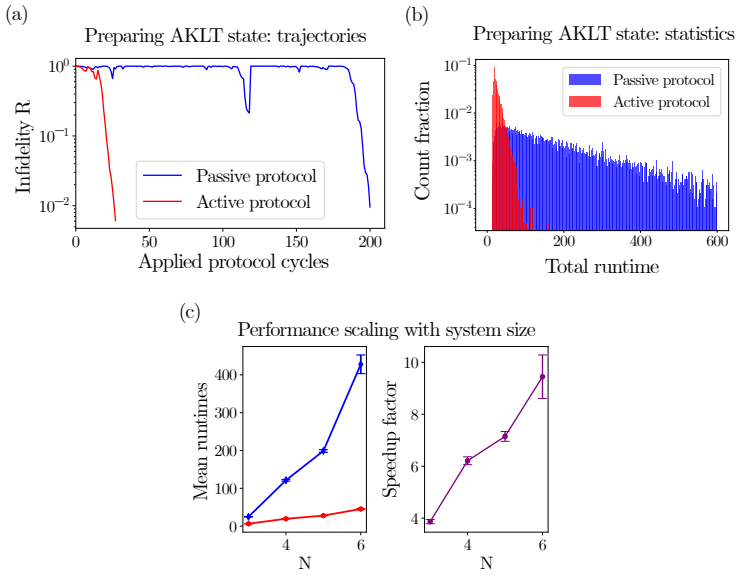

Figure 7.2: (a) Infidelity as a function of the protocol cycle for active and passive protocol runs towards a 5-spin AKLT state. These example runs are characterized by the duration similar to the mean protocol durations of respective protocols ($\approx 199 \pm 4$ for passive and $\approx 28 \pm 0.5$ for active protocol) (b) Histograms of protocol durations t for the preparation of the five-spin AKLT, with accuracy given by infidelity $R < \epsilon = 0.01$. An exponential decaying profile, characteristic of a Poissonian process, can be clearly observed (note the log scale). Note that all recorded runs for an active protocol lasted far less than the mean duration of a passive protocol (200 cycles). Each histogram was compiled from 10^4 simulated runs; the figure is truncated at 600 cycles for better presentation. (c) Scaling of the active protocol's advantage with system size N. A speedup factor tends to increase significantly as the system scales, with factor 9.5 being the estimated speedup at 6 spins. The error bars represent 95% confidence intervals due to sampling error in numerical simulation. 10^4 samples were collected to simulate both protocols $N = 3, 4, 5$, and 10^3 at $N = 6$. Similarly to the above, the infidelity threshold is $\epsilon = 0.01$.

$$H_{\mathrm{AKLT}} = \sum_i H_{i,i+1} = \sum_i \left[\vec{S}_i \cdot \vec{S}_{i+1} + \frac{1}{3} \left(\vec{S}_i \cdot \vec{S}_{i+1} \right)^2 \right], \tag{7.4.2}$$

Here, we assume periodic boundary conditions, implying a single ground

state [207]. The Hamiltonian (7.4.2) is a parent Hamiltonian (as defined in Sec. 7.3.2), where each term $H_{i,i+1}$ has four degenerate ground states $|\phi_a^i\rangle$ with eigenvalue $-2/3$, and 5 excited states $|\theta_b^i\rangle$ with energy $4/3$.

Hereafter, we consider the all-down product state as our starting state. We first design a passive steering protocol for AKLT state preparation, following the parent Hamiltonian construction from Sec. 7.3.2. For simplicity, we restrict ourselves to a less general version of Eq. (7.3.7), and use the following family of coupling operators (cf. Refs. [195, 202]):

$$V(\mathbf{c}, i) = |\phi_4^i\rangle\langle\theta_5^i| + \sum_{\alpha,\beta=1,..4} c_{\alpha\beta}|\phi_\alpha^i\rangle\langle\theta_\beta^i|, \tag{7.4.3}$$

with $c_{\alpha\beta}$ constrained to be an orthogonal matrix, to make sure that the interaction is of constant strength and thus no bias is introduced in the construction. In a passive steering protocol, we will alternate between different values of i, while drawing instances of orthogonal matrices $\mathbf{c}$ at random. For an active feedback strategy to be used on top of this, we propose a greedy policy relative to $C(\rho_s) = R(\rho_s, |\psi_0\rangle)$ to select $\mathbf{c}$. In both passive and active protocol, we assume each coupling to be applied multiple times until one either receives a click, or no-click for an asymptotically long time. Such a repeated application of a single coupling is then counted as a single protocol cycle. We take this approach for a practical purpose because simulating such protocols is more accessible numerically.

Thus simulated, the relative performance of the passive and the active policies (Fig. 7.2) shows a strong advantage of the active policy. In particular, the speedup factor is steadily increasing with system size (Fig. 7.2c), reaching the value of 9.5 for $N = 6$.

7.4.1 Discussion: orthogonality catastrophe and alternative cost functions

The approach defined above harbors a potential challenge. For the greedy procedure to be effective, it should always yield a nonzero bias in favor of a specific $V_s^{(\mathrm{greed})}(\mathbf{p})$ (or a small subset thereof). In other words, the landscape of the cost function $C(\rho_s)$ should not be flat — and some cost functions may yield better landscapes than others. In particular, applying the infidelity measure $R(\rho_s, |\psi_0\rangle)$ is, in general, fundamentally flawed. Indeed, a $(2^N - 1)$-dimensional subspace of states in the N-body Hilbert space is orthogonal to the target state. Let us consider the case when the starting state belongs to that subspace. This situation would in general not change after a single steering cycle with a local coupling $V_s(\mathbf{p})$. For our

purposes, it implies that the infidelity measure R is equal to 1 for a large manifold of states, and there might be no direction of increase that would allow us to choose an appropriate coupling. The most direct example of this can be observed when applying the greedy policy to non-frustrated steering (see Sec. 7.3.1). For simplicity, let us again take the product state of N qubits $|00..0\rangle$ as the target state, the state $|11..1\rangle$ as the starting state, and the couplings $V(i) = \sigma_i^-$ for steering. Only after such steering protocol results in N successful click events, $R(\rho_s, |\psi_0\rangle)$ gains a nonzero value. Thus before $N-1$ clicks, the greedy policy for $R^{(\mathrm{inf})}$ will not be capable of providing a meaningful decision for the next coupling. Strongly enhanced by the system size, this phenomenon is reminiscent of Anderson's orthogonality catastrophe [211].

As a remedy to this deficiency, a "subsystem infidelity" measure can be introduced:

$$R_{\mathcal{S}}(\rho_s, |\psi_0\rangle) = \sum_{\sigma \in \mathcal{S}} \left[1 - \mathrm{tr}\left(\sqrt{\sqrt{\rho_{0,\sigma}}\rho_{s,\sigma}\sqrt{\rho_{0,\sigma}}} \right)^2 \right], \tag{7.4.4}$$

where $\rho_{0,\sigma}$ ($\rho_{s,\sigma}$) is the reduced density matrix of the target state (current state) with respect to subsystem σ. $\mathcal{S}$ is the family of subsystems from which σ are drawn; the choice of $\mathcal{S}$ depends per target state. In the case of the $|11..1\rangle \to |00..0\rangle$ protocol described above, the appropriate $\mathcal{S}$ would be the set of individual spins. Unlike R, such quantity $R_{\mathcal{S}}$ changes every time when a click occurs in this protocol. As a result, the greedy policy with respect to the local $R_{\mathcal{S}}$ would yield the partial-termination protocol of Sec. 7.3.1, significantly boosting the preparation of such a product state.

By continuity with the case of the product state target, such preference for $R_{\mathcal{S}}$ should extend to the weakly-entangled target states, and maybe to some highly-entangled targets. However, we did not see a manifestation of this in the case of our AKLT simulation, where using $R_{\mathcal{S}}$ as a cost function did not yield any improvement compared to R. As a likely explanation for this, the orthogonality catastrophe should become manifest only at large system sizes, where the classical simulation of the protocol is also hindered. However, we expect that some practical target states may still develop a noticeable performance difference between $R_{\mathcal{S}}$ and R, similarly to the case of the product state target. A further study of this question constitutes a promising direction for future work.

7.5 Hilbert-space orienteering map: Quantum State Machine

In this section, we present an orienteering tool that is an alternative to cost-function minimization: mapping out the steering transformations with a Quantum State Machine (QSM) construction. We then illustrate navigation in many-body Hilbert space, employing this machinery to the preparation of the highly entangled W-state of three qubits.

7.5.1 QSM generalities

Every transformation of the system's state, $\Lambda^{(\mathrm{cl})}_{V_s(\mathbf{p})}$ and $\Lambda^{(\mathrm{nc})}_{V_s(\mathbf{p})}$, associated to steering with a specific coupling $V_s(\mathbf{p})$ in a given readout scenario [click or no-click, respectively, see Eqs. (7.2.9) and (7.2.10)], can be represented with a directed graph with complex weights. For this, we notice that every such steering transformation conserves the purity of the state. Therefore, it is convenient to encode transformations $\Lambda^{(\mathrm{cl,\ ncl})}_{V_s}$ in their action on Hilbert space basis states $|\phi_\alpha\rangle$:

$$\Lambda^{(\mathrm{cl,ncl})}_{V_s}(|\phi_\alpha\rangle) = \frac{1}{\sqrt{p^{(\mathrm{cl,ncl})}}} \sum_\beta L^{(\mathrm{cl,ncl})}_{\alpha\beta} |\phi_\beta\rangle \tag{7.5.1}$$

$$L^{(\mathrm{cl})}_{\alpha\beta} = \langle\phi_\beta|\delta t V_s|\phi_\alpha\rangle, \tag{7.5.2}$$

$$L^{(\mathrm{ncl})}_{\alpha\beta} = \langle\phi_\beta|1 - \delta t^2 V_s^\dagger V_s/2|\phi_\alpha\rangle, \tag{7.5.3}$$

where $p^{(\mathrm{cl})}$ ($p^{(\mathrm{ncl})}$) is the probability of a click (non-click) readout upon this steering action. Note that in Eq. (7.5.1), we extended the action of Λ_{V_s} to pure states by a slight abuse of notation compared to Eq. (7.2.9).

The graph representation for steering action $\Lambda^{(\mathrm{cl,ncl})}_{V_s(\mathbf{p})}$, or a steering graph, is then directly obtained from the amplitudes $L^{(\mathrm{cl,ncl})}_{\alpha\beta}$. The vertices in such a graph correspond to the Hilbert space basis states, and the edges describe the steering transformations. The edges are directed and weighted with complex amplitudes, the edge $\alpha \to \beta$ being weighted with amplitude $L^{(\mathrm{cl,\ ncl})}_{\alpha\beta}$ (edges weighted with zero amplitudes are excluded from the graph). Implying this definition, we will use the notation $L^{(\mathrm{cl,\ ncl})}$ for the steering graphs themselves. For basic examples of steering graphs, please refer to Fig. 7.3.

