

# Applications of topology to Weyl semimetals and quantum computing

PROEFSCHRIFT

TER VERKRIJGING VAN  
DE GRAAD VAN DOCTOR AAN DE UNIVERSITEIT LEIDEN,  
OP GEZAG VAN RECTOR MAGNIFICUS PROF. MR. C.J.J.M. STOLKER,  
VOLGENS BESLUIT VAN HET COLLEGE VOOR PROMOTIES  
TE VERDEDIGEN OP DONDERDAG 20 JUNI 2019  
KLOKKE 15.00 UUR

DOOR

Thomas Eugene O'Brien

GEBOREN TE OTAHUHU (NIEUW-ZEELAND) IN 1990

Promotores: Prof. dr. C. W. J. Beenakker  
Prof. dr. L. DiCarlo (Technische Universiteit Delft)  
Co-promotor: Dr. İ. Adagideli (Sabancı University, Istanbul)

Promotiecommissie: Prof. dr. J. Aarts  
Prof. dr. A. Achúcarro  
Dr. V. Cheianov  
Dr. M. Veldhorst (Technische Universiteit Delft)  
Prof. dr. F. K. Wilhelm-Mauch (Saarland University)

Casimir PhD series, Delft-Leiden 2019-16

ISBN 978-90-8593-402-8

An electronic version of this thesis can be found  
at <https://openaccess.leidenuniv.nl>

This work has been partially funded by the Netherlands  
Organization for Scientific Research (NWO/OCW), and by the  
European Research Council (ERC).

Printed by: ProefschriftMaken || [www.proefschriftmaken.nl](http://www.proefschriftmaken.nl)

Cover: Quantum circuits for interleaved surface code X and Z stabilizer  
measurements on a Möbius strip, forcing their intersection.

*To my family; Mum, Dad, Tess, and to Louise.*





# Contents

<b>1. Introduction</b>	<b>1</b>
1.1. Preface . . . . .	1
1.2. Topological phases of matter . . . . .	2
1.2.1. Topological quantum computation with Majorana zero-modes . . . . .	4
1.2.2. Weyl semimetals . . . . .	7
1.3. Quantum error correction . . . . .	11
1.3.1. Toy example - the repetition code . . . . .	12
1.3.2. Stabilizer codes . . . . .	15
1.3.3. Stabilizer code dynamics . . . . .	18
1.3.4. Topological quantum error correcting codes . . . . .	19
1.3.5. Decoding topological codes . . . . .	22
1.4. Quantum algorithms . . . . .	24
1.4.1. Quantum phase estimation . . . . .	24
1.4.2. Variational quantum eigensolvers . . . . .	26
1.5. Quantum computing with superconducting qubits . . . . .	27
1.5.1. Transmon architecture . . . . .	27
1.5.2. Sources of errors . . . . .	29
1.6. Outline of this thesis . . . . .	31
1.6.1. Part I . . . . .	31
1.6.2. Part II . . . . .	33
1.6.3. Part III . . . . .	35
 <b>I. Magnetotransport in topological semimetals</b>	 <b>39</b>
 <b>2. Magnetic breakdown and Klein tunneling in a type-II Weyl semimetal</b>	 <b>41</b>
2.1. Introduction . . . . .	41
2.2. Semiclassical quantization . . . . .	43
2.3. Magnetic breakdown . . . . .	44
2.4. Conclusion . . . . .	50
2.A. Low-energy limit of the four-band model Hamiltonian of a type-II Weyl semimetal . . . . .	52

2.B. Topological protection of the special magnetic field axis for Klein tunneling between electron and hole pockets . . . .	53
2.C. Klein tunneling for pairs of connected type-II Weyl points	54
<b>3. Superconductivity provides access to the chiral magnetic effect of an unpaired Weyl cone</b>	<b>57</b>
3.1. Introduction . . . . .	57
3.2. Pathway to single-cone physics . . . . .	58
3.3. Model Hamiltonian of a Weyl superconductor . . . . .	59
3.4. Flux bias into the single-cone regime . . . . .	60
3.5. Magnetic response of an unpaired Weyl cone . . . . .	61
3.6. Consistency of a nonzero equilibrium electrical current and vanishing particle current . . . . .	64
3.7. Numerical simulation . . . . .	65
3.8. Extensions . . . . .	66
3.9. Conclusion . . . . .	67
3.A. Charge renormalization in a superconducting Weyl cone .	67
3.A.1. Block diagonalization . . . . .	68
3.A.2. Current and charge operators . . . . .	69
3.A.3. Effective Hamiltonian in the zeroth Landau level .	71
3.A.4. Renormalized charge for the CME . . . . .	73
3.B. Surface Fermi arc . . . . .	74
3.B.1. Boundary condition . . . . .	74
3.B.2. Construction of the surface state . . . . .	75
3.B.3. Surface dispersion relation . . . . .	77
3.C. Derivation of the renormalized-charge formula for the CME	79
3.C.1. On-shell and off-shell contributions . . . . .	79
3.C.2. Application to the zeroth Landau level . . . . .	81

## II. Topological codes and quantum error correction 83

<b>4. Density-matrix simulation of small surface codes under current and projected experimental noise</b>	<b>85</b>
4.1. Introduction . . . . .	85
4.2. Results . . . . .	86
4.2.1. Error rates for Surface-17 under current experimental conditions . . . . .	86
4.2.2. Optimization of logical error rates with current experimental conditions . . . . .	88

4.2.3.	Projected improvement with advances in quantum hardware . . . . .	90
4.3.	Discussion . . . . .	92
4.3.1.	Computational figure of merit . . . . .	92
4.3.2.	Decoder performance . . . . .	93
4.3.3.	Other observations . . . . .	94
4.3.4.	Effects not taken into account . . . . .	95
4.4.	Methods . . . . .	96
4.4.1.	Simulated experimental procedure . . . . .	96
4.4.2.	Error models . . . . .	101
4.A.	Full circuit diagram for Surface-17 implementation . . . .	104
4.B.	Parameters of error models . . . . .	105
4.B.1.	Qubit idling . . . . .	106
4.B.2.	Photon decay . . . . .	106
4.B.3.	Single-qubit $R_y(\pi/2)$ rotations . . . . .	107
4.B.4.	Flux noise . . . . .	107
4.B.5.	C-Z gates . . . . .	109
4.B.6.	Measurement . . . . .	110
4.C.	Effect of over-rotations and two-qubit phase noise on logical error rate . . . . .	112
4.D.	Calculation of decoder upper bound . . . . .	112
4.E.	Hardware requirements of simulation . . . . .	114
4.F.	Homemade MWPM decoder with asymmetric weight calculation . . . . .	115
4.G.	Implementation of a look-up table decoder . . . . .	118
4.H.	Details of lowest-order approximation . . . . .	119
<b>5.</b>	<b>Adaptive weight estimator for quantum error correction in a time-dependent environment</b>	<b>123</b>
5.1.	Introduction . . . . .	123
5.2.	Quantum error correction and the repetition code . . . . .	124
5.3.	Weight inference from error syndromes . . . . .	126
5.3.1.	Formulation of the inversion problem . . . . .	126
5.3.2.	Solution for edges connecting pairs of vertices . . . .	128
5.3.3.	Solution for boundary edges . . . . .	129
5.4.	Implementation of the adaptive decoder . . . . .	130
5.4.1.	Convergence in the large-time limit . . . . .	130
5.4.2.	Performance in a time-dependent environment . . . .	131
5.5.	Conclusion . . . . .	132

<b>6. Neural network decoder for topological color codes with circuit level noise</b>	<b>135</b>
6.1. Introduction . . . . .	135
6.2. Description of the problem . . . . .	136
6.2.1. Color code . . . . .	136
6.2.2. Error model . . . . .	138
6.2.3. Fault-tolerance . . . . .	139
6.2.4. Flag qubits . . . . .	140
6.3. Neural network decoder . . . . .	141
6.3.1. Learning mechanism . . . . .	141
6.3.2. Decoding algorithm . . . . .	141
6.4. Neural network performance . . . . .	143
6.4.1. Power law scaling of the logical error rate . . . . .	143
6.4.2. Performance on realistic data . . . . .	144
6.5. Conclusion . . . . .	145
6.A. Quantum circuits . . . . .	145
6.A.1. Circuits for the Pauli error model . . . . .	145
6.A.2. Measurement processing for the density-matrix error model . . . . .	146
6.B. Details of the neural network decoder . . . . .	148
6.B.1. Architecture . . . . .	148
6.B.2. Training and evaluation . . . . .	149
6.B.3. Pauli frame updater . . . . .	152
6.C. Results for distance-5 and distance-7 codes . . . . .	153
 <b>III. Quantum Algorithms for digital quantum simulation</b>	 <b>155</b>
<b>7. Majorana-based fermionic quantum computation</b>	<b>157</b>
7.1. Introduction . . . . .	157
7.2. Description of the architecture . . . . .	158
7.3. Quantum algorithms . . . . .	161
7.4. Conclusion . . . . .	165
7.A. Preparing extended ancilla qubits for quantum phase estimation . . . . .	166
7.B. Algorithm to perform Trotter steps in $O(N^3)$ time . . . . .	167
7.C. Details of parallel circuit for Hubbard model . . . . .	169