Since the weights $L^{(\mathrm{cl})}_{\alpha\beta}$ are proportional to the matrix elements of coupling operator V_s while $L^{(\mathrm{ncl})}_{\alpha\beta}$ can be expressed via V_s as well, the graph

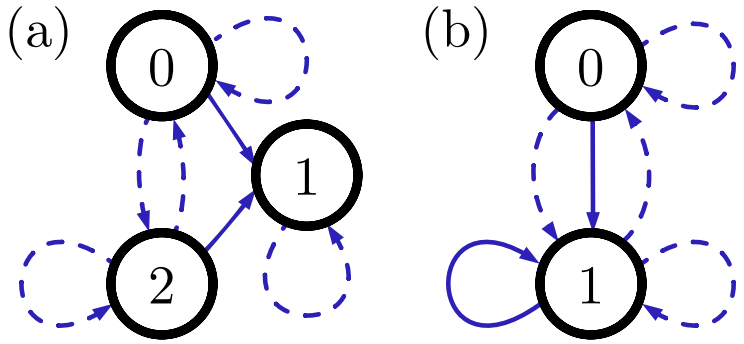

Figure 7.3: Examples of steering graphs (see definition in Sec. 7.5.1): (a) Steering graphs on a 3-level system, corresponding to the coupling $V_s = \gamma(|1\rangle\langle 0| + |1\rangle\langle 2|)$. Graph $L^{(\mathrm{cl})}$ for click action is denoted with solid arrows and the graph $L^{(\mathrm{ncl})}$ for no-click action by dashed arrows (a particular representation of graph coloring). Note that due to the identity operator in Eq. (7.5.3), every vertex is decorated with a self-loop from the $L^{(\mathrm{ncl})}$ graph. To see how the rest of $L^{(\mathrm{ncl})}$ can be deduced from $L^{(\mathrm{cl})}$ (cf. discussion in Sec. 7.5.1), consider the example of $e_{02}^{(ncl)}$ (dashed arrow from state 0 to 2). According to the graphical approach from Sec. 7.5.1, one is to follow edge $e_{01}^{(cl)}$ (solid arrow from 0 to 1) forward and then $e_{21}^{(cl)}$ (solid arrow from 2 to 1) backward - and thus manages to travel from state 0 to 2, in correspondence to $e_{02}^{(ncl)}$. (b) Steering graphs on a 2-level system, as defined by the coupling $V_s = \gamma(|1\rangle\langle 1| + |1\rangle\langle 0|)$. Following the same rule as above, inter-vertex edges of $L^{(\mathrm{ncl})}$ can be deduced from $L^{(\mathrm{cl})}$. For example, by following the edge $e_{11}^{(\mathrm{cl})}$ forward and then the edge $e_{01}^{(\mathrm{cl})}$ backward, one performs a transition from state 1 to state 0, thus reproducing the edge $e_{10}^{(\mathrm{ncl})}$ from $L^{(\mathrm{ncl})}$.

$L^{(\mathrm{ncl})}$ for the no-click action can be inferred entirely from the graph $L^{(\mathrm{cl})}$ for the click action. In particular, due to the term $\propto V_s^\dagger V_s$, graph $L^{(\mathrm{ncl})}$ contains an edge $e_{ij}^{(\mathrm{ncl})}$ from vertex v_i to v_j, if a graph $L^{(\mathrm{cl})}$ contains edges $e_{ik}^{(\mathrm{cl})}$ and $e_{jk}^{(\mathrm{cl})}$ (see Fig. 7.3). Heuristically, to yield a $L^{(\mathrm{ncl})}$-edge, one has to first follow a $L^{(\mathrm{cl})}$-edge forward, and then another $L^{(\mathrm{cl})}$-edge backward. Furthermore, due to the additional identity operator term in Eq. (7.5.3), any graph for the no-click steering action will also include self-loops on each vertex.

The steering graphs introduced above can now be used to create a Quantum State Machine. For n_V couplings $V_s(\mathbf{p})$ in the steering kit, there exist $2n_V$ graphs corresponding to steering maps $\Lambda_{V_s(\mathbf{p})}^{(\mathrm{cl},\,\mathrm{ncl})}$, because of the two possible measurement outcomes for each of the couplings. The QSM

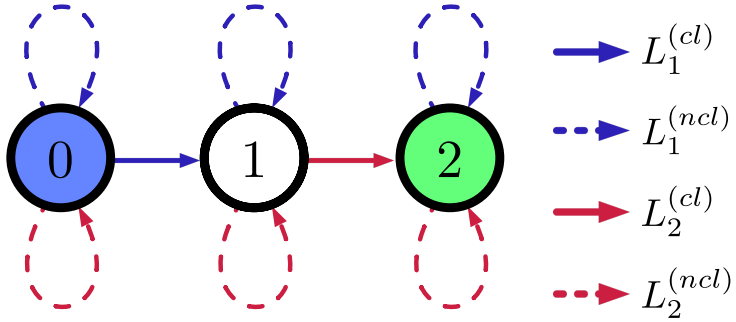

Figure 7.4: A basic example of the QSM multigraph, describing the steering kit for a three-state system. The steering options are represented by the coupling operators $V_1 = \gamma_1|1\rangle\langle 0|$ and $V_2 = \gamma_2|2\rangle\langle 1|$. The starting state is 0, marked in blue, and the target state is 2, marked in green. The optimal coordination policy of the two steering operations is straightforward: one needs to first repeatedly apply the V_1-steering until a click is obtained, and then the V_2-steering until a click is obtained. Compared to the passive steering which iterates between V_1 and V_2 regardless of measurement outcomes, this directly yields a 2-fold speedup in the average performance.

for the steering protocol is then obtained as a collection of these graphs. It can be represented as a colored multigraph, where each steering graph is represented as a single-color subgraph (Fig. 7.4). Consequently, in a QSM multigraph there may be multiple edges going from any vertex α into any other vertex β (making it a multigraph rather than a simple graph), but at most one such edge for each color.

Let us now consider our original task of finding the accelerated navigation protocol. To make use of the QSM construction in this context, we will restrict our consideration to bases $\{|\phi_\beta\rangle\}$ where one of the basis states is the target state $|\psi_0\rangle$ itself. In such a case, state $|\psi_0\rangle$ corresponds to a marked vertex in the graph, and the goal of the steering protocol becomes to drive the system state to that vertex. The goal of optimizing this protocol may then look similar to a known problem of finding the shortest path to the marked vertex on a weighted graph. This problem is standard in graph theory and can be solved as such. Can such a solution be used to design the navigation protocol?

As we will see in Sec. 7.5.2, this analogy is not complete, since the quantum evolution on the graph goes beyond the simple path-on-the-graph

picture. This aspect creates an obstacle to directly applying the graph exploration algorithms to facilitate our protocol speed-up. Fortunately, in some cases, this difficulty can be properly accounted for, as we will see in Sec. 7.5.3. In those cases, the "semi-classical heuristics" of graph exploration may indeed be applied. Finally, in Sec. 7.5.4, we will apply this approach to the preparation of the W-state, with a factor 12.5 improvement compared to the passive protocol using the active navigation protocol.

7.5.2 Quantum subgraphs in a QSM

Let us now compare our QSM navigation task to the standard problem of graph exploration. Our goal is to identify the differences between the two, which prevent us from applying the graph exploration techniques directly to QSM navigation. First of all, the state of the system in graph exploration is at all times represented by a single vertex. The system in a QSM, on the other hand, is generally represented by a superposition over multiple vertices. Furthermore, in graph exploration, the state is modified by following one of the edges. A steering action in a QSM, in contrast, corresponds to a whole collection of edges – i.e., a single steering graph in the QSM multigraph.

Some steering graphs may induce quantum effects, such as superposition and interference. For instance, the steering action whose graph contains two outgoing edges from a given vertex (e.g., vertex 0 for graph $L_1^{(\mathrm{cl})}$ in Fig. 7.5a), can create a nontrivial quantum superposition. If a state is given by a superposition of multiple vertex states, it may further undergo quantum interference. In particular, this can be facilitated by a steering action whose graph contains a vertex with two incoming edges (e.g., vertex 4 for graph $L_2^{(\mathrm{cl})}$ in Fig. 7.5a). In general, a notion of "superposition subgraphs" and "interference subgraphs" of a steering graph can be defined:

1. Superposition subgraph is a subgraph of a steering graph span by multiple (more than one) edges outgoing from a single vertex.

2. Interference subgraph is a subgraph of a steering graph span by multiple edges incoming to a single vertex.

Collectively, we will refer to such interference and superposition subgraphs of a single steering graph as its quantum subgraphs. If the quantum subgraphs are absent in the QSM, we will refer to it as a classical QSM. In other words, in a classical QSM, each vertex has at most one outgoing and at most one incoming edge of any given color.

If a QSM is classical, optimization of the navigation protocol can essentially be reduced to classical graph exploration. For a simple example of a classical QSM and the way to optimize the respective state preparation, consider the 3-level steering actions described in Fig. 7.4. Note that optimization of the classical QSM also applies to the case when the starting state is a superposition of multiple vertex states. If the steering operations contain no quantum subgraphs, the quantum superposition is equivalent to a probabilistic mixture for the sake of the protocol optimization, and the optimal navigation pattern can be extracted accordingly.

As the form of the steering graph depends on the choice of basis, it is conceivable that the number of quantum subgraphs in such a graph in some cases can be reduced by changing the basis (compare Fig. 7.5a and b). However, using a change of basis to remove all the quantum subgraphs in an arbitrary QSM is generally impossible (see Fig. 7.5).

7.5.3 Coarse-grained QSM. Semiclassical heuristic for navigation

We now focus on the steering protocols whose QSM cannot be made classical via a basis transformation. In such a case, it may still be possible to optimize it via a classical graph exploration heuristic. For that, we propose to coarse-grain the QSM by grouping subsets of its vertices into single block-vertices. The coarse-grained QSM would consist of graphs drawn between such block-vertices. The block-vertex containing the target vertex can be considered as the target block-vertex.

An inter-block edge between two block-vertices is drawn, if the original QSM has at least one edge connecting the vertices inside the respective block-vertices. For the coarse-graining to be useful for our purposes, it should be done in such a way that all of the resulting QSM graphs have a classical structure. Namely, the coarse-grained graph should not have quantum subgraphs, e.g. realizing superposition or interference between the block-vertices (in analogy to Sec. 7.5.2). To satisfy this requirement, the following rule for vertex grouping can be employed (cf. Fig. 7.6): *if two edges of the same color are simultaneously coming in or out of a given vertex, the two vertices at the other ends of these edges should be grouped within one effective block-vertex.* This rule manifestly yields basis-dependent groupings, since the very presence of quantum subgraphs in a QSM is basis-dependent. Thus, a smart choice of the basis may allow for an efficient and simpler coarse-grained graph. Designing a general explicit algorithm for finding the optimum basis for an arbitrary QSM is a highly non-trivial task. Heuristically speaking, a convenient choice of the

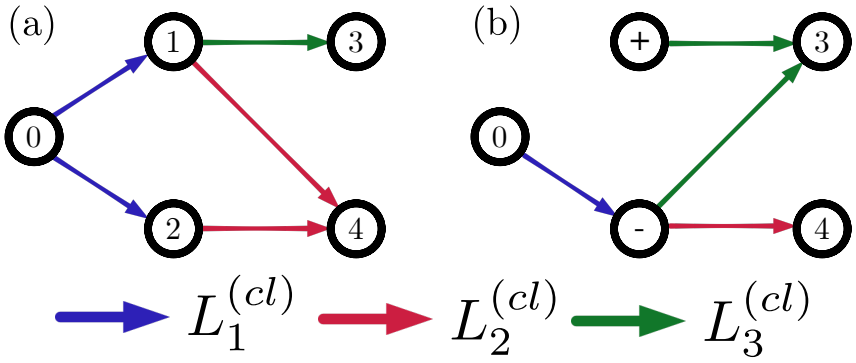

Figure 7.5: Possible configurations of quantum subgraphs in a QSM, exemplified by the 5-vertex subgraph of an example QSM. (a) The click-action graphs for the three coupling operators $V_{1,2,3}$ that form the steering kit. The operators have the form $V_1 = \gamma_1(|1\rangle - |2\rangle)\langle 0|$, $V_2 = \gamma_2|4\rangle(\langle 1| - \langle 2|)$, $V_3 = \gamma_3|3\rangle\langle 1|$. The graphs for the no-click actions are not shown, as their form can be deduced from the graphs for click actions. In the present basis, the V_1-click is manifest as a superposition, the V_2-click – as an interference, and the V_3-click corresponds to a semiclassical evolution. (b) Quantum State Machine for the steering kit from the previous panel, depicted in a different basis. The basis transformation is $|\pm\rangle = (|1\rangle \pm |2\rangle)/\sqrt{2}$. In this case, the basis transformation removes the quantum elements in the $L_{1,2}^{(cl)}$ graphs, however, it turns $L_3^{(cl)}$ into an interference element. Note that there is no basis transformation that would turn such a QSM into a classical one. This statement follows from the uniqueness of the Jordan canonical form for operators V_2 and V_3.

basis should be the one that results in the minimum number of quantum subgraphs in a QSM before coarse-graining.

For the coarse-grained graph to be effectively classical, we desire to ignore details of the system evolution inside the subspace of a given block-vertex. Specifically, we aim to view every block-vertex as an effective single state of the system and assume that every edge allows transporting the system between such block-vertex states with no obstacles. If this was directly possible, and since the coarse-grained QSM by definition contains no quantum subgraphs, optimization of its exploration would have become a classical task. However, such an approximation scheme needs more careful justification. Every block fundamentally corresponds to a Hilbert subspace, and an inter-block edge is given by a $N_1 \times N_2$ matrix of coefficients (where N_1 and N_2 are the internal dimensionalities of the

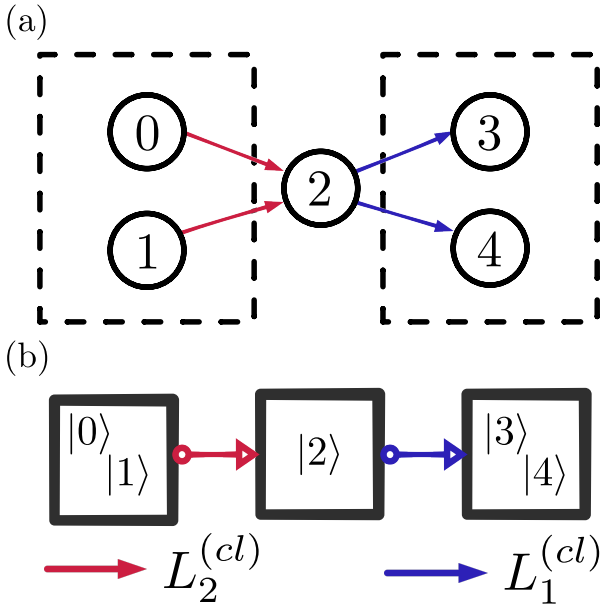

Figure 7.6: Semiclassical coarse-graining applied to a QSM. (a) A 5-state part of a QSM with two quantum subgraphs: interference subgraph realized by $L_2^{(cl)}$ and a superposition subgraph realized by $L_1^{(cl)}$. Since pairs of states $\{|0\rangle,\ |1\rangle\}$ and $\{|3\rangle,\ |4\rangle\}$ fall under conditions described in Sec. 7.5.3, these are to be grouped together in a coarse-grained QSM. (b) Simplified depiction of a coarse-grained QSM, obtained from (a).

linked blocks). Characterizing these effectively with single amplitudes may lead to erroneous navigation policies. In particular, one state internal to a block-vertex might be untouched by an inter-block edge, i.e., it only yields zero matrix elements in a matrix characterizing the edge. If the edge is outgoing, a system initialized in the said state would not be able to escape the block-vertex using that edge alone (see Fig. 7.7). This is in direct conflict with characterizing blocks and inter-block edges with single amplitudes. For an incoming edge, a similar problem may arise: some states inside a block-vertex might not get populated when that edge is activated. This may become detrimental for the navigation protocol based on a coarse-grained QSM, especially if the unavailable state in question is the final target of the protocol.

Such difficulties may be overcome, if some of the couplings given in

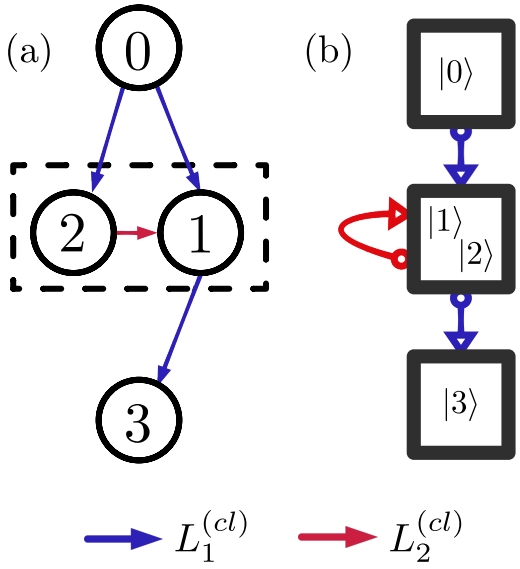

Figure 7.7: Illustration of ancillary couplings in the context of QSM coarse-graining (a) A 4-state part of a QSM that is subject to coarse-graining, featuring non-trivial actions by couplings denoted as V_1 and V_2. States $|1\rangle$ and $|2\rangle$ are to be grouped together since they are both targets in a superposition subgraph span by edges e_{02} and e_{01} of a steering graph $L_1^{(\mathrm{cl})}$. (b) The coarse-grained version of the QSM from (a). The block $\{|1\rangle, |2\rangle\}$ is connected to state $|3\rangle$ through an outgoing edge of $L_1^{(\mathrm{cl})}$. However, from microscopic point of view exemplified in (a), no population can be transferred from state $|2\rangle$ to $|3\rangle$ unless the click action $\Lambda_2^{(\mathrm{cl})}$ is realized first. Therefore, including and applying V_2 as an ancillary coupling is required for a valid semiclassical coarse-graining of this QSM.

a QSM allow for an internal mixing of the subspace (represented by a self-loop on the block-vertex in the respective $L^{(\mathrm{cl})}$-graph). Applying such a coupling in the protocol would allow one to make the block-vertex accessible to all the edges that are connected to it (see Fig. 7.7), via a sufficient number of clicks. In the scenarios described above, where additional couplings are needed to turn a block-vertex into an effective single vertex, we will refer to such couplings as ancillary couplings. Note that given a steering kit, there is no guarantee that the ancillary couplings needed for exploration of every block-vertex, are available. For simplicity, in this chapter, we restrict our further consideration to the coarse-grained

QSMs, where the ancillary couplings happen to be present wherever needed. Every block-vertex can then be made accessible to the outgoing edges, and the target state is ensured to be reachable once the target block is reached. In this case, we consider the coarse-grained QSM as effectively semiclassical.

To design an active steering policy within the coarse-grained approach, we note that the navigation protocol has the following structure. The system state can be transported between block-vertices, and eventually steered to the target block-vertex. After that, either the target state is reached already (one can obtain this information from the simulated copy of the system), or it can be reached after applying ancillary couplings on the target block-vertex. The cost of the protocol can now be broken into two parts. The first is the cost of exploring the coarse-grained graph using the inter-vertex edges. The second is the dwell time inside the block-vertices, which is spent applying the ancillary couplings. If we could find the route through the graph that minimizes the combination of these two components, it would solve our optimization problem exactly. However, because of the presence of the degrees of freedom that are internal to the block-vertices, the coarse-grained geometrical information does not allow for such a precise solution. In other words, both the inter-vertex travel time and the block-vertex dwell time depend on the microscopic details of the evolution.

Instead of studying such quantum-mechanical microscopics, we propose a semiclassical approximation to this calculation. Specifically, we assign every inter-block edge a characteristic traversal time, and every block-vertex a characteristic dwell time. For this, we use the matrices for click transitions between blocks i and j (the case of ancillary couplings given by $i = j$). Let us loosely denote these as $L^{(\mathrm{cl})}_{i,\alpha;j,\beta}$, implying that only matrix elements with states from blocks i and j are included. In that case, the effective transition amplitude between blocks i and j can be defined as operator norm $L^{(\mathrm{cl})}_{i,j} = \|L^{(\mathrm{cl})}_{i,\alpha;j,\beta}\|$, and characteristic traversal (dwell if $i = j$) time $\tau_{i,j} = (L^{(\mathrm{cl})}_{i,j})^{-2}$. Note that this reduces to the average traversal time for the case of a genuinely classical graph, with an amplitude $\gamma\delta t$ connecting two states implying duration of $\tau = \frac{1}{\gamma^2\delta t^2}$ for traversal (cf. Sec. 7.2.2).

With characteristic times $\tau_{i,j}$ assigned, the time-cost of following a specific path through this graph can be estimated as a combined characteristic time of all the edges and vertices crossed along the way. The desired path will be the one that optimizes this expected time. On the one hand, this

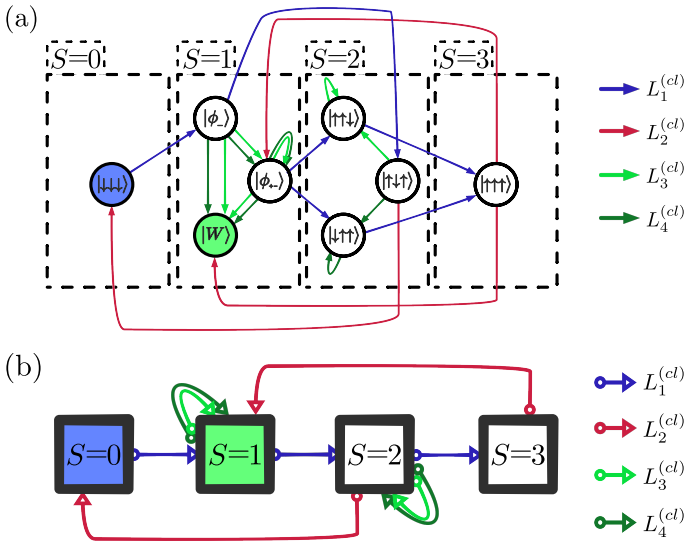

Figure 7.8: Measurement-driven navigation towards the 3-qubit W-state: QSM representation. (a) Steering with couplings Eqs. (7.5.5)-(7.5.8). The vertices in the single-excitation subspace are given by states $|W\rangle$, $|\phi_{-}\rangle \equiv \frac{1}{\sqrt{2}}(|100\rangle - |001\rangle)$, and $|\phi_{+-}\rangle \equiv \frac{1}{\sqrt{6}}(|100\rangle - 2|010\rangle + |001\rangle)$. (b) The coarse-grained version of the above QSM. The vertices are labeled by the excitation number. From perspective of Sec. 7.5.3, couplings 2 and 3 play the ancillary role. Indeed, those couplings mix the internal structure of the block-vertices, allowing one to eventually steer the state to the target $|W\rangle$.

may result in a different navigation protocol compared to what is optimal from the complete quantum-mechanical analysis. On the other hand, such a first-principles analysis is prohibitively hard, and we expect that our semiclassically derived protocol will still be considerably quicker than its completely passive version. One example of such an improved protocol is given below.