<b>8. Quantum phase estimation for noisy, small-scale experiments</b>	<b>173</b>
8.1. Introduction . . . . .	173
8.2. Quantum phase estimation . . . . .	175
8.3. Classical data analysis . . . . .	179
8.3.1. Time-series analysis . . . . .	182
8.3.2. Efficient Bayesian analysis . . . . .	187
8.3.3. Experiment design . . . . .	189
8.4. Results without experimental noise . . . . .	190
8.4.1. Single eigenvalues . . . . .	190
8.4.2. Example behaviour with multiple eigenvalues . . . . .	192
8.4.3. Estimator scaling with two eigenvalues . . . . .	194
8.4.4. Many eigenvalues . . . . .	196
8.5. The effect of experimental noise . . . . .	197
8.5.1. Depolarizing noise . . . . .	198
8.5.2. Realistic circuit-level noise . . . . .	199
8.6. Discussion . . . . .	203
Appendices . . . . .	205
8.A. Derivation of the identity in Eq. (8.25) . . . . .	206
8.B. Variance calculations for time-series estimator . . . . .	207
8.C. Fourier representation for Bayesian updating . . . . .	209
8.C.1. Bayesian updating for multi-eigenvalue starting state	211
8.D. Convergence of the (noiseless) time-series analysis in case of multiple eigenvalues. . . . .	214
8.E. Details of realistic simulation . . . . .	214
8.E.1. Error model and error parameters . . . . .	216
<b>Bibliography</b>	<b>219</b>
<b>Acknowledgments</b>	<b>249</b>
<b>Samenvatting</b>	<b>253</b>
<b>Summary</b>	<b>257</b>
<b>Curriculum Vitæ</b>	<b>261</b>
<b>List of publications</b>	<b>263</b>



# 1. Introduction

## 1.1. Preface

Over the past decades (and beyond), condensed matter theory has concerned itself with the classification of physical systems into phases of matter, from which generic features can be identified. This is exemplarized by the recent efforts in studying topological physics: systems classified by global properties that remain invariant under local deformation or perturbation. Motivated originally by quantum Hall systems [1], many such phases have been discovered: topological insulators [2, 3], Majorana wires [31, 35], Weyl semimetals [4, 5, 92, 93] and quantum spin Hall systems [6, 7] being perhaps the most famous examples. These systems promise various exotic properties; anomalous responses to driving fields [62], conducting states immune to localization [2], and perhaps most strikingly, the possibility for excitations that satisfy non-Abelian statistics [8]. However, in lieu of microscopic details of a specific system, the above properties are necessarily qualitative, and quantitative features tend to vary between two systems described by the same toy model.

The field of quantum information, concerned with precise control of single degrees of freedom in small quantum systems, is in direct contrast the above broad strokes. Recent advances in experimental work focus on improving manufacture and measurement to below the one percent mark (in state-of-the-art devices) for all gates and measurement pulses [167, 169, 183], requiring unique, individual control over separate qubits. Then, as the quantum community pushes ever closer to the goal of outperforming classical computers at first any [9, 10] and then useful [263] tasks, we become focused on stretching any advantage as far as can be, optimizing algorithms [266, 290, 291], error mitigation techniques [11–15], and classical control software in lieu of fault-tolerant qubits. This has also made quantum information results-oriented, and less concerned with bigger picture questions like classifying all topological phases of matter.

However, despite the gap above, the two fields remain intertwined and co-dependent. Many quantum information platforms are based on condensed matter systems, such as superconductors [88] and quantum

## 1. Introduction

dots [16]. Possibly the most famous such platform is that of Majorana zero modes [35] (or Majorana bound states), which promise topological protection for quantum operations, possibly negating some of the need for the above fine tuning. On a slightly higher level, probably the most studied quantum error correcting codes are topological codes, which find their roots in Kitaev's celebrated toric code in condensed matter [220]. In return, quantum computers promise to improve our ability to model condensed matter physics, by solving the many-body problem (as originally proposed by Feynman [17]). Also, ideas about entanglement entropy and mutual information have allowed serious advances already in our understanding of such quantum systems. It is very notable that the same mathematics permeates the two fields; in particular the Pauli matrices  $\{X, Y, Z\} \equiv \{\sigma_x, \sigma_y, \sigma_z\}$ , which are present as spins in condensed matter systems and qubits in quantum information.

This thesis explores various topics on both sides of the boundary between condensed matter and quantum information, and some directly at the interface between the two. It has been split into three parts, containing chapters loosely centered on three main topics. This introduction now attempts to give some background required for the various topics therein.

## 1.2. Topological phases of matter

One of the most successful studies in theoretical physics has been the classification of phases of matter: regions in the parameter space of physical systems which display similar characteristics. This is important, as we are often uninterested in precise details of a physical system. Experiments tend to contain some level of unknown disorder, results requiring fine-tuning tend not to be reproducible, and broad features are much more applicable to the universe outside the lab.

To make connection between topological and non-topological phase transitions, let us consider the textbook example of the difference between metals and insulators [24, 25]. Let us take a toy model of a  $D$ -dimensional non-interacting lattice Hamiltonian

$$\hat{H}_{\text{NI}} = \sum_{\vec{i}, \vec{j}, a, b} H_{\vec{i}, a, \vec{j}, b} \hat{c}_{i, a}^\dagger \hat{c}_{j, b}, \quad (1.1)$$

where  $\hat{c}_{i, a}^\dagger$  ( $\hat{c}_{i, a}$ ) creates (annihilates) an electron on site  $\vec{i}$  in spin-orbital



## 1.2. Topological phases of matter

$a^*$ . Following Bloch's theorem [26], if  $\hat{H}_{\text{NI}}$  is invariant under  $D_{\text{TI}}$  linearly independent translations  $\vec{i} \rightarrow \vec{i} + \vec{\delta}^n$ , it may be block-diagonalized by a Fourier transform of the creation and annihilation operators

$$\hat{c}_{\vec{k},\alpha} = \sum_{j_1, \dots, j_{D_{\text{TI}}}}^{\infty} \exp\left(i \sum_n j_n \vec{\delta}_n \cdot \vec{k}\right) \hat{c}_{i_\alpha + \sum_n j_n \vec{\delta}_n, a_\alpha}. \quad (1.2)$$

Here,  $\alpha$  is taken from a maximal *unit cell* of spin-orbitals  $(i_\alpha, a_\alpha)$  that cannot be joined by translations of  $\vec{\delta}_n$ , and  $\vec{k}$  is taken from the Brillouin zone:

$$\hat{c}_{\vec{k}',\alpha} \equiv \hat{c}_{\vec{k},\alpha} \quad \text{if} \quad \vec{k} \cdot \vec{\delta}_n = \vec{k}' \cdot \vec{\delta}_n \pmod{2\pi} \quad (1.3)$$

We may then write

$$\hat{H}_{\text{NI}} = \oplus_{\vec{k}} \hat{H}_{\text{NI}}(\vec{k}), \quad \hat{H}_{\text{NI}}(\vec{k}) = \sum_{\alpha, \beta} H_{\alpha, \beta}(\vec{k}) \hat{c}_{\vec{k}, \alpha}^\dagger \hat{c}_{\vec{k}, \beta}. \quad (1.4)$$

This construction is then a map from the  $D_{\text{TI}}$ -dimensional torus to a set of *energy bands*. Following the Fermi-Dirac principle, only a single electron may occupy each spin-orbital in each point of the Brillouin zone, which implies that only a single electron per unit cell may occupy any given energy band.. Then, the many-body ground state can be found by filling up each energy band with  $E_\alpha(\vec{k}) < 0$ . One may add a *chemical potential*  $\mu \sum_{\alpha, \vec{k}} \hat{c}_\alpha^\dagger(\vec{k}) \hat{c}_\alpha(\vec{k})$  to adjust this. Physically this is performed by doping the system, or by applying an external potential. Band theory dictates that when all energy bands are far from zero energy the system is an insulator, and when some bands pass through  $E = 0$  the system is a metal.

As physics is local, phases are necessarily local: a single system may be split into different phases in different spatial regions. A boundary, being caught between two phases, may then display different physical properties to the regions on either side. In the above example, one may have two local insulators at a PN junction, such that the Fermi level must pass through a band as we shift between the two regions, creating a conducting surface. One can then ask the question whether it is possible to have a boundary without a (obvious) phase transition: is it possible to have two systems that look locally identical, but necessitate a boundary between whenever they touch? An equivalent question is whether phases can be non-local: can two phases of matter exist such that a local observer within one phase cannot

---

\*Lattice models do not generally require that every site contains the same set of orbitals, in which case the sum over  $a$  and  $b$  is different for each choice of sites  $\vec{i}$  and  $\vec{j}$  in Eq. 1.1.

## 1. Introduction

determine which of the two they are in? The answer to these questions is yes, and such non-local phases are called ‘topological’ [27]. We note that the above example is explicitly not an example of such a topological phase; one may distinguish phases by the local electron density per unit cell  $\sum_a \langle \hat{c}_a^\dagger \hat{c}_a \rangle$ . Importantly, this means that the conducting interface in a PN junction is not robust against disorder or interactions, which lead to other types of insulating behaviour [28, 29]. By contrast, topological phase transitions are robust against such behaviour.