7.5.4 W-state preparation

To illustrate the principles of the QSM framework, we consider the coarse-graining approach to the navigation of a 3-qubit state from a trivial $|000\rangle$

state to a so-called W-state [212] that has the following form:

$$W = \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle). \tag{7.5.4}$$

For the steering kit, we choose the following family of couplings (assuming labels A, B, and C for the qubits):

$$V_1 = \sigma_A^+ - \sigma_C^+, \tag{7.5.5}$$
$$V_2 = \sigma_A^- \sigma_C^-, \tag{7.5.6}$$
$$V_3 = \sigma_A^- \sigma_B^+ - P_A^0 P_B^1, \tag{7.5.7}$$
$$V_4 = \sigma_B^+ \sigma_C^- - P_B^1 P_C^0. \tag{7.5.8}$$

Here, $\sigma^{\pm} = \frac{1}{2}(\sigma^x \pm i\sigma^y)$ and $P^a = |a\rangle\langle a|$, $a = 0, 1$. A passive version of the protocol would amount to blindly alternating between the steering actions with different V_i, which does yield the target state, given that the steering is applied a sufficient number of times (see Fig. 7.8.b).

To design a feedback policy, we now consider a QSM representation of the steering kit. It is shown in Fig. 7.8a. Note that this QSM has multiple quantum subgraphs. Therefore, to employ a feedback policy, it should be subjected to the subspace-clustering coarse-graining technique, as outlined in Sec. 7.5.3. It proves useful to cluster the Hilbert space by the total excitation number, which results in a semiclassical QSM, as desired (Fig. 7.8b). Given the starting state of the evolution, it is then straightforward to design the policy that leads to the target state:

1. Repeat V_1-steering until a click is obtained;
2. Repeat V_3-steering until the target state is reached (with fidelity error below ϵ).

This protocol moves the state of the system from the zero excitation state to the single-excitation subspace (part 1) and then takes the system to the W-state in that subspace (part 2). Note that this employs only two couplings out of four, a simplification that is only possible with an active steering protocol. In a passive protocol, as all couplings are employed cyclically, multiple V_1-clicks may accidentally occur before the target state is reached (Fig. 7.9a), and, therefore, other couplings are needed to reduce the excitation number back to 1. In Fig. 7.9b, the performance histogram is given for a large number of numerical trials for the active and passive protocols. Both are run with $\delta t = 0.1$ for the target fidelity error $\epsilon = 0.01$. We are primarily interested in the average runtimes of the active and the passive protocol. These are $N_{\text{act}} \simeq 365 \pm 3$ and $N_{\text{pas}} \simeq 4600 \pm 100$ cycles of the protocol, respectively, yielding a speedup factor of around 12.5.

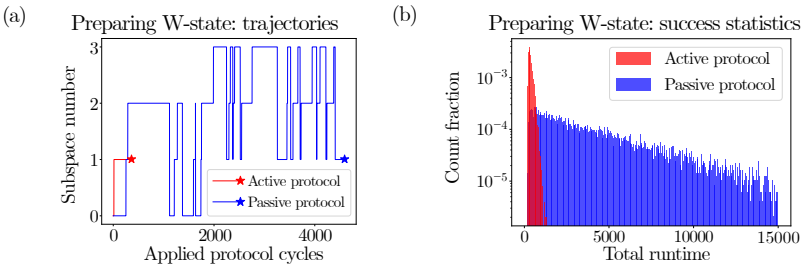

Figure 7.9: Performance of the active navigation protocol based on the QSM representation of steering towards the W-state, as described in the text. (a) Typical trajectories of the passive and active protocols, in terms of the excitation number sector that is occupied by the system state. Displayed are trajectories that yield the runtimes approximately equal to average runtimes of 365 (active) and 4600 (passive). The first V_1-click in the displayed run of the active protocol occurs as early as the 14th cycle, which is not legible from the plot. (b) The histogram over the protocol runtimes, required for a passive and an active protocol to achieve r=0.01 infidelity to the target state. Similarly to the Fig. 7.2 for AKLT target state, note the clear advantage for each recorded run of the active protocol compared to the mean duration $\simeq 4.6 \cdot 10^3$ of the passive protocol. Each histogram was obtained from 10^4 numerical simulations, and truncated at 15000 cycles for better presentation.

7.6 Discussion and conclusions

In this work, we have put forward the concept of measurement-driven active-decision steering of quantum states. We have developed steering protocols in which the measurement readouts are used to adjust the measurement protocol on-the-go, yielding significant acceleration of state preparation, with improved fidelity, compared with passive steering. The possibility of exploiting the readouts explored here is the great advantage of measurement-based steering over drive-and-dissipation (largely equivalent to "blind" steering) state preparation. While our approach has sweeping applicability, here we have chosen to focus on active measurement-driven steering as applied to the most challenging case of many-body quantum systems with entangled target states.

To satisfy physical (locality) constraints on system-detector couplings, we have proposed a scheme, based on parent Hamiltonian construction, for identifying feasible couplings. Employing such couplings, we have developed and analyzed Hilbert-space-orientation techniques for measurement-

driven steering. A central ingredient here has been to develop feedback policies based on detector readouts. One such Hilbert-space path-finding technique is based on a cost function, evaluating the running fidelity to the target state. We have shown a substantial (up to 9.5-fold) speedup of steering, employing this approach for preparation of the ground state of the AKLT model. A second protocol comprises mapping out the available measurement actions onto a Quantum State Machine (QSM), using a coarse-grained version of the corresponding graphs in Hilbert space. We have given an example of an entangled state preparation which shows acceleration by a factor of 12.5 compared to passive steering.

While we have limited ourselves here to a few examples, our schemes are of general applicability. They open the door to the design of efficient and high-quality state engineering, adiabatic state manipulation, and, possibly, quantum information processing. Moreover, steering protocols are subject to errors, both "static" (choice of steering parameters) and "dynamics" (noise). Active decision-making steering may be designed to reduce the effect of such errors. One may envision a host of directions to generalize and develop these ideas. For example, the greedy minimization of our cost function may be further improved by finding other metrics of local "steepest decent." Further, one may systematically investigate less local (less greedy) optimization of the cost function, e.g., looking n steps ahead. Another potential advantage of our protocols relies on the following observation: in the context of passive steering, one imposes constraints concerning locality (e.g., how many spins can be coupled to a local detector), and certain types of coupling terms. Given such constraints, not all target states are reachable. The introduction of active steering may overcome this handicap of target-state accessibility.

One may combine the dynamics incorporated here with the inherent unitary evolution of the system at hand (due to a system-only Hamiltonian). Consider the context of passive (blind) measurement-induced steering, which, in the continuum time limit, leads to Lindbladian dynamics. Then, the addition of Hamiltonian dynamics enriches the variability of steering, allowing, for example, to obtain mixed states by design [213]. It is intriguing to investigate how the addition of Hamiltonian dynamics extends or improves active steering, thus marrying the frameworks of closed-loop quantum control for Hamiltonian-based state preparation and active-decision measurement-based steering.

Further extensions of our approach include applications of QSM protocols to larger and more complex systems, going beyond a three-qubit setup. Optimizing such protocols may involve automatization of the creation and analysis of QSMs, e.g., for finding an optimal basis automatically, in

similarity with quantum annealing, but now at the level of measurement operators. Finally, one may envision using machine learning to find more optimized navigation protocols (see [214, 215] for related work in the context of Hamiltonian feedback and open-loop control). Given the delayed-reward setting at hand, a reinforcement learning strategy such as Q-learning [216] or SARSA [217] might be the most appropriate choice.

Bibliography

[1] M.A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2010).

[2] D. DiVincenzo, *The Physical Implementation of Quantum Computation*, Fortschr. Phys. **48**, 771 (2000).

[3] C. Nayak, S.H. Simon, A. Stern, M. Freedman, and S.D. Sarma, *Non-Abelian Anyons and Topological Quantum Computation*, Rev. Mod. Phys. **80**, 1083 (2008).

[4] A.Y. Kitaev, *Fault-tolerant quantum computation by anyons*, Annals Phys. **303**, 2 (2003).

[5] M.H. Freedman, *Quantum Computation and the localization of Modular Functors*, arXiv:quant-ph/0003128.

[6] J. Alicea, *New directions in the pursuit of Majorana fermions in solid state systems* , Rep. Prog. Phys. **75**, 076501 (2012).

[7] D. Culcer, A. C. Keser, Y. Li, and G. Tkachov, *Transport in two-dimensional topological materials: recent developments in experiment and theory*, 2D Mater. **7**, 022007 (2020).

[8] A. Vuik, B. Nijholt, A. R. Akhmerov, and M. Wimmer, *Reproducing topological properties with quasi-Majorana states*, SciPost Phys. **7**, 061 (2019).

[9] Y. Huang, F. Setiawan, and J.D. Sau, *Disorder-induced half-integer quantized conductance plateau in quantum anomalous Hall insulator-superconductor structures*, Phys. Rev. B **97**, 100501 (2018).

[10] M. Kayyalha, D. Xiao, R. Zhang, J. Shin, J. Jiang, F. Wang, Y. F. Zhao, R. Xiao, L. Zhang, K. M. Fijalkowski, P. Mandal, M. Winnerlein, C. Gould, Q. Li, L. W. Molenkamp, M. H. W. Chan, N. Samarth, and C. Z. Chang, *Absence of evidence for chiral Majorana modes in quantum anomalous Hall-superconductor devices*, Science **367**, 64 (2020).

[11] J. Preskill, *Quantum Computing in the NISQ era and beyond*, Quantum **2**, 79 (2018).

[12] S. Lloyd, *Universal Quantum Simulators*, Science **273**, 1073 (1996).

[13] I. Buluta and F. Nori, *Quantum Simulators*, Science **326**, 108 (2009).

[14] B. P. Lanyon, C. Hempel, D. Nigg, M. Müller, R. Gerritsma, F. Zähringer, P. Schindler, J. T. Barreiro, M. Rambach, G. Kirchmair, M. Hennrich, P. Zoller, R. Blatt, and C. F. Roos, *Universal digital quantum simulation with trapped ions*, Science **334**, 57 (2011).

[15] N. Schuch and F. Verstraete, *Computational complexity of interacting electrons and fundamental limitations of density functional theory*, Nat. Phys. **5**, 732 (2009).

[16] A.Y. Kitaev, A. Shen, M.N.Vyalyi, *Classical and Quantum Computation*, Amer. Math. Soc., 2002.

[17] M. Troyer, U.-J. Wiese, *Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations*, Phys. Rev. Lett. **94**, 170201 (2005).

[18] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, *Quantum Algorithms for Quantum Chemistry and Quantum Materials Science*, Chem. Rev. **120**, 12685 (2020).

[19] X.-G. Wen, *Quantum Field Theory of Many-Body Systems: From the Origin of Sound to an Origin of Light and Electrons* (Oxford University Press, 2007).

[20] B.A. Bernevig and T. L. Hughes, *Topological Insulators and Topological Superconductors* (Princeton University Press, 2013).

[21] C.W.J. Beenakker, *Search for Majorana fermions in superconductors*, Annu. Rev. Con. Mat. Phys. **4**, 113 (2013).

[22] M.H. Freedman, M. Larsen, and Z. Wang, *A Modular Functor Which is Universal for Quantum Computation*, Commun. Math. Phys. **227**, 605 (2002).

[23] A.P. Mackenzie and Y. Maeno, *p-wave superconductivity*, Physica B: Cond. Mat. **280**, 148 (2000).

[24] M. Sato, Y. Ando, *Topological superconductors: a review*, Rep. Prog. Phys. **80**, 076501 (2017).

[25] L. Fu and C.L. Kane, *Superconducting Proximity Effect and Majorana Fermions at the Surface of a Topological Insulator*, Phys. Rev. Lett. **100**, 096407 (2008).

[26] M.R. Norman, *The Challenge of Unconventional Superconductivity*, Science **332**, 196 (2011).

[27] Q. L. He, L. Pan, A. L. Stern, E. Burks, X. Che, G. Yin, J. Wang, B. Lian, Q. Zhou, E. S. Choi, K. Murata, X. Kou, T. Nie, Q. Shao, Y. Fan, S.-C. Zhang, K. Liu, J. Xia, and K. L. Wang, *Chiral Majorana fermion modes in a quantum anomalous Hall insulator–superconductor structure*, Science **357**, 294 (2017).