The vague notion of ‘similar characteristics’ in topological phases is quantified by a *topological invariant* [27]. This is a similar object to the order parameters of Landau [30] which characterise regular phases. However, there are two key differences:

1. Two phases characterised by an order parameter  $o$  have  $o = 0$  in one phase and  $o > 0$  in the other, such that  $o \rightarrow 0$  as one approaches the phase boundary from either side. Phases characterised by a topological invariant  $\chi$  have strictly  $\chi = \chi_1$  in one phase and  $\chi = \chi_2$  in the other, and the topological invariant ceases to be defined at the boundary (as it cannot be multivalued).
2. Order parameters are (typically) local - they may be calculated as an expectation value of a sum of strictly local terms. Topological invariants are by definition non-local, and cannot be written as such an expectation value.

### 1.2.1. Topological quantum computation with Majorana zero-modes

A system with a genuine topological phase transition is the one-dimensional Kitaev chain [31]. This system (Fig. 1.1) has the Bogoliubov-de Gennes superconducting Hamiltonian [32, 33]

$$\hat{H}_{\text{NI}} = \sum_i \left( -\mu \hat{c}_i^\dagger \hat{c}_i - t \hat{c}_i^\dagger \hat{c}_{i+1} + t \hat{c}_{i+1}^\dagger \hat{c}_i + \Delta \hat{c}_i^\dagger \hat{c}_{i+1}^\dagger - \Delta \hat{c}_i \hat{c}_{i+1} \right), \quad (1.5)$$

One may diagonalize such a Hamiltonian as in Eq. 1.4, but in order to account for the non-particle conserving terms, we must make use of the Bogoliubov transformation between a particle with wavevector  $k$  and its hole at wavevector  $-k$ . This requires that the non-interacting blocks of

the Hamiltonian take the form\*

$$\hat{H}_{\text{NI}}(k) = t(k)\hat{c}_k^\dagger\hat{c}_k + t(k)\hat{c}_{-k}\hat{c}_{-k}^\dagger + \Delta(k)\hat{c}_k^\dagger\hat{c}_{-k}^\dagger + \Delta^*(k)\hat{c}_{-k}\hat{c}_k. \quad (1.6)$$

At  $k = 0$  and  $k = \pi$ ,  $k \equiv -k$ , and so  $\Delta(0) = \Delta(\pi) = 0$  (as  $\hat{c}_k^2 = \hat{c}_k^{\dagger 2} = 0$ ). This implies that the particle and hole sectors of the Hamiltonians are split; if gapped, the two eigenstates of these Hamiltonians are either entirely electron or entirely hole. Note that in the BdG formulation, only positive energy eigenstates correspond to allowed excitations of the BCS ground state [33]. Let us denote the particle-hole parity of the positive energy eigenstate of  $\hat{H}_{\text{NI}}(k)$  as  $P(\hat{H}_{\text{NI}}(k))$ .<sup>†</sup> The only way then to continuously tune  $t$ ,  $\mu$  and  $\Delta$  to change from  $P(\hat{H}_{\text{NI}}(0)) = -1$  to  $P(\hat{H}_{\text{NI}}(0)) = +1$  is to pass through a gapless state (and the same for  $P(\hat{H}_{\text{NI}}(\pi))$ ). If two phases have the same product  $\chi = P(\hat{H}_{\text{NI}}(0))P(\hat{H}_{\text{NI}}(\pi))$ , it turns out that continuously tuning from one to the other is equivalent to the PN transition described earlier (i.e. non-topological). In particular, such a phase transition produces two zero energy states at  $k = 0$  and  $k = \pi$ ; these may be gapped out by any perturbation that couples these two wavevectors [34]. However, a phase transition in which the sign of  $\chi$  changes necessitates a single zero energy state, which in turn cannot be coupled and gapped. Such a zero-energy crossing cannot be gapped out by any mechanism that does not break particle-hole symmetry (allowing the state to couple to its hole counterpart) [31]. This scenario is complicated in realistic settings by additional bands (and in particular electron spin), however under appropriate conditions the topological phase can be retained [36–38], and experimental signatures (such as those shown in Fig. 1.1) have been obtained that suggest its existence in various physical systems [35, 39, 40] (although this remains somewhat contentious [41, 42]).

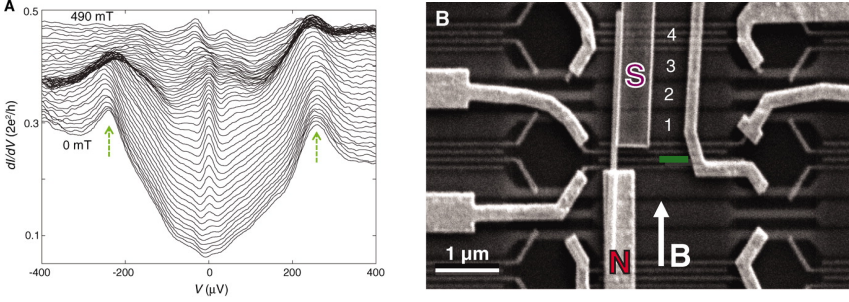
A one-dimensional chain has two ends, and so the zero-energy excitation at a topological-normal phase transition is split evenly across the two ends. This implies that Majorana zero-modes  $\hat{\gamma}_{1,2}$  that commute with  $\hat{H}_{\text{NI}}$  are exponentially localised to either end of the chain. Such operators are Hermitian, and anti-commute with each other. Majorana zero-modes may be coupled by proximity, which adds terms of the form  $E \prod_i \hat{\gamma}_i$  to the Hamiltonian. However, particle-hole symmetry requires such a term to be a product of an even number of Majorana zero-modes only. Given a system

---

\*Please note that compared to Eq. 1.4, we have dropped the vector symbol from the wavevector as we are in one dimension, and lost the  $\alpha$  index as we have only one fermion per unit cell.

<sup>†</sup>This is equivalent to the sign of the Pfaffian of the Hamiltonian in its antisymmetric form [31].

## 1. Introduction



**Figure 1.1.:** Experimental realization of the Kitaev chain in a superconducting nanowire. (left) Conductance spectrum; the peak at zero bias is evidence for the existence of a Majorana state at zero energy. (right) The experimental device; when the central wire is in contact with the superconductor (labelled S), and the external magnetic field is correctly tuned, it is hypothesized that the low-energy effective theory is that of the Kitaev chain. Figure taken with permission from Ref. [35].

of three Majorana zero-modes  $\gamma_{1,2,3}$  (say at the ends of three wires), such interactions are insufficient to gap all three states; a linear combination of the zero-modes will continue to commute with the Hamiltonian [34]. (In larger systems with the possibility of strong correlations, this is raised from a 2-fold to an 8-fold periodicity [34].) Particle-hole symmetry also prevents odd products of Majorana zero-modes from being measured; such a measurement would project the system into a state which is not a parity-eigenstate. However, a product of two Majorana zero-modes defines a fermion,  $\frac{1}{2}(1-i\gamma_i\gamma_j)$ , which has a well-defined parity in a superconductor. If  $2N$  such Majorana zero-modes share a common superconducting phase, their ground state is  $2^N/2$ -fold degenerate (up to corrections exponentially small in the separation of the Majorana zero-modes), with the broken degeneracy being due to the Coulomb energy of the superconductor [34]. If such a system is decoupled from other Majorana zero-modes and grounded, it will decohere into a parity eigenstate [43] (separate from any decoupled fermions); and may be measured by connection to a quantum dot or lead [44].

The most interesting characteristic of Majorana zero-modes is their evolution under exchange. Majorana zero-modes  $\gamma_i, \gamma_j$  may be swapped by tuning external potentials adiabatically. This entails the evolution of the ground state under a unitary  $U(0, t)$  such that  $U(0, t)\gamma_i|\psi(0)\rangle = e^{i\phi}\gamma_j|\psi(t)\rangle$  [43, 45, 46]. If done adiabatically, the system remains in the

ground state. This then requires a time both long enough to prevent excitation to higher-energy states, and short enough to prevent accumulation of phase from coupling Majorana zero-modes (which in turn requires Majorana zero-modes be sufficiently separated at all times). The phase  $\phi$  is key here; one could imagine gauging it away by choosing  $|\psi(t)\rangle$  differently, but if  $|\psi(t)\rangle = |\psi(0)\rangle$  this is not possible. In this case, the remaining *Berry phase* is a physical object, and could be observed by performing the same operation on superpositions of different  $|\psi(0)\rangle$ . In particular, one may show that the Berry phase incurred by a unitary that takes a single Majorana zero-mode on a single loop around  $N$  other Majorana zero-modes is  $\phi = N\pi$ . Such a phase is a property of the topology of the path (namely that it wraps around other Majorana zero-modes), and is unaffected by local perturbations. We are free to choose any gauge for our unitaries that produces this final result; one such choice is to assign the operator  $e^{i\frac{\pi}{4}\gamma_i\gamma_j}$  to the exchange of Majorana zero-modes  $i$  and  $j$ . With this choice of gauge, one may realise the entire Clifford group on the parity-protected sector of  $2N$  Majorana zero-modes by braiding [47]. As fault-tolerant universal quantum computation may be achieved with noiseless Clifford gates and a single noisy non-Clifford unitary [48], this makes Majorana zero-modes an attractive option for quantum devices. In chapter 7 we will study this in more detail for the purposes of digital quantum simulation of fermionic systems, and show that such quantum computers allow for simpler such algorithms than devices using bosonic qubit modes.

### 1.2.2. Weyl semimetals

In the previous section, a protected zero-energy mode emerged at the boundary of a band insulator. One could ask whether it is possible to arrange for such a mode to remain inside the bulk of the system instead of the surface. To be specific, this corresponds to constructing an overlap between two (or more [49], but in this work, we only consider two) energy bands that cannot be lifted by local perturbations. Such a two-band system is described by a Hamiltonian  $\hat{H}(\vec{k})$  that is a two-state system for each  $\vec{k}$ . Modulo a chemical potential, a band overlap implies that  $\hat{H}(\vec{k}_0) = 0$  for some  $\vec{k}_0$ . Perturbation theory suggests such a degeneracy would be unstable against any perturbation [25], which would split the states by introducing some non-trivial term  $\delta\hat{H} = \sum_a E_a\sigma_a$  for  $a = x, y, z$ . However, suppose if, the Hamiltonian takes the structure of the (3D) *Weyl*

## 1. Introduction

Hamiltonian\* [50],

$$\hat{H}_{\text{Weyl}}(\vec{k}) = \vec{\sigma} \cdot J \cdot (\vec{k} - \vec{k}_0), \quad (1.7)$$

where  $J$  is a invertible  $3 \times 3$  matrix. Such a Hamiltonian has a conical bandstructure that overlaps at  $\vec{k} = \vec{k}_0$  (see Fig. 1.2). In this system, the perturbation described above could be absorbed into the Hamiltonian by shifting  $\vec{k}_0 \rightarrow \vec{k}'_0 = \vec{k}_0 - J^{-1}\vec{E}$ . This result would then continue to have a zero mode, although the position (the Weyl point) of this mode would shift from  $\vec{k}_0$  to  $\vec{k}'_0$ . One might consider a perturbation that varies in  $\vec{k}$  -  $\delta\hat{H}(\vec{k}) = \sum_a E_a(\vec{k})\sigma_a + E_0(\vec{k})$ . So long as  $E_a(\vec{k})$  is bounded above, a solution exists to the coupled equations  $J \cdot \vec{k} + E_a(\vec{k}) = 0$ , and at each such point  $\vec{k}_{\text{new}}$  a Weyl cone exists to first order in  $\vec{k} - \vec{k}_{\text{new}}$ . A system with such a bandstructure near the Fermi surface is known as a *Weyl semimetal*.

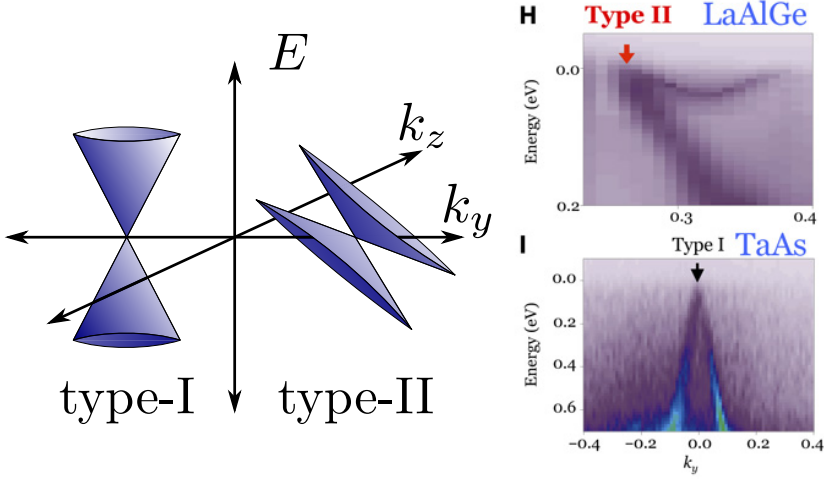
The term  $E_0(\vec{k})$  does not change the band gap and thus does not affect the cone position. However, it may distort the Weyl cone instead. In particular, the first-order contribution  $E_0(\vec{k}) \approx \vec{a} \cdot (\vec{k} - \vec{k}_0)$  tilts the cone in the direction  $\vec{a}$  [51–59]. (One should note that this ‘tilting’ involves making the cone broader, rather than rotating it in the  $\vec{a} - E$  plane. The latter would eventually fix the number of bands at a point  $\vec{k}$  to be  $\neq 2$ , which implies they can no longer be generated by a  $2 \times 2$  Hamiltonian.) When  $\|\vec{a}\|^2 > \|J\vec{a}\|$ , this tilting pushes both bands below the Weyl point in the  $\vec{a}$  direction (and both above the Weyl point in the  $-\vec{a}$  direction - see Fig. 1.2). This *type-II* Weyl cone has a non-trivial Fermi surface, which is conical when the chemical potential is fixed to the Weyl point (as opposed to the Fermi surface of an un-tilted Weyl cone, which is a point). In Chapter 2 we study the effect this Fermi surface might have in magnetic transport experiments.

A topological invariant that captures the above physics is the existence or non-existence of the Weyl point. However, one is not generally presented with a Hamiltonian of the form Eq. 1.7. Instead, to construct this invariant rigorously, one may integrate the Berry curvature [61] over a closed surface containing the Weyl point, to give the *chiral charge* [154]. If  $|\Psi_+(\vec{k})\rangle$  is the ground state of  $\hat{H}_{\text{Weyl}}$ , this is defined as

$$\chi = \frac{i}{2\pi} \oint_{\mathcal{K}} \cdot \nabla_{\vec{k}} \times \langle \Psi(\vec{k}) | \nabla_{\vec{k}} | \Psi(\vec{k}) \rangle \cdot d\vec{\mathcal{K}} \quad (1.8)$$

---

\*We choose to represent the Pauli matrices as  $\sigma_x$ ,  $\sigma_y$ ,  $\sigma_z$  here as this is more common in the literature, and as we use the same notation in Part I.



**Figure 1.2.:** Bandstructure for tilted (or type-II) and untilted (or type-I) Weyl cones. (left) Weyl cone cartoon, showing a two-dimensional cut through the bandstructure of type-I and type-II Weyl cones, where the type-II Weyl cone is tilted in the  $k_y$ -direction. (right) Experimental observation of type-I Weyl cones in TaAs (bottom) and type-II Weyl cones in LaAlGe (top) via ARPES measurements. Right figure taken with permission from Ref. [60]

## 1. Introduction

The integrand here is divergence-free as long as  $|\Psi(\vec{k})\rangle$  is well-defined, so Gauss' theorem implies this integral is independent of the exact choice of  $\mathcal{K}$ , and that a non-zero  $\chi$  *must* imply a singularity at some point  $\vec{k}_0$  inside the surface.

One may evaluate the integral by substituting the wavevector  $\vec{k}' = J(\vec{k} - \vec{k}_0)$ , which sets  $\hat{H}_{\text{Weyl}} = \vec{\sigma} \cdot \vec{k}'$ . This sends  $\nabla_{\vec{k}} \rightarrow J\nabla_{\vec{k}'}$ , and  $d\vec{K} \rightarrow J^T \det(J)^{-1} d\vec{K}'$ , which cancel out in the integral. If  $\vec{k}'$  is written in polar co-ordinates  $(k_r, k_\theta, k_\phi)$ , we have

$$|\Psi_+(\vec{k}')\rangle = \sin(k'_\theta/2)|0\rangle - \cos(k'_\theta/2)e^{ik'_\phi}|1\rangle. \quad (1.9)$$

Our integral is then

$$\begin{aligned} \chi &= \frac{-1}{2\pi} \oint_{\mathcal{K}} \nabla_{\vec{k}'} \times \left( 0, 0, \frac{\cos^2(k'_\theta/2)}{k'_r \sin(k'_\theta)} \right) \cdot d\vec{K} \\ &= \frac{1}{2\pi} \oint_{\mathcal{K}} \left( \frac{1}{2k_r'^2}, 0, 0 \right) \cdot d\vec{K}. \end{aligned} \quad (1.10)$$

We are free to choose our surface of integration, so let us take a unit sphere in  $\vec{k}'$ . However, the direction of the surface element matters - our surface element originally pointed out, and if  $\det(J) < 0$  this will be inverted under the transformation, so:

$$\chi = \frac{\text{sign}(\det(J))}{2\pi} \int_0^{2\pi} dk'_\phi \int_0^\pi \frac{1}{2} \sin(k'_\theta) dk'_\theta = \text{sign}(\det(J)) \quad (1.11)$$

That the chiral charge  $\chi$  is signed is no accident. We observed that  $\chi$  may only be affected by passing Weyl points (or other singularities) through the surface of integration  $\mathcal{K}$ , and so if it were absolute, it would be impossible to perturb a bandstructure to create a Weyl point in the first place. Instead, two Weyl points of opposite chiral charge may combine and annihilate to create an entirely gapped spectrum [154]. Importantly, if one chooses  $\mathcal{K}$  to be the Brillouin zone of a periodic system, the integration cancels [154], implying that the sum charge of all Weyl points in a bandstructure must be 0. This is important, as under applied electromagnetic fields, single Weyl cones produce particles or holes, depending on their sign [62, 63]. This is known as the chiral magnetic effect, and must cancel out to conserve charge [64, 65]. In Chapter 3, we will explore how one can recover this as a charge current in a superconductor, where this conservation law no longer applies.