[28] J. Shen, J. Lyu, J. Z. Gao, Y.-M. Xie, C.-Z. Chen, C. Cho, O. Atanov, Z. Chen, K. Liu, Y. J. Hu, K. Y. Yip, S. K. Goh, Q. L. He, L. Pan, K. L. Wang, K. T. Law, and R. Lortz, *Spectroscopic fingerprint of chiral Majorana modes at the edge of a quantum anomalous Hall insulator/superconductor heterostructure*, PNAS **117**, 238 (2020).

[29] W. Ji and X.-G. Wen, *Conductance Plateau without 1D Chiral Majorana Fermions*, Phys. Rev. Lett. **120**, 107002 (2018).

[30] A.A. Abrikosov, *The magnetic properties of superconducting alloys*, J. Phys. Chem. Solids **2**, 199 (1957).

[31] Y. Aharonov and D. Bohm, *Significance of Electromagnetic Potentials in the Quantum Theory*, Phys. Rev. **115**, 485 (1959).

[32] D. A. Ivanov, *Non-abelian statistics of half-quantum vortices in p-wave superconductors*, Phys. Rev. Lett. **86**, 268 (2001).

[33] B. Lian, X.-Q. Sun, A. Vaezi, X.-L. Qi, and S.-C. Zhang, *Topological quantum computation based on chiral Majorana fermions*, PNAS **115**, 10938 (2018).

[34] E.H. Hall, *On a New Action of the Magnet on Electric Currents*, Am. J. Math. **2**, 287 (1879).

[35] K. v. Klitzing, G. Dorda, and M. Pepper, *New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance*, Phys. Rev. Lett. **45**, 494 (1980).

[36] D. C. Tsui, H. L. Stormer, and A. C. Gossard, *Two-Dimensional Magnetotransport in the Extreme Quantum Limit*, Phys. Rev. Lett. **48**, 1559 (1982).

[37] B.I. Halperin, *Quantized Hall conductance, current-carrying edge states, and the existence of extended states in a two-dimensional disordered potential*, Phys. Rev. B **25**, 2185 (1982).

[38] S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, 1995).

[39] R.B. Laughlin, *Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations*, Phys. Rev. Lett. **50**, 1395 (1983).

[40] D.J. Clarke, J. Alicea, and K. Shtengel, *Exotic non-Abelian anyons from conventional fractional quantum Hall states*, Nat. Comm. **4**, 1348 (2013).

[41] W.J. Cook, W.H. Cunningham, W.R. Pulleyblank, and A. Schrijver, *Combinatorial Optimization, First Edition* (John Wiley & Sons, 1998).

[42] F. Barahona, *On the computational complexity of Ising spin glass models*, J. Phys. A: Math. Gen. **15**, 3241 (1982).

[43] J. Kempe, A. Kitaev, and O. Regev, *The Complexity of the Local Hamiltonian Problem*, SIAM J. Comput. **35**, 1070 (2006).

[44] P. Wocjan and S. Zhang, *Several natural BQP-Complete problems*, arXiv:quant-ph/0606179.

[45] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, *Quantum Computation by Adiabatic Evolution*, arXiv:quant-ph/0001106.

[46] A.Y. Kitaev, *Quantum measurements and the Abelian Stabilizer Problem*, arXiv:quant-ph/9511026.

[47] J.R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, *The theory of variational hybrid quantum-classical algorithms*, New J. Phys. **18**, 023023 (2016).

[48] V. Verteletskyi, T.-C. Yen, and A. F. Izmaylov, *Measurement Optimization in the Variational Quantum Eigensolver Using a Minimum Clique Cover*, J. Chem. Phys. **152**, 124114 (2020).

[49] X. Bonet-Monroig, R. Babbush, and T.E. O'Brien, *Nearly Optimal Measurement Scheduling for Partial Tomography of Quantum States*, Phys. Rev. X **10**, 031064 (2020).

[50] J. Cotler and F. Wilczek, *Quantum Overlapping Tomography*, Phys. Rev. Lett. **124**, 100401 (2020).

[51] J. M. Kübler, A. Arrasmith, L. Cincio, and P. J. Coles, *An Adaptive Optimizer for Measurement-Frugal Variational Algorithms*, Quantum **4**, 263 (2020).

[52] A. Arrasmith, L. Cincio, R.D. Somma, and P. J. Coles, *Operator Sampling for Shot-frugal Optimization in Variational Algorithms*, arXiv:2004.06252.

[53] K.J. Sung, J. Yao, M. P. Harrigan, N. C. Rubin, Z. Jiang, L. Lin, R. Babbush, and J. R. McClean, *Using models to improve optimizers for variational quantum algorithms*, Quantum Sci. Technol. **5**, 044008 (2020).

[54] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. Love, and A. Aspuru-Guzik, *Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz*, Quantum Sci. Technol. **4**, 014008 (2018).

[55] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, *An adaptive variational algorithm for exact molecular simulations on a quantum computer*, Nat. Comm. **10**, 3007 (2019).

[56] D. Wecker, M. B. Hastings, and M. Troyer, *Towards Practical Quantum Variational Algorithms*, Phys. Rev. A **92**, 042303 (2015).

[57] K. A. Brueckner, *Many-Body Problem for Strongly Interacting Particles. II. Linked Cluster Expansion*, Phys. Rev. **100**, 36 (1955).

[58] H. F. Trotter, *On the product of semi-groups of operators*, Proc. Ams. Math. Soc. **10**, 545 (1959).

[59] M. Suzuki, *General theory of fractal path integrals with applications to many-body theories and statistical physics*, J. Math. Phys. **32**, 400 (1991).

[60] J. D. Whitfield, J. Biamonte, A. Aspuru-Guzik *Simulation of Electronic Structure Hamiltonians Using Quantum Computers*, Mol. Phys. **109**, 735 (2011).

[61] F. Verstraete, M. M. Wolf, and J. I. Cirac, *Quantum computation and quantum-state engineering driven by dissipation*, Nat. Phys. **5**, 633 (2009).

[62] P.O. Boykin, T. Mor, V. Roychowdhury, F. Vatan, and R. Vrijen, *Algorithmic cooling and scalable NMR quantum computers*, PNAS **99**, 3388 (2002).

[63] D. Kielpinski, B. E. King, C. J. Myatt, C. A. Sackett, Q. A. Turchette, W. M. Itano, C. Monroe, D. J. Wineland, and W. H. Zurek, *Sympathetic cooling of trapped ions for quantum logic*, Phys. Rev. A **61**, 032310 (2000).

[64] M. Popp, J. J. Garcia-Ripoll, K. G. H. Vollbrecht, and J. I. Cirac, *Ground-state cooling of atoms in optical lattices*, Phys. Rev. A **74**, 013622 (2006).

[65] M. Metcalf, J. E. Moussa, W. A. de Jong, and M. Sarovar, *Engineered thermalization and cooling of quantum many-body systems*, Phys. Rev. Research **2**, 023214 (2020).

[66] M. Raghunandan, F. Wolf, C. Ospelkaus, P. O. Schmidt, and H. Weimer, *Initialization of quantum simulators by sympathetic cooling*, Sci. Adv. **6**, eaaw9268 (2020).

[67] Y. Aharonoov, D. Z. Albert, and L. Vaidman, *How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100*, Phys. Rev. Lett. **60**, 1351 (1988).

[68] B. Tamir and E. Cohen, *Introduction to Weak Measurements and Weak Values*, Quanta **2**, 7 (2013).

[69] J. Zhang, Y.-X. Liu, R.-B. Wu, K. Jacobs, F. Nori, *Quantum feedback: theory, experiments, and applications*, Phys. Rep. **679**, 1 (2017).

[70] F. Ticozzi, L. Viola, *Analysis and synthesis of attractive quantum Markovian dynamics*, Automatica **45**, 2002 (2009).

[71] H.M. Wiseman, *Quantum theory of continuous feedback*, Phys. Rev. A **49**, 2133 (1994).

[72] V. P. Belavkin, *Towards the theory of control in observable quantum systems*, Autom. Remote Control **44**, 178 (1983).

[73] H. M. Wiseman and G. J. Milburn, *Quantum theory of optical feedback via homodyne detection*, Phys. Rev. Lett. **70**, 548 (1993).

[74] K. Jacobs, *Feedback control using only quantum back-action*, New J. Phys. **12**, 043005 (2010).

[75] S. Ashhab and F. Nori, *Control-free control: Manipulating a quantum system using only a limited set of measurements*, Phys. Rev. A **82**, 062103 (2010).

[76] A. Stern, *Anyons and the quantum Hall effect — A pedagogical review*, Ann. Phys. **323**, 204 (2008).

[77] G. Moore and N. Read, *Nonabelions in the fractional quantum hall effect*, Nucl. Phys. B **360**, 362 (1991).

[78] N. Read and D. Green, *Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries and the fractional quantum Hall effect*, Phys. Rev. B **61**, 10267 (2000).

[79] Sanghun An, P. Jiang, H. Choi, W. Kang, S. H. Simon, L. N. Pfeiffer, K. W. West, and K. W. Baldwin, *Braiding of Abelian and non-Abelian anyons in the fractional quantum Hall effect*, arXiv:1112.3400.

[80] R. L. Willett, C. Nayak, K. Shtengel, L. N. Pfeiffer, and K. W. West, *Magnetic field-tuned Aharonov-Bohm oscillations and evidence for non-Abelian anyons at* $\nu = 5/2$, Phys. Rev. Lett. **111**, 186401 (2013).

[81] R. M. Lutchyn, E. P. A. M. Bakkers, L. P. Kouwenhoven, P. Krogstrup, C. M. Marcus, and Y. Oreg, *Realizing Majorana zero modes in superconductor-semiconductor heterostructures*, Nature Rev. Mater. **3**, 52 (2018).

[82] S. Das Sarma, M. Freedman, and C. Nayak, *Topologically protected qubits from a possible non-Abelian fractional quantum Hall state*, Phys. Rev. Lett. **94**, 166802 (2005).

[83] A. Stern and B. I. Halperin, *Proposed experiments to probe the non-Abelian* $\nu = 5/2$ *quantum Hall state*, Phys. Rev. Lett. **96**, 016802 (2006).

[84] P. Bonderson, A. Kitaev, and K. Shtengel, *Detecting non-Abelian statistics in the* $\nu = 5/2$ *fractional quantum Hall state*, Phys. Rev. Lett. **96**, 016803 (2006).

[85] G. E. Volovik, *Fermion zero modes on vortices in chiral superconductors*, JETP Letters. **70**, 609 (1999).

[86] A. Kitaev, *Unpaired Majorana fermions in quantum wires*, Phys. Usp. **44** (suppl.), 131 (2001).

[87] R. M. Lutchyn, J. D. Sau, and S. Das Sarma, *Majorana fermions and a topological phase transition in semiconductor-superconductor heterostructures*, Phys. Rev. Lett. **105**, 077001 (2010).

[88] Y. Oreg, G. Refael, and F. von Oppen, *Helical liquids and Majorana bound states in quantum wires*, Phys. Rev. Lett. **105**, 177002 (2010).

[89] P. Bonderson, M. Freedman, and C. Nayak, *Measurement-only topological quantum computation*, Phys. Rev. Lett. **101**, 010501 (2008).

[90] J. Alicea, Y. Oreg, G. Refael, F. von Oppen, and M. P. A. Fisher, *Non-Abelian statistics and topological quantum information processing in 1D wire networks*, Nature Phys. **7**, 412 (2011).