## 1.3. Quantum error correction

The primary reason why quantum computers have not yet surpassed their classical counterparts is due to the fragility of quantum states. As a quantum computation proceeds, the physical system accumulates errors, which are unwanted operations on the quantum state. Accumulation of these operations will destroy any information contained in the quantum state. This necessitates the spending of additional computational resources to detect and correct errors as they accumulate.

The above is true also of classical computers, and the field of quantum error correction finds its roots in the field of error correction (now classical error correction) that began with the advent of the digital computer. In both fields, error correction is applied by repeating three steps:

1. Errors are accumulated over a short piece of the computation.
2. Errors are diagnosed by probing the system.
3. Errors are either corrected, or kept track of for latter correction if this is lower cost.

This procedure is not usually perfect; ‘correcting’ a particularly noisy computation may lead to a remaining *logical* error. One wishes to be able to make the *logical error rate* arbitrarily small. It is generally accepted that this is achievable if the error rate can be made exponentially small in size of the additional resources required to perform error correction [66–68].

Four main complications occur when shifting error correcting techniques to quantum devices.

1. Quantum computers are more noisy than their classical counterparts. Error rates in modern classical computers are now of the order of  $10^{-14}$  per operation, driven by interference from cosmic rays [69]. By comparison, state-of-the-art error rates in quantum computers are between 0.5 and 5 percent [212]. As these error rates are somewhat fundamental - classical bits are by nature less fragile than qubits - it should be expected that quantum computers will always rely more heavily upon error correction than their classical counterparts.
2. Measurement of the quantum state causes its collapse. If the measurements to diagnose errors on a quantum computer were to gain information about the desired logical state, they would be as bad for the quantum system as the errors themselves. This makes error

## 1. Introduction

detection on a quantum computer far more costly than on a classical computer, as it must be designed to extract only information pertaining to errors.

3. Quantum states are analog and not digital. This means that quantum states can be subject to the gradual accumulation of noise over a period of time, in comparison to a classical bit which is only ever discretely flipped by errors. This gradual accumulation of error is a lethal problem for classical analog computers [66]; how can one determine whether a process  $x \rightarrow x + \delta x$  has occurred or not? The fact that this may be avoided by quantum mechanics is one of the few desirable side-effects of the Born collapse for quantum computing.
4. Due to the no-cloning theorem [70–72], quantum information cannot be copied. This prevents any error correction by making redundant copies of quantum states, which necessitates thinking about error correction in a different manner.

The above issues made quantum computing appear nigh-impossible during the 90s. Though one could obtain a quantum speedup for certain tasks, it appeared that one would need exponential resources to achieve this [20, 21], negating any advantage. A breakthrough occurred when it was realised that one could choose non-local degrees of freedom on which to store quantum information [238]. Then, if error processes are local and uncorrelated, unwanted operations that do not perturb any local degrees of freedom are suppressed exponentially. If a local error occurs, the perturbed local degrees of freedom may be measured to diagnose and correct the error without disturbing the protected quantum information. In particular, this local measurement discretizes a local error into either having occurred or having not occurred, preventing its gradual accumulation as described above.

### 1.3.1. Toy example - the repetition code

A comparison of quantum and classical error correction may be demonstrated through the repetition code. This is a toy example as it cannot correct all local errors on a quantum computer. Nevertheless, it demonstrates the means by which quantum error correction may be performed despite issues (2-4) above. (Issue (1) does not threaten a critical failure of error correction, but causes an increase in the resource requirements for QEC.)

### 1.3. Quantum error correction

In the classical  $n$ -bit repetition code, a bit-worth  $x_L = 0, 1$  of information is copied over  $n$  physical bits to give one of two *code words*  $x_L x_L \dots x_L$ . Classical errors accumulate to flip  $m$  physical bits, resulting in the system state no longer being a code-word. One may diagnose the errors via a ‘majority vote’; choosing to return to either  $00 \dots 0$  or  $11 \dots 1$  by the least number of flips. If  $n$  is odd, this choice is always well defined, and a logical error occurs if  $m \geq (n+1)/2$ . If errors accumulate stochastically at a uniform rate  $p$  on each logical bit, the chance of a logical error is then, to lowest order in  $p$

$$p_L^{(n)} = \binom{n}{(n+1)/2} p^{(n+1)/2} (1-p)^{(n-1)/2}. \quad (1.12)$$

To see that this decreases exponentially with  $n$ , we calculate

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{p_L^{(n+2)}}{p_L^{(n)}} &= \lim_{n \rightarrow \infty} \frac{(n+2)!((n+1)/2)!((n-1)/2)!}{n!((n+3)/2)!((n+1)/2)!} p(1-p) \\ &= \lim_{n \rightarrow \infty} \frac{4(n+2)(n+1)}{(n+3)(n+1)} p(1-p) = 4p(1-p). \end{aligned} \quad (1.13)$$

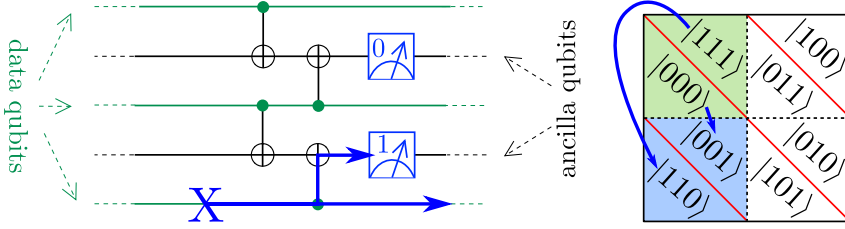
We see that this is less than 1 for all  $p \in [0, 1]$ , exponential error suppression is achieved.

Though similar to the above, the quantum version of the repetition code requires considerably different description to its classical counterpart. A schematic of the Hilbert space of the  $n = 3$ -qubit repetition code is given in Fig. 1.3. One cannot arbitrarily copy a quantum state  $a_0|0\rangle + a_1|1\rangle \rightarrow (a_0|0\rangle + a_1|1\rangle)^{\otimes n}$  (issue (4)). However, it is possible to copy the basis states to codeword states:

$$a_0|0\rangle + a_1|1\rangle \rightarrow a_0|00 \dots 0\rangle + a_1|11 \dots 1\rangle. \quad (1.14)$$

Then, measuring any single physical qubit is sufficient to cause the collapse of the state (issue (2)). However, coherent measurements of the operators  $Z_i Z_{i+1}$  do not cause any such collapse, as the logical state is an eigenstate of each such measurement. A bit flip error on qubit  $i$  is uniquely identifiable, as it shifts the state into a new eigenspace of the measurements  $Z_i Z_{i+1}$  (if  $i < n$ ) and  $Z_{i-1} Z_i$  (if  $i > 1$ ). This measurement further discretizes small coherent errors (issue (3)); if say, qubit 1 is rotated by a small error  $e^{i\delta X_1}$ ,

## 1. Introduction



**Figure 1.3.:** The three qubit repetition code. (left) a circuit to measure the two stabilizers of the code using ancilla qubits. Any X error propagates onto a unique combination of the ancilla qubits, and is detected as shown. (right) The physical qubit Hilbert space split into the four possible choices for the logical Hilbert space. Weight-1 errors may shift the qubit across the dashed lines, but crossing the red lines requires a weight-3 operation. Following the shown X error and its detection, a logical state prepared in an initial (green) subspace will be coherently shifted to the blue subspace, whilst preserving any quantum information encoded between the two logical states. Z errors, however, create undetectable logical errors by causing the two logical states to pick up a phase, making this not a true quantum error correcting code. Figure taken from Chapter 5 with significant edits.

the system state evolves to

$$\cos(\delta)(a_0|00\dots 0\rangle + a_1|11\dots 1\rangle) + i\sin(\delta)(a_0|10\dots 0\rangle + a_1|01\dots 1\rangle), \quad (1.15)$$

and measurement of  $Z_1 Z_2$  either projects the system to a state where no error has occurred (with probability  $1 - \delta^2$ ), or to a state where a discrete bit-flip has occurred (with probability  $\delta^2$ ). If these are the only errors, this detection is sufficient for diagnosis in a very similar manner to the classical case, which we investigate in Chapter 5. However, error modes in realistic quantum computers are not usually restricted to rotations around a single axis. Any error of the form  $e^{i\theta Z_i}$  on any qubit will rotate the logical state, and so the repetition code enhances the possibility of these errors, rather than reducing them. It turns out that this is not systemic in quantum computing, but just a feature of the repetition code itself. In the next sections, we will describe codes that can detect and diagnose all local errors, and thus perform true quantum error correction.

### 1.3.2. Stabilizer codes

Throughout the rest of this section, we will refer to measurements made on a device with capital letters, and transformations that occur on a device in lower case. This implies in particular to the Pauli matrices  $\{\sigma_0, \sigma_x, \sigma_y, \sigma_z\} \equiv \{\mathbb{I}, x, y, z\} \equiv \{\mathbb{I}, X, Y, Z\}$ , as on set of mathematical objects can refer to both transformations and measurements. This is an important distinction; if a system is in an eigenstate of  $Z$ , then after performing (the transformation)  $x$  it remains in an eigenstate (though it is no longer the same eigenstate). By contrast, after ‘performing’ (the measurement)  $X$ , the system is no longer in an eigenstate of  $Z$ , as  $[X, Z] = 0$ .

A large class of important quantum error correcting codes are stabilizer codes. This class was identified following the theorem, due to Gottesman [68], that if a QEC code can correct unitary errors  $\hat{e}_1, \hat{e}_2$ , it can correct any quantum channel  $\Lambda$  that can be written in terms of with Kraus matrices that are linear combinations of  $\hat{e}_1$  and  $\hat{e}_2$ . Then, as  $\Pi^N = \{\sigma_0, \sigma_x, \sigma_y, \sigma_z\}^{\otimes N}$  spans  $\mathbb{C}^{2^N}$ , one may rewrite the problem of correcting arbitrary CPTP maps on  $N$  qubits to one of correcting elements of  $\Pi^N$ . The latter set, being discrete and finite, is much more tractable than the former.

An  $[N, K]$ -stabilizer code is defined by a set  $S$  of  $N - K$  Hermitian stabilizer measurements  $\hat{S}_j \in \Pi^N$  that satisfy the following two properties [68]:

1.  $[\hat{S}_j, \hat{S}_k] = 0$ .
2.  $\pm \hat{S}_j$  cannot be generated by the other  $\hat{S}_k$ .

The code space  $\mathcal{H}_L$  is then the space of all  $+1$  eigestates of the  $\hat{S}_j$ , which exists due to the first property. The code space is stabilized by the group  $\mathcal{S}$  generated by the  $\hat{S}_j$

$$\mathcal{S} := \{\hat{S}_1 \hat{S}_2 \dots, \hat{S}_1, \hat{S}_2, \dots \in S\} \equiv \{\hat{O} \in \Pi^n, \hat{O}|\psi\rangle = |\psi\rangle \forall |\psi\rangle \in \mathcal{H}_L\} \quad (1.16)$$

Each stabilizer cuts the  $N$ -qubit Hilbert space  $\mathbb{C}^{2^N}$  into two eigenspaces of equal size. The second stabilizer property implies that cuts from subsequent stabilizers continue this division [68], and the code space thus has dimension  $2^{N-(N-K)} = 2^K$ . This may then be identified as a space containing  $K$  logical qubits.

Now, suppose we have a state  $\rho \in \mathcal{H}_L$  in our code space. An error

## 1. Introduction

channel  $\Lambda = \{\hat{\Lambda}_i\}$  sends  $\rho$  to

$$\sum_i \hat{\Lambda}_i^\dagger \rho \hat{\Lambda}_i. \quad (1.17)$$

Let us further decompose each  $\hat{\Lambda}_i$  into Paulis  $\hat{e}_j \in \Pi_n$  -  $\hat{\Lambda}_i = \sum_j \Lambda_{i,j} \hat{e}_j$ . If a subsequent measurement of  $\hat{S}$  produces a reading of +1, the state is projected into

$$\frac{1}{N} (1 + \hat{S}) \sum_{i,j,k} \Lambda_{i,j} \hat{e}_j \rho \hat{e}_k \Lambda_{k,i}^* (1 + \hat{S}), \quad (1.18)$$

with  $N$  a normalization constant. Next, note that each  $\hat{e}_j$  either commutes or anti-commutes with  $\hat{S}$ . If  $\{\hat{S}, \hat{e}_j\} = 0$ , we have  $(1 + \hat{S})\hat{e}_j \rho = 0$ , as  $\rho$  was in the +1 eigenstate of  $\hat{S}$ . Thus, the only terms that appear in Eq. 1.18 are those of operators that commute with  $\hat{S}$ .

If subsequent measurements of all stabilizers returns readings of +1, this effect cumulates, and the only terms remaining in Eq. 1.18 are those commuting with all stabilizers. Furthermore, if  $\hat{e}_j$  is generated by the stabilizers,  $\hat{e}_j \rho = \rho$ , and  $\hat{e}_j$  acts as the identity. Thus, the only non-trivial action of the error channel on  $\rho$  comes from the set of *pure errors*

$$\mathcal{E} = \{\hat{e} \in \Pi^n, \hat{e} \notin \mathcal{S}, [\hat{e}, \hat{S}] = 0 \ \forall \hat{S} \in \mathcal{S}\}. \quad (1.19)$$

The resulting state is then

$$\rho_0 = \frac{1}{N} \sum_i \Lambda_i^0 \rho \Lambda_i^{0\dagger} \quad (1.20)$$

where

$$\Lambda_i^0 = \left( \sum_{\hat{e}_j \in \mathcal{S}} \Lambda_{i,j} + \sum_{\hat{e}_j \in \mathcal{E}} \Lambda_{i,j} \hat{e}_j \right), \quad N = \text{Trace} \left[ \sum_i \Lambda_i^0 \rho \Lambda_i^0 \right]. \quad (1.21)$$

If  $\Lambda_{i,j} = 0$  for all pure errors  $\hat{e}_j \in \mathcal{E}$ ,  $\rho_0 = \rho$ , and the error channel has no effect (when the stabilizer measurements all return +1).

Errors on quantum computers are generally local, as the underlying physical processes are local. Assuming that errors are Markovian (i.e. uncorrelated in time), an error channel acting between time  $t_i$  and  $t_f$  may be approximated as  $\Lambda = \prod_{t=t_i}^{t_f} \Lambda(t)$ . Locality implies here that terms  $\hat{\Lambda}(t)_{i,j}$  in the Kraus decomposition are 0 unless  $\hat{e}_j$  is distance  $k$  (has  $k$

### 1.3. Quantum error correction

non-trivial tensor factors) or less for some  $k$ . Then, suppose that no pure error  $\hat{e}_j \in \mathcal{E}$  is less than  $d$ -local, for some *code distance*  $d$ . Pure errors in the full channel  $\Lambda$  must be generated by products of no less than  $d/k$  non-trivial terms from each  $\Lambda(t)$ . In general, these terms will be of a magnitude  $p_{\text{phys}} \ll 1$ , and the size of the pure error in  $\Lambda$  will scale proportional to  $p_{\text{phys}}^{d/k}$ . The logical error rate  $\epsilon_L$  is then also exponentially suppressed in the code distance  $d$ .

The above demonstrates the ability to exponentially suppress errors in the situation where none are detected. One must also consider the situation where stabilizer measurements do not return  $+1$ . Repeating the previous reasoning, the state  $\rho$  is then mapped to

$$\rho_{\vec{s}} = \frac{1}{N} \sum_i \Lambda_i^{\vec{s}} \rho \Lambda_i^{\vec{s}\dagger}, \quad (1.22)$$

where the *syndrome*  $\vec{s} \in \mathbb{Z}_2^{N-K}$  contains a 1 for each  $\hat{S}_j$  that measured  $-1$  (and a 0 otherwise), and

$$\Lambda_i^{\vec{s}} = \left( \sum_{\hat{e}_j \in E_{\vec{s}}} \Lambda_{i,j} \hat{e}_j \right), \quad E_{\vec{s}} = \{\hat{e} \in \Pi^N, \{\hat{e}, \hat{S}_j\} = 0 \text{ iff } s_j = 1\}. \quad (1.23)$$

There exists a minimum distance  $d_{\vec{s}}$  for errors in  $E_{\vec{s}}$ , so we may choose some  $\hat{e}_{\vec{s}} \in E_{\vec{s}}$  with distance  $d_{\vec{s}}$ . Applying this error to  $\rho_{\vec{s}}$  ‘corrects’ it, sending it back to the code space  $\mathcal{H}_L$ . Lowest-order terms in the resulting pure error must be a product of the correction  $\hat{e}_{\vec{s}}$  and an error  $\hat{e}_{\text{other}} \in E_{\vec{s}}$  of distance  $(d - d_{\vec{s}})$ . As  $\hat{e}_{\vec{s}}$  was chosen to be a lowest-distance element of  $E_{\vec{s}}$ , we have that  $(d - d_{\vec{s}}) \geq d_{\vec{s}}$ . In particular, for odd  $d$ ,  $(d - d_{\vec{s}}) \geq (d + 1)/2$ .  $\hat{e}_{\text{other}}$  must then have been generated by  $(d + 1)/(2k)$  products of physical errors, and the error rate in turn should scale as

$$\epsilon_L \propto p_{\text{phys}}^{(d+1)/2k}. \quad (1.24)$$

Note that this implies that (when  $d$  is odd) *all* errors in  $E_{\vec{s}}$  of minimal distance  $d_{\vec{s}}$  are equivalent - the product of two such errors acts as the identity on  $\mathcal{H}_L$ , and is therefore a stabilizer.

The above is only a sketch of a proof that errors may be exponentially suppressed at a polynomial resource cost. One must also confirm that it is possible to create stabilizer codes of arbitrary distance  $d$  with polynomial resources. This may be done by construction; many families of such codes exist [74, 75, 207, 238], and codes may be additionally created from parity

## 1. Introduction

codes of classical error correction [73].