[91] B. van Heck, A. R. Akhmerov, F. Hassler, M. Burrello, and C. W. J. Beenakker, *Coulomb-assisted braiding of Majorana fermions in a Josephson junction array*, New J. Phys. **14**, 035019 (2012).

[92] S. Vijay and L. Fu, *Braiding without braiding: Teleportation-based quantum information processing with Majorana zero-modes*, Phys. Rev. B **94**, 235446 (2016).

[93] T. Karzig, C. Knapp, R. M. Lutchyn, P. Bonderson, M. B. Hastings, C. Nayak, J. Alicea, K. Flensberg, S. Plugge, Y. Oreg, C. M. Marcus, and M. H. Freedman, *Scalable designs for quasiparticle-poisoning-protected topological quantum computation with Majorana zero-modes*, Phys. Rev. B **95**, 235305 (2017).

[94] A. R. Akhmerov, J. Nilsson, and C. W. J. Beenakker, *Electrically detected interferometry of Majorana fermions in a topological insulator*, Phys. Rev. Lett. **102**, 216404 (2009).

[95] J. Nilsson and A. R. Akhmerov, *Theory of non-Abelian Fabry-Perot interferometry in topological insulators*, Phys. Rev. B **81**, 205110 (2010).

[96] D. J. Clarke and K. Shtengel, *Improved phase gate reliability in systems with neutral Ising anyons*, Phys. Rev. B **82**, 180519(R) (2010).

[97] C.-Y. Hou, F. Hassler, A. R. Akhmerov, and J. Nilsson, *Probing Majorana edge states with a flux qubit*, Phys. Rev. B **84**, 054538 (2011).

[98] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, *Chiral topological superconductor from the quantum Hall state*, Phys. Rev. B **82**, 184516 (2010).

[99] L. Fu and C. L. Kane, *Probing neutral Majorana fermion edge modes with charge transport*, Phys. Rev. Lett. **102**, 216403 (2009).

[100] A. C. Potter and L. Fu, *Anomalous supercurrent from Majorana states in topological insulator Josephson junctions*, Phys. Rev. B **88**, 121109(R) (2013).

[101] S. Park and P. Recher, *Detecting the exchange phase of Majorana bound states in a Corbino geometry topological Josephson junction*, Phys. Rev. Lett. **115**, 246403 (2015).

[102] P. Fendley, M. P. A. Fisher, and C. Nayak, *Edge states and tunneling of non-Abelian quasiparticles in the $\nu = 5/2$ quantum Hall state and $p + ip$ superconductors*, Phys. Rev. B **75**, 045317 (2007).

[103] E. Grosfeld and A. Stern, *Observing Majorana bound states of Josephson vortices in topological superconductors*, PNAS **108**, 11810 (2011).

[104] J. Keeling, I. Klich, and L. S. Levitov, *Minimal excitation states of electrons in one-dimensional wires*, Phys. Rev. Lett. **97**, 116403 (2006).

[105] C. W. Groth, M. Wimmer, A. R. Akhmerov, and X. Waintal, *Kwant: A software package for quantum transport*, New J. Phys. **16**, 063065 (2014).

[106] P. W. Brouwer, *Scattering approach to parametric pumping*, Phys. Rev. B **58**, R10135(R) (1998).

[107] B. Tarasinski, D. Chevallier, J. A. Hutasoit, B. Baxevanis, and C. W. J. Beenakker, *Quench dynamics of fermion-parity switches in a Josephson junction*, Phys. Rev. B **92**, 144306 (2015).

[108] J. Dubois, T. Jullien, F. Portier, P. Roche, A. Cavanna, Y. Jin, W. Wegscheider, P. Roulleau, and D. C. Glattli, *Minimal-excitation states for electron quantum optics using levitons*, Nature **502**, 659 (2013).

[109] Junying Shen, Jian Lyu, Jason Zheshen Gao, Chui-Zhen Chen, Chang-woo Cho, Lei Pan, Zhijie Chen, Kai Liu, Y. J. Hu, K. Y. Yip, S. K. Goh, Qing Lin He, Kang L. Wang, Kam Tuen Law, and Rolf Lortz, *Spectroscopic evidence of chiral Majorana modes in a*

quantum anomalous Hall insulator/superconductor heterostructure, arXiv:1809.04752.

[110] C. W. von Keyserlingk, S. H. Simon, and B. Rosenow, *Enhanced bulk-edge Coulomb coupling in fractional Fabry-Perot interferometers*, Phys. Rev. Lett. **115**, 126807 (2015).

[111] E. Bocquillon, V. Freulon, F. D. Parmentier, J.-M Berroir, B. Plaçais, C. Wahl, J. Rech, T. Jonckheere, T. Martin, C. Grenier, D. Ferraro, P. Degiovanni, and G. Fève, *Electron quantum optics in ballistic chiral conductors*, Ann. Physik **526**, 1 (2014).

[112] S. Das Sarma, M. Freedman, and C. Nayak, *Majorana zero modes and topological quantum computation*, npj Quantum Inf. **1**, 15001 (2015).

[113] M. Leijnse and K. Flensberg, *Introduction to topological superconductivity and Majorana fermions*, Semicond. Sci. Technol. **27**, 124003 (2012).

[114] A. Stern, F. von Oppen, and E. Mariani, *Geometric phases and quantum entanglement as building blocks for non-Abelian quasiparticle statistics*, Phys. Rev. B 70, 205338 (2004).

[115] M. Stone and S.-B. Chung, *Fusion rules and vortices in $p_x + ip_y$ superconductors*, Phys. Rev. B **73**, 014505 (2006).

[116] C. W. J. Beenakker, P. Baireuther, Y. Herasymenko, I. Adagideli, Lin Wang, and A. R. Akhmerov, *Deterministic creation and braiding of chiral edge vortices*, arXiv:1809.09050.

[117] G. Strübi, W. Belzig, M.-S. Choi, and C. Bruder, *Interferometric and noise signatures of majorana fermion edge states in transport experiments*, Phys. Rev. Lett. **107**, 136403 (2011).

[118] L. Chirolli, J. P. Baltanás, and D. Frustaglia, *Chiral Majorana interference as a source of quantum entanglement*, Phys. Rev. B **7**, 155416 (2018).

[119] Yan-Feng Zhou, Zhe Hou, Peng Lv, X.C. Xie, and Qing-Feng Sun, *Magnetic flux control of chiral Majorana edge modes in topological superconductor*, Sci. China-Phys. Mech. Astron. **61**, 127811 (2018).

[120] I. Klich, *An elementary derivation of Levitov's formula*, in: *Quantum Noise in Mesoscopic Physics*, NATO Science Series II, **97**, 397 (2003).

[121] J. E. Avron, S. Bachmann, G. M. Graf, and I. Klich, *Fredholm determinants and the statistics of charge transport*, Commun. Math. Phys. **280**, 807 (2008).

[122] I. Klich, *A note on the full counting statistics of paired fermions*, J. Stat. Mech. P11006 (2014). When comparing formulas, note that Klich has a factor of two in the anticommutator of Majorana operators.

[123] C. W. J. Beenakker, *Annihilation of colliding Bogoliubov quasiparticles reveals their Majorana nature*, Phys. Rev. Lett. **112**, 070604 (2014).

[124] M. E. Fisher and R. E. Hartwig, *Toeplitz determinants: Some applications, theorems, and conjectures*, Adv. Chem. Phys. **15**, 333 (1968).

[125] R. E. Hartwig and M. E. Fisher, *Asymptotic behavior of Toeplitz matrices and determinants*, Rational Mech. Anal. **32**, 190 (1969).

[126] C. W. J. Beenakker, *Random-matrix theory of Majorana fermions and topological superconductors*, Rev. Mod. Phys. **87**, 1037 (2015).

[127] J. S. Bell, *On the Einstein Podolsky Rosen paradox*, Phys. Phys. Fiz. **1**, 195 (1964).

[128] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, *Bell nonlocality*, Rev. Mod. Phys. **86**, 419 (2014).

[129] B. S. Cirel'son, *Quantum generalizations of Bell's inequality*, Lett. Math. Phys. **4**, 93 (1980).

[130] J. Uffink, *Quadratic Bell Inequalities as Tests for Multipartite Entanglement*, Phys. Rev. Lett. **88**, 230406 (2002).

[131] B. S. Tsirel'son, *Quantum analogues of the Bell inequalities. The case of two spatially separated domains*, J. Sov. Math. **36**, 557–570 (1987).

[132] L. J. Landau, *Empirical two-point correlation functions*, Found. Phys. **18**, 449 (1988).

[133] Ll. Masanes, *Necessary and sufficient condition for quantum-generated correlations*, arXiv:quant-ph/0309137.

[134] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Proposed Experiment to Test Local Hidden-Variable Theories*, Phys. Rev. Lett. **23**, 880 (1969).

[135] M. Navascués, S. Pironio, and A. Acin, *A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations*, New J. Phys. **10**, 073013 (2008).

[136] J. Silman, S. Pironio, and S. Massar, *Device-Independent Randomness Generation in the Presence of Weak Cross-Talk*, Phys. Rev. Lett. **110**, 100504 (2013).

[137] N. H. Lindner, E. Berg, G. Refael, and A. Stern, *Fractionalizing Majorana Fermions: Non-Abelian Statistics on the Edges of Abelian Quantum Hall States*, Phys. Rev. X **2**, 041002 (2012).

[138] M. Cheng, *Superconducting proximity effect on the edge of fractional topological insulators*, Phys. Rev. B **86**, 195126 (2012).

[139] A. Vaezi, *Fractional topological superconductor with fractionalized Majorana fermions*, Phys. Rev. B **87**, 035132 (2013).

[140] R. S. K. Mong, D. J. Clarke, J. Alicea, N. H. Lindner, P. Fendley, C. Nayak, Y. Oreg, A. Stern, E. Berg, K. Shtengel, and M. P. A. Fisher, *Universal Topological Quantum Computation from a Superconductor-Abelian Quantum Hall Heterostructure*, Phys. Rev. X **4**, 011036 (2014).

[141] J. Klinovaja and D. Loss, *Parafermions in an Interacting Nanowire Bundle*, Phys. Rev. Lett. **112**, 246403 (2014).

[142] J. Klinovaja and D. Loss, *Time-reversal invariant parafermions in interacting Rashba nanowires*, Phys. Rev. B **90**, 045118 (2014).

[143] J. Alicea and P. Fendley, *Topological Phases with Parafermions: Theory and Blueprints*, Annu. Rev. Condens. Matter Phys. **7**, 119 (2016).

[144] A. Romito and Y. Gefen, *Ubiquitous Nonlocal Entanglement with Majorana Zero Modes*, Phys. Rev. Lett. **119**, 157702 (2017).

[145] J. Bulte, A. Bednorz, C. Bruder, and W. Belzig, *Noninvasive Quantum Measurement of Arbitrary Operator Order by Engineered Non-Markovian Detectors*, Phys. Rev. Lett. **120**, 140407 (2018).

[146] A. Carmi and E. Cohen, *On the Significance of the Quantum Mechanical Covariance Matrix*, Entropy **20**, 500 (2018).

[147] A. Bednorz and W. Belzig, *Proposal for a cumulant-based Bell test for mesoscopic junctions*, Phys. Rev. B **83**, 125304 (2011).

[148] Ll. Masanes, A. Acin, and N. Gisin, *General properties of nonsignaling theories*, Phys. Rev. A **73**, 012112 (2006).

[149] A. Carmi and E. Cohen, *Relativistic independence bounds nonlocality*, Sci. Adv. **5**, eaav8370 (2019).

[150] K. Snizhko, R. Egger, and Y. Gefen, *Measurement and control of a Coulomb-blockaded parafermion box*, Phys. Rev. B **97**, 081405 (2018).