One must also confirm that the proportionality factor in  $\epsilon_L \propto p_{\text{phys}}^{(d+1)/2k}$  (which is combinatorial in the number of qubits) does not overwhelm the exponential suppression. This is highly non-trivial, and we refer the reader to [66] for further discussion of this point. In general, one requires  $p_{\text{phys}}$  to be smaller than some *threshold* error rate [76] for such correction. However, thresholds for classical codes may be as high as  $p_{\text{phys}} = 50\%$ , while the highest-known thresholds for quantum error correction are around  $p_{\text{phys}} = 1\%$  (with  $p_{\text{phys}}$  being the average error rate per physical gate or measurement).

### 1.3.3. Stabilizer code dynamics

Traditionally, stabilizer error correction is broken up into cycles, each cycle containing an error accumulation step, a stabilizer measurement step, and a correction step [68]. More recently, it has been realised that the correction step is often unnecessary. The results of the previous section hold regardless of the chosen eigenvalue of the stabilizers;  $+1$  is as good as  $-1$ . For a syndrome  $\vec{s}$  of a  $d$ -distance code, we can define the syndrome space

$$\mathcal{H}_{\vec{s}} = \{|\psi\rangle \in \mathcal{H}_{\text{phys}}, \hat{S}_j|\psi\rangle = -1^{s_j}|\psi\rangle\}, \quad (1.25)$$

and then the only Pauli operators that send  $\mathcal{H}_{\vec{s}} \rightarrow \mathcal{H}_{\vec{s}}$  are precisely the pure errors  $\mathcal{E}$ . This implies that the logical error rate has the same scaling in  $d$  regardless of which code space is used. (For most codes we expect the logical error rate to be identical for all  $\mathcal{H}_{\vec{s}}$ , but this is not necessarily the case.)

Then, let us consider  $t = 1, \dots, T$  of error correction of a state starting in  $\mathcal{H}_L$ , giving syndromes  $\vec{s}(t)$  (and defining  $\vec{s}(0) = \vec{0}$ ). If one can choose for each  $t$  a minimal-weight error  $\hat{e}_{\vec{s}(t-1), \vec{s}(t)}$  that maps  $\mathcal{H}_{\vec{s}(t)} \rightarrow \mathcal{H}_{\vec{s}(t-1)}$ , the product  $\hat{F}(t) = \prod_{s=1}^t \hat{e}_{\vec{s}(t-1), \vec{s}(t)}$  will take the final state back to the logical Hilbert space. This product  $\hat{F}(t)$  is known as a Pauli frame. A Pauli frame does not need to be the shortest-distance error mapping  $\mathcal{H}_{\vec{s}(t)} \rightarrow \mathcal{H}_L$ , as it has to respect the path that the state took in time. After each cycle  $t$ , the state accumulates a pure error of size proportional to  $p_{\text{phys}}^{d-d(t)}$ , where  $d(t)$  is the distance of  $\hat{e}_{\vec{s}(t-1), \vec{s}(t)}$ . This is bounded above by  $p_{\text{phys}}^{(d+1)/2}$ , as required.

With the logical Hilbert space  $\mathcal{H}_L$  defined, one may ask what the logical operators acting on  $\mathcal{H}_L$  are in terms of the operators on the physical Hilbert space. This choice has a large amount of freedom, as we are



technically not concerned with the behaviour of an operator  $\hat{O}$  on any other part of the Hilbert space. However, following the above idea that all syndrome spaces  $\mathcal{H}_{\vec{s}}$  should be treated equivalently, one wishes to choose a set of operators that does so. Furthermore, one wishes to choose operators that are not significantly altered by local errors. The most natural such choice is to draw from the set of pure errors  $\mathcal{E}$ . Each element of  $\mathcal{E}$  is block diagonal on the syndrome spaces  $\mathcal{H}_{\vec{s}}$ , and when combined with the stabilizer operators, this forms a basis for all such block diagonal operators. Individual elements  $\hat{O}_L \in \mathcal{E}$  have the following advantage: lowest-distance errors mapping  $\mathcal{H}_{\vec{s}(t)} \rightarrow \mathcal{H}_{\vec{s}(t+1)}$  either all commute with  $\hat{O}_L$  or all anticommute with  $\hat{O}_L$  (as the product of two such errors is a stabilizer, which necessarily commutes with  $\hat{O}_L$ ). This implies that the difference between applying  $\hat{O}_L$  before or after such an error is simply to apply a global sign change to the state  $\rho$ , which is irrelevant.

The above does not entirely fix the freedom in choosing logical operators; given a stabilizer  $\hat{S} \in \mathcal{S}$ ,  $\hat{S}\hat{O}_L$  acts identically to  $\hat{O}_L$  on every syndrome space. This is a gauge degree of freedom; it is entirely at the users discretion which logical operators to choose. The only requirement is that the choice of operators give a map from  $\Pi^K \rightarrow \Pi^N$  that respects the algebraic structure of the logical operators (a homomorphism).

### 1.3.4. Topological quantum error correcting codes

An obvious motif for designing quantum error correcting codes is topology. The general notion here is to generate a stabilizer group  $\mathcal{S}$  via nearby Pauli operators on a  $D$ -dimensional qubit lattice such that the lowest-distance pure errors occupy  $D - 1$ -dimensional boundaries. If  $D > 1$ , the size of these boundaries grows with the system size, and so does the code distance. This is only true however if the pure errors extend across the entire boundary, which in general occurs if the dimension  $K$  of the logical Hilbert space  $\mathcal{H}_L$  is constant in the system size. Such a code is known as a topological quantum error correcting code, or a topological code for short. These codes are of interest as they require only local stabilizer measurements, and can be scaled without any change to the local structure of the code, allowing for modular designs [77].

In this work, we focus on the study of  $D = 2$  topological codes. Higher dimensional codes are an active area of research, and have some advantages over two-dimensional codes [78, 174]. However, most current quantum hardware is built on a two-dimensional architecture, and the challenges of making vertical interconnects or long-range couplings usually presents too great a challenge to overcome in the near-term. Two-dimensional codes

## 1. Introduction

also have the highest known thresholds [177], which are highly attractive given current noise limitations [212].

Probably the most well-studied topological code is the 2-dimensional surface code, first introduced in [207]. This code is defined on a checker-board of  $d - 1 \times d - 1$  squares of alternating colour (Fig. 1.4). Qubits are placed at the corners of each square, forming a  $d \times d$  lattice. Each square on the board defines a stabilizer measurement, being the tensor product of either  $Z$  or  $X$  on the four surrounding qubits. One can observe that neighbouring squares of different stabilizer type share two qubits, and thus the stabilizer measurements commute as required. Every square edge around the perimeter of the surface gives a remaining local degree of freedom (either  $ZZ$  or  $XX$ ). This can be pinned by making every second such degree of freedom a stabilizer (noting that this pins the alternating degrees of freedom, which do not commute) [79, 176, 180]. This pinning fixes one pair of edges to be  $Z$ -edges, and the other  $X$ -edges. This leaves no local pure errors, but the total number of stabilizer measurements can be calculated as  $(d - 1)^2 + 2d = d^2 - 1$ . The remaining pure errors can be constructed by drawing a path  $\lambda$  along the board edges from one  $X(Z)$ -edge to another, and taking the tensor product of  $X(Z)$  on each qubit crossed, such that  $\lambda$  crosses an even number of vertices on every  $Z(X)$  stabilizer. The freedom in choosing the path is the gauge degree of freedom in the code discussed earlier; one may deform between any two paths by applying some combination of stabilizer operators. As the minimum-length path between the two edges is a straight one, a surface code on a  $d \times d$  lattice will have a distance  $d$ .

A second well-known class of topological codes are color codes, first studied in [75]. In two dimensions, color codes may be generically defined on any tiling of a sphere, given that: [228]

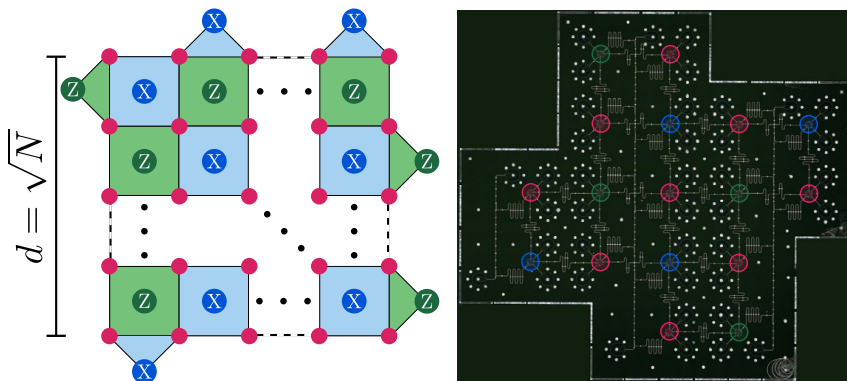
1. Every vertex has exactly 3 edges leading to it.
2. The tiling is 3-colourable.

A possible such tiling is shown in Fig. 1.5. As the Euler characteristic of the sphere is 2, and the number of edges  $n_e$  is  $3/2$  the number of vertices  $n_v$ , this gives a tiling with  $n_v$  qubits and

$$n_f = 2(3/2n_v - n_v + 2) = n_v + 4 \quad (1.26)$$

faces.