[151] L.-B. Fu, *General Correlation Functions of the Clauser-Horne-Shimony-Holt Inequality for Arbitrarily High-Dimensional Systems*, Phys. Rev. Lett. **92**, 130404 (2004).

[152] L.-B. Fu, J.-L. Chen, and X.-G. Zhao, *Maximal violation of the Clauser Horne-Shimony-Holt inequality for two qutrits*, Phys. Rev. A **68**, 022323 (2003).

[153] E. Arthurs and J. L. Kelly, *On the Simultaneous Measurement of a Pair of Conjugate Observables*, Bell Syst. Tech. J. **44**, 725 (1965).

[154] L. M. Johansen and P. A. Mello, *Quantum mechanics of successive measurements with arbitrary meter coupling*, Phys. Lett. A **372**, 5760–5764 (2008).

[155] M. A. Ochoa, W. Belzig, and A. Nitzan, *Simultaneous weak measurement of noncommuting observables: a generalized Arthurs-Kelly protocol*, Sci. Rep. **8**, 15781 (2018).

[156] K. J. Resch and A. M. Steinberg, *Extracting Joint Weak Values with Local, Single-Particle Measurements*, Phys. Rev. Lett. **92**, 130402 (2004).

[157] J. S. Lundeen and A. M. Steinberg, *Experimental Joint Weak Measurement on a Photon Pair as a Probe of Hardy's Paradox*, Phys. Rev. Lett. **102**, 020404 (2009).

[158] D. Litinski, *Magic State Distillation: Not as Costly as You Think*, Quantum **3**, 205 (2019).

[159] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J.Love, A. Aspuru-Guzik, and J. L. O'Brien, *A variational eigenvalue solver on a photonic quantum processor*, Nat. Comm. **5**, 4213 (2014).

[160] J. McClean, S. Boixo, V. Smelyanskiy, R. Babbush, and H. Neven, *Barren plateaus in quantum neural network training landscapes* , Nat. Comm. **9**, 4812 (2018).

[161] P.-L. Dallaire-Demers, J. Romero, L. Veis, S. Sim, and A. Aspuru-Guzik, *Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer*, Quantum Sci. Technol. **4**, 045005 (2019).

[162] E. Farhi, J. Goldstone, and S. Gutmann, *A Quantum Approximate Optimization Algorithm*, arXiv:1411.4028.

[163] S. Lloyd, *Quantum approximate optimization is computationally universal* , arXiv:1812.11075.

[164] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M.Chow, and J. M. Gambetta, *Hardware-efficient Variational Quantum Eigensolver for Small Molecules and Quantum Magnets* , Nature **549**, 242 (2017).

[165] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. Rol, C. Bultink, X. Fu, C. Price, V. Ostroukh, N. Muthusubramanian, A. Bruno, M. Beekman, N. Haider, T. O'Brien, and L. DiCarlo, *Error Mitigation by Symmetry Verification on a Variational Quantum Eigensolver*, Phys. Rev. A **100**, 010302 (2019).

[166] G. Guerreschi and M. Smelyanskiy, *Practical optimization for hybrid quantum-classical algorithms*, arXiv:1701.01450.

[167] O. Higgott, D. Wang, and S. Brierley, *Variational Quantum Computation of Excited States*, Quantum **3**, 156 (2019).

[168] S. Endo, T. Jones, S. McArdle, X. Yuan, and S. Benjamin, *Variational quantum algorithms for discovering Hamiltonian spectra*, Phys. Rev. A **99**, 062304 (2019).

[169] K. M. Nakanishi, K. Fujii, and S. Todo, *Sequential minimal optimization for quantum-classical hybrid algorithms*, Phys. Rev. Research **2**, 043158 (2020).

[170] D. Gottesman, *Stabilizer Codes and Quantum Error Correction*, PhD Dissertation, California Institute of Technology (1997).

[171] B.T. Gard, L. Zhu, G.S. Barron, N.J. Mayhall, S.E. Economou, and E. Barnes, *Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm*, NPJ Quantum Inf. **6**, 10 (2020).

[172] J. Kirkwood and L. Thomas, *Expansions and phase transitions for the ground state of quantum Ising lattice systems*, Commun. Math. Phys. **88**, 569 (1983).

[173] S. Bravyi, D. DiVincenzo, and D. Loss, *Polynomial-time algorithm for simulation of weakly interacting quantum spin systems*, Commun. Math. Phys. **284**, 481 (2008).

[174] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, *Elucidating reaction mechanisms on quantum computers*, Proc. Nat. Acad. Sci. USA **114**, 7555 (2017).

[175] J.-S. Xu, M.-H. Yung, X.-Y. Xu, S. Boixo, Z.-W. Zhou, C.-F. Li, A. Aspuru-Guzik, and G.-C. Guo, *Demon-like algorithmic quantum cooling and its realization with quantum optics*, Nat. Photonics **8**, 113 (2014).

[176] S. McArdle, T. Jones, S. Endo, Y. Li, S. Benjamin, and X. Yuan, *Variational ansatz-based quantum simulation of imaginary time evolution*, npj Quant. Inf. **5**, 75 (2019).

[177] M. Motta, C. Sun, A. T. K. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandao, and G. K.-L. Chan, *Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution*, Nat. Phys. **16**, 205 (2020).

[178] O. Kyriienko, *Quantum inverse iteration algorithm for programmable quantum simulators*, npj Quantum Information **6**, 7 (2020).

[179] S. Kretschmer, K. Luoma, and W.T. Strunz, *Collision model for non-Markovian quantum dynamics*, Phys. Rev. A **94**, 012106 (2016).

[180] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, *Efficient Quantum Algorithms for Simulating Sparse Hamiltonians*, Comm. Mat. Phys. **270**, 359 (2007).

[181] M. Suzuki, *Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems*, Comm. Mat. Phys. **51**, 183 (1976).

[182] Z.-C. Yang, A. Rahmani, A. Shabani, H. Neven, and C. Chamon, *Optimizing Variational Quantum Algorithms using Pontryagin's Minimum Principle*, Phys. Rev. X **7**, 021027 (2017).

[183] A. Bapat and S. Jordan, *Bang-bang control as a design principle for classical and quantum optimization algorithms*, Quant. Inf. Comp. **19**: 424-446 (2019).

[184] F. G. Brandao and K. M. Svore, *Quantum Speed-Ups for Solving Semidefinite Programs*, IEEE FOCS **555**, 415 (2017).

[185] J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, *Integrated photonic quantum technologies*, Nat. Photonics **14**, 273 (2020).

[186] R. van Handel, J. K. Stockton, and H. Mabuchi, *Modelling and feedback control design for quantum state preparation*, J. Opt. B **7**, 10 (2005).

[187] I. Bloch, *Quantum coherence and entanglement with ultracold atoms in optical lattices*, Nat. **453**, 1016 (2008).

[188] L. C. Kwek, Z. Wei, and B. Zeng, *Measurement-Based Quantum Computing with Valence-Bond-Solids*, Int. J. Mod. Phys. B (2012).

[189] J.-W. Pan, C. Simon, C. Brukner, A. Zeilinger, *Entanglement purification for quantum communication* , Nature **410**, 1067 (2001).

[190] X.-Y. Luo, Y.-Q. Zou, L.-N. Wu, Q. Liu, M.-F. Han, M. K. Tey, and L. You, *Deterministic entanglement generation from driving through quantum phase transitions*, Science **355** 620 (2017).

[191] J.G. Bohnet, B.C. Sawyer, J. W. Britton, M.L. Wall, A.M. Rey, M. Foss-Feig, and J. J. Bollinger, *Quantum spin dynamics and entanglement generation with hundreds of trapped ions*, Science **352**, 1297 (2016).

[192] R. Stockill, M.J. Stanley, L. Huthmacher, E. Clarke, M. Hugues, A.J. Miller, C. Matthiesen, C. Le Gall, and M. Atatüre, *Phase-Tuned Entangled State Generation between Distant Spin Qubits*, Phys. Rev. Lett. **119**, 010503 (2017).

[193] J.K. Stockton, R. van Handel, and H. Mabuchi, *Deterministic Dicke-state preparation with continuous measurement and control*, Phys. Rev. A **70**, 022106 (2004).

[194] C. Marr, A. Beige, and G. Rempe, *Entangled state preparation via dissipation-assisted adiabatic passages*, Phys. Rev. A **68**, 033817 (2003).

[195] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller, *Preparation of entangled states by quantum Markov processes*, Phys. Rev. A **78**, 042307 (2008).

[196] A. Pechen and H. Rabitz, *Teaching the environment to control quantum systems*, Phys. Rev. A **73**, 062102 (2006).

[197] X.-Q. Shao, T.-Y. Zheng, and S. Zhang, *Engineering steady three-atom singlet states via quantum-jump-based feedback*, Phys. Rev. A **85**, 042308 (2012).

[198] Y. Lin, J.P. Gaebler, F. Reiter, T.R. Tan, R. Bowler, Y. Wan, A. Keith, E. Knill, S. Glancy, K. Coakley, A.S. Sørensen, D. Leibfried, and D.J. Wineland, *Preparation of Entangled States through Hilbert Space Engineering*, Phys. Rev. Lett. **117**, 140502 (2016).

[199] Z. Liu, L. Kuang, K. Hu, L. Xu, S. Wei, L. Guo, and X.-Q. Li, *Deterministic creation and stabilization of entanglement in circuit QED by homodyne-mediated feedback control*, Phys. Rev. A **82**, 032335 (2010).

[200] A. Pechen, N. Il'in, F. Shuang, and H. Rabitz, *Quantum control by von Neumann measurements*, Phys. Rev. A **74**, 052102 (2006).

[201] L. Roa, A. Delgado, M. L. Ladrón de Guevara, and A. B. Klimov, *Measurement-driven quantum evolution*, Phys. Rev. A **73**, 012322 (2006).

[202] S. Roy, J.T. Chalker, I.V. Gornyi, and Y. Gefen, *Measurement-induced steering of quantum systems*, Phys. Rev. Research **2**, 033347 (2020).

[203] M. Ippoliti, M. J. Gullans, S. Gopalakrishnan, D. A. Huse, and V. Khemani, *Entanglement phase transitions in measurement-only dynamics*, Phys. Rev. X **11**, 011030, (2021).

[204] T. Grigoletto and F. Ticozzi, *Stabilization via feedback switching for quantum stochastic dynamics*, arXiv:2012.08712.

[205] A. Larrouy, S. Patsch, R. Richaud, J.-M. Raimond, M. Brune, C. P. Koch, and S. Gleyzes, *Fast Navigation in a Large Hilbert Space Using Quantum Optimal Control*, Phys. Rev. X **10**, 021058 (2020).

[206] S. Fu, G. Shi, A. Proutiere, and M. R. James, *Feedback Policies for Measurement-based Quantum State Manipulation*, Phys. Rev. A **90**, 062328 (2014).

[207] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki *Rigorous results on valence-bond ground states in antiferromagnets*, Phys. Rev. Lett. **59**, 799 (1987).

[208] O. Zilberberg, A. Romito, D. J. Starling, G. A. Howland, C. J. Broadbent, J. C. Howell, and Y. Gefen, *Null Values and Quantum State Discrìmination* , Phys. Rev. Lett. **110**, 170405 (2013).

[209] M. Hein, J. Eisert, and H. J. Briegel, *Multi-party entanglement in graph states*, Phys. Rev. A **69**, 062311 (2004).

[210] M.Fannes, B.Nachtergaele, and R. F. Werner, *Finitely correlated states on quantum spin chains*, Commun. Math. Phys. **144**, 443 (1992).

[211] P. W. Anderson, *Infrared Catastrophe in Fermi Gases with Local Scattering Potentials*, Phys. Rev. Lett. **18**, 1049 (1967).

[212] W. Dür, G. Vidal, and J. I. Cirac, *Three qubits can be entangled in two inequivalent ways*, Phys. Rev. A **62**, 062314 (2000).

[213] P. Kumar, K. Snizhko, and Y. Gefen, *Engineering two-qubit mixed states with weak measurements*, Phys. Rev. Research **2**, 042014(R) (2020).

[214] S. Borah, B. Sarma, M. Kewming, G. J. Milburn, and J. Twamley, *Measurement-Based Feedback Quantum Control with Deep Reinforcement Learning for a Double-Well Nonlinear Potential*, Phys. Rev. Lett. **127**, 190403 (2021).

[215] D. Bondarenko, P. Feldmann, *Quantum autoencoders to denoise quantum data*, Phys. Rev. Lett. **124**, 130502 (2020).

[216] C.J.C.H. Watkins, *Learning from delayed rewards*, Ph.D. dissertation, Psychology Dept. Univ. of Cambridge, UK (1989).

[217] G.A. Rummery, M. Niranjan, *On-line Q-learning using connectionist systems*, Engineering Dept. Univ. of Cambridge, UK (1994).

Samenvatting

Dit proefschrift behandelt fundamentele aspecten van de bouw en toepassing van digitale quantum computers. Door de wetten van de quantum mechanica te benutten, kunnen deze apparaten op een efficiënte wijze sommige problemen oplossen die te moeilijk zijn voor gewone computers. Hoewel ze veelbelovend zijn, bevinden digitale quantum computers zich nog in het beginstadium van hun ontwikkeling.

Één manier om deze apparaten te realiseren is om materialen te gebruiken die deeltjes bevatten die anyonen worden genoemd. Hoewel de theorie voorspelt dat er met deze deeltjes op een heel robuuste manier gerekend kan worden (door gebruik te maken van hun niet-Abelse uitwisselingsstatistiek), is de praktische uitwerking tot nu toe niet mogelijk gebleken. Deze problematiek is de motivatie voor hoofdstukken 2 en 3, waarin we een nieuwe strategie ontwikkelen om anyonen te detecteren. De methode maakt gebruik van zogenaamde topologische supergeleiders, die de mogelijkheid bieden om anyonen langs de rand van het materiaal te laten lopen in zogenaamde Majorana kanalen. De niet-Abelse uitwisselingsstatistiek kan gedetecteerd worden door de anyonen te laten passeren langs een magnetische vortex in het binnenste van de supergeleider (zie hoofdstuk 2). Het is ook mogelijk om het ene anyon langs het andere te laten passeren (zie hoofdstuk 3). Een voordeel van onze implementatie is dat detectie volledig elektrisch kan plaatsvinden.

Hoofdstuk 4 behandelt hetzelfde thema van de anyonen van een andere soort, zogenaamde parafermionen. De theorie voorspelt dat deze verschijnen aan de randen van een halfgeleider in een sterk magneetveld, als het fractionele quantum Hall effect optreedt. We beschrijven een methode om de correlaties (verstrengeling) tussen de parafermionen te meten. Verstrengeling is een intrinsiek quantum mechanisch effect, zonder analogon in de klassieke fysica.

De volgende hoofdstukken betreffen toepassingen van quantum computers. Hiervoor onderzoeken we de klasse van algoritmes die berusten op een variatieprincipe. Zo'n principe stelt ons in staat om de grondtoestand (laagste energie-toestand) van een quantum mechanisch systeem te vinden. In hoofdstuk 5 onderzoeken we het cruciale onderdeel van het variatie algoritme, namelijk het vinden van een circuit van qubits dat als eerste

benadering ("Ansatz") voor de grondtoestand kan dienen. We ontwikkelen een methode om het gewenste circuit zo efficiënt mogelijk te realiseren. Onze nieuwe methode is getest op het Ising model en is aantoonbaar meer efficiënt dan de bestaande methode. Een andere manier om de grondtoestand te bereiken is door het natuurlijke proces van afkoelen op de quantum computer na te bootsen. In hoofdstuk 6 laten we zien dat een enkele qubit de koelende werking van een warmtebad kan nabootsen. We ontwikkelen twee verschillende algoritmes voor het koelproces: één die geschikt is voor de huidige generatie van quantumcomputers, maar een beperkte nauwkeurigheid heeft, en een andere die nauwkeuriger is maar hogere eisen stelt aan de quantumcomputer. In hoofdstuk 7 beschrijven we een techniek om een willekeurige toestand (niet noodzakelijk de grondtoestand) op een quantumcomputer te realiseren, door gebruik te maken van een serie van metingen die het systeem maar heel weinig verstoren (een zogenaamde "zwakke" meting). Dit is een bekende techniek, ons doel in dit hoofdstuk is om sneller tot de gewenste toestand te geraken. Hiervoor vergelijken we twee methodes, die beiden de bestaande methodes een factor tien versnellen.

Summary

This thesis deals with fundamental aspects of the construction and application of digital quantum computers. By harnessing the laws of quantum mechanics, these devices could solve some problems that are too difficult for ordinary computers. Although promising, digital quantum computers are still in the early stages of their development.

One way to realize such hardware is to use materials with special energy excitations called anyons. Theory predicts that these excitations can encode and process quantum information robustly via their mutual interchange, also referred to as braiding. However, existing proposals for anyon braiding prove challenging to realize in an experiment. This issue motivates chapters 2-4, in which we develop new strategies to detect anyons.

The first strategy uses a so-called topological superconductor, which offers the possibility to run anyons (magnetic vortices) along its edge. The effect of the braiding statistics can be detected by passing these anyons by an immobile vortex in the interior of the superconductor (see chapter 2). It is also possible to pass one itinerant anyon by the other (see chapter 3). An advantage of our implementation is that the detection can be realized using only electrical fields.

The second strategy concerns parafermions — a different kind of anyon, which can be realized at the edges of a semiconductor in a strong magnetic field (so-called fractional Quantum Hall regime). To detect parafermions, we introduced several quantities that characterize their quantum correlations, or entanglement. One can perform a weak-measurement protocol to probe these quantities in an experiment.

In addition to the physics of digital quantum hardware, we investigated the potential strategies for its utilization. One promising application of quantum computers is simulating a quantum system's ground state (its lowest-energy state). Such a simulation is a computational problem that is highly relevant to material science and quantum chemistry. Chapters 5-7 introduce several new approaches to preparing ground states using digital quantum computers.

Chapter 5 is concerned with so-called variational quantum algorithms. These algorithms employ a tunable quantum circuit as a ground state

template ("ansatz"), approximating the ground state of a simulated system by optimizing the output energy of the circuit. We focus on constructing efficient ansatz circuits. The first step is designing an ansatz that provably spans the space of all quantum states via a minimal number of digital quantum operations. This ansatz allows a reduction to more practical ansatzes, in agreement with the linked-cluster theorem, which certifies the efficiency of the reduced ansatz. We confirm the efficiency of our method compared to existing alternatives by a numerical study of the weakly coupled quantum Ising model.

Another approach to preparing a ground state on a quantum computer is to simulate the natural process of cooling. Chapter 6 outlines a cooling-like algorithm that simulates a cold bath with a single qubit. We pin down the critical challenges of this approach, which come from the limited simulation time and the digital nature of quantum hardware. We propose two variants of Quantum Digital Cooling algorithms. The first – BangBang approach – is suitable for near-term applications but it is inaccurate. The second approach – LogSweep – aims to be asymptotically accurate but is only suitable for applications in the fault-tolerant regime. Numerically, applying Quantum Digital Cooling to the quantum Ising model shows a scalable performance of LogSweep within a broad interval of model parameters.

Quantum hardware also allows one to prepare the desired quantum state by employing weak measurements. In chapter 7, we propose accelerating such a method of state preparation by deciding the subsequently applied measurements on the go. Because of the vastness of the space of quantum states, optimizing this decision proves challenging. We offer two heuristic approaches to this "navigation" problem. A "compass" method uses the geometrical distance to the target state and aims to reduce it in a greedy fashion. The second method uses a graph, representing measurement operations with edges and quantum states with vertices. Using a numerical test, we identify the potential to accelerate the state preparation at least 10-fold by each of these two heuristics.

Curriculum Vitæ

I was born in Kyiv, Ukraine, in 1995. There, I attended primary and secondary school and received my high school education from the Natural Sciences Lyceum no. 145.

From 2012 to 2016, I studied physics at Kyiv National Taras Shevchenko University (Quantum Field Theory track). During this Bachelor's program, I completed an internship in the group of Carlo Beenakker in Leiden. Upon receiving my diploma in 2016, I took on a Master's program in Leiden. Then I interned at the Weizmann Institute in the group of Yuval Gefen and participated in several summer schools. I obtained my Master's degree summa cum laude in 2018.

In February 2018, I started my Ph.D. studies in Carlo Beenakker's group at Leiden University. My work first concentrated on the theory of non-conventional braiding platforms, then shifted towards the problem of preparing many-body states using quantum digital hardware. In the summer of 2021, I completed a 3-month internship at Google Quantum AI supervised by Thomas O'Brien and Vadim Smelyanskiy. During my Ph.D. studies, I also gave teaching assistance in a Condensed Matter Theory Master's course and two extra-curricular Advanced Theoretical Physics courses. In addition to that, I supervised one Bachelor and one Master student. I traveled to Kyiv, giving research-level introductory courses for university students, as well as popular science lectures for a broader audience. I attended several schools and conferences, presenting my work in Italy, Spain, Austria, Germany, USA, Israel, Ukraine and the Netherlands.

List of publications

- Y. Gerasimenko, B. Tarasinski, and C. W. J. Beenakker. *Attractor-repeller pair of topological zero-modes in a nonlinear quantum walk.* Physical Review A **93**, 022329 (2016).
- Y. Herasymenko, K. Snizhko, and Y. Gefen. *Universal quantum noise in adiabatic pumping.* Physical Review Letters **120**, 226802 (2018).
- C. W. J. Beenakker, N. V. Gnezdilov, E. Dresselhaus, V. P. Ostroukh, Y. Herasymenko, I. Adagideli, and J. Tworzydło. *Valley switch in a graphene superlattice due to pseudo-Andreev reflection.* Physical Review B **97**, 241403 (2018).
- C. W. J. Beenakker, P. Baireuther, Y. Herasymenko, I. Adagideli, L. Wang, and A. R. Akhmerov. *Deterministic creation and braiding of chiral edge vortices.* Physical Review Letters **122**, 146803 (2019). [Chapter 2]
- A. Carmi, Y. Herasymenko, E. Cohen, and K. Snizhko. *Bounds on nonlocal correlations in the presence of signaling and their application to topological zero modes.* New Journal of Physics **21**, 073032 (2019). [Chapter 4]
- C. W. J. Beenakker, A. Grabsch, and Y. Herasymenko. *Electrical detection of the Majorana fusion rule for chiral edge vortices in a topological superconductor.* SciPost Physics **6**, 022 (2019). [Chapter 3]
- Y. Herasymenko and T. E. O'Brien. *A diagrammatic approach to variational quantum ansatz construction.* Quantum **5**, 596 (2021). [Chapter 5]
- S. Polla, Y. Herasymenko, and T. E. O'Brien. *Quantum digital cooling.* Physical Review A **104**, 012414 (2021). [Chapter 6]
- Y. Herasymenko, I. Gornyi, and Y. Gefen. *Measurement-driven navigation in many-body Hilbert space: Active-decision steering.* arXiv:2111.09306. [Chapter 7]