If one then deletes a single qubit and the three faces associated, the tiling may be flattened onto the plane. To each vertex we associate a qubit, and

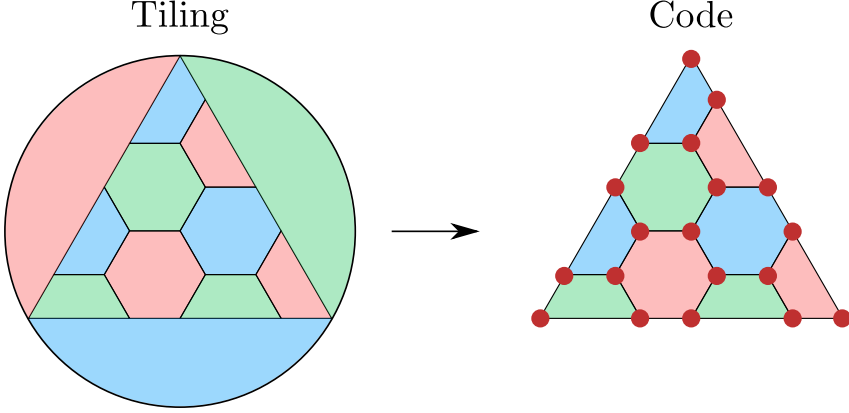


**Figure 1.4.:** Schematic of the surface code and an experimental implementation. (left) The surface code is defined by fixing a  $d \times d$  lattice of data qubits, and colouring the squares in the formed lattice with a checkerboard pattern. This colouring gives the stabilizers of the code, which are four-fold  $XXXX$  and  $ZZZZ$  parity checks on the qubits surrounding any given square. Two-fold  $XX$  and  $ZZ$  checks are performed on the edges, leaving a single logical degree of freedom. (right) A prototype implementation of the distance-3 surface code, Surface-17, on a superconducting transmon architecture, from TU Delft. Transmons are circled and color-coded as per the schematic on the left. Photo credit: Alessandro Bruno.

to each face we associate two stabilizers, being the product of all  $X$  and all  $Z$  on surrounding vertex qubits. This gives a total of  $N = n_v - 1$  physical qubits and  $N - K = n_v - 2$  stabilizer measurements. It is non-trivial to check that the stabilizer measurements do not generate each other, but this is indeed the case [80, 228], and the resulting code encodes a single logical qubit. The logical  $X_L(Z_L)$  operators in the resulting code are the product of individual  $X(Z)$  Pauli operators on the edges of the excised faces.

Though similar, some key differences exist between the two codes described above. The most obvious is that of qubit number: a distance- $d$  color code requires roughly  $3/4$  times the number of physical qubits for a distance- $d$  surface code [79]. A second important concern is that of the threshold physical error rate  $p_t$ . With perfect stabilizer measurements, both codes have equal thresholds [79] of around 12%. However, this is not the case when noisy measurements and decoding are considered, as will be discussed in the next sections: the surface code significantly outperforms the color code. A final distinction is the ease at which one may

## 1. Introduction



**Figure 1.5.:** (right) A tiling of the sphere for the color code. (left) The reduction to a distance-5 encoding of a single logical qubit. Dots correspond to data qubit location, and coloured tiles correspond to stabilizer measurements (the type of measurement is independent of the colour). Taken from Chapter 6 with significant changes.

perform logical operations. One typically desires logical operations to be transversal; to be able to be performed via either single-physical-qubit rotations or two-qubit gates between corresponding physical qubits in the case of logical two-qubit gates. On the color code, all Clifford gates may be performed transversally [228], which is the best possible outcome for any 2-dimensional code. By contrast, the surface code may only perform Pauli operations, the Hadamard gate, and the CNOT gate transversally, which does not generate the entire Clifford set. Although this makes the color code possibly more attractive in the long-term, the ease of decoding the surface code, and its large threshold  $p_t$ , have made it a more typical choice for current QEC experiments [77, 173, 212].

### 1.3.5. Decoding topological codes

The task of taking cycles of syndrome data  $\vec{s}(t)$  and calculating a Pauli frame  $\hat{F}(t)$  to correct the system is performed by a classical algorithm called a *decoder*. Performing such a task optimally for a general error model on a general code is known to be NP-complete [81]. However, for some codes and some restricted error models, polynomial-time optimal decoding is possible [82]. Moreover, decoding need not be optimal. A suboptimal decoder  $D$  will have a reduced error rate  $\epsilon_L^{(D)} < \epsilon_L^{(\text{opt})}$ , but if  $\epsilon_L^{(D)}$  decreases

exponentially in the code distance  $d$ , this may be compensated for by increasing  $d$ . Of course, optimizing classical algorithms is in general cheaper than improving physical hardware. As such, ‘good enough’ for the decoder efficiency [212],

$$\eta_D = \epsilon_L^{(D)} / \epsilon_L^{(\text{opt})}, \quad (1.27)$$

really depends on who is asking the question.

Though a decoder is tasked with updating and returning a Pauli frame  $\hat{F}$ , this contains a lot of unnecessary information. The redundancy is due to a gauge degree of freedom: for all  $\hat{S} \in \mathcal{S}$ ,  $\hat{S}\hat{F} \equiv \hat{F}$ . This redundancy may be removed if one shifts to the Heisenberg frame, and considers the action of errors upon the set of logical Pauli operators. In particular, suppose one has an uncorrected state  $\rho \in \mathcal{H}_{\bar{s}}$ , a Pauli frame  $\hat{F}$ , and wishes to measure a logical Pauli operator  $\hat{O}_L$ . The probability of the post-corrected measurement returning  $+1$  is equal to

$$p^{(\text{EC})}(\hat{O}_L = +1) = \text{Trace} \left[ (\mathbb{I} + \hat{O}_L) \hat{F} \rho \hat{F}^\dagger \right]. \quad (1.28)$$

If  $[\hat{F}, \hat{O}_L] = 0$ , this is equal to the probability  $p(\hat{O}_L = +1)$  of measuring  $\hat{O}_L$  without performing such a correction, while if  $\{\hat{F}, \hat{O}_L\} = 0$ , this is equal to  $p(\hat{O}_L = -1)$ . Thus, no Pauli frame or correction is required for such a measurement. One may simply measure  $\hat{O}_L$ , and task the decoder with deciding the logical parity: whether the labels attached to the output dial should be  $\pm 1$  or  $\mp 1$ . Though many decoding algorithms inherently rely on a Pauli frame representation, shifting to a Heisenberg frame can be important for machine learning decoders, which tend to find the redundancy in a Pauli frame confusing [83, 211, 247]. A Heisenberg frame is also incredibly useful when one wishes to consider Clifford operations on logical qubits, as logical parities map directly through these transformations with no need to consider the action on the underlying qubits.

Algorithms for fast decoding of quantum error correcting codes tend to be code-specific. For the surface code, the problem of correcting independent  $x$  and  $z$  errors may be mapped to the minimum-weight perfect matching problem on a graph, leading to thresholds above 1% for Pauli channel noise models [173, 177]. In Chapter 4, we study how this mapping performs for realistic error models of superconducting transmon qubits, and in Chapter 5, we study how this mapping may be optimized by analysing the correlations in repeated syndrome measurements. Other codes do not unfortunately permit such a mapping. In lieu of this, much

## 1. Introduction

recent interest has gone into using machine learning to solve the decoding problem [211, 246–251, 253]. In Chapter 6 we demonstrate that this is a viable strategy for the color code, and study the performance of the resulting decoder.

## 1.4. Quantum algorithms

Although fault-tolerant quantum computers still seem a relatively distant fantasy, much preparation is still needed for their potential arrival. In particular, as the cost of quantum error correction is significantly high, algorithm optimization can pay dividends by reducing the resources needed to solve useful problems (and by finding what useful problems are indeed low-cost). Algorithm development on quantum computers is significantly different to that on classical computers, as quantum operations are unitary and continuous, so quantum coders have to work with a different set of building blocks to develop their code. In this section, we describe a few said building blocks.

### 1.4.1. Quantum phase estimation

Possibly the most important algorithm for quantum computing is that of quantum phase estimation [22, 84]. Let us assume we have a means of performing a unitary  $\hat{U}$  on  $N_{\text{sys}}$  qubits, conditional on an ancilla qubit (I.e., the unitary  $\mathbb{I} \oplus \hat{U}$ ). The unitary has eigenphases  $\phi_j$ , and corresponding eigenstates  $|\phi_j\rangle$

$$\hat{U}|\phi_j\rangle = e^{i\phi_j}|\phi_j\rangle, \quad (1.29)$$

Quantum phase estimation aims to estimate the value of a single target eigenphase  $\phi_t$ . To obtain a speedup over the best classical methods, this estimation needs an accuracy  $\epsilon \propto 1/\text{Poly}(N_{\text{sys}})$  in time  $T \propto \text{Poly}(N_{\text{sys}})$  with failure probability  $\delta \propto 1/\text{Poly}(N_{\text{sys}})$ . To achieve such speed, we require a means of preparing an initial state  $|\Psi_0\rangle = \sum_j a_j |\phi_j\rangle$  with  $|a_t|^2 \gg 0$  (otherwise this problem becomes QMA-hard [313]). Specifically, one may replace  $\text{Poly}(N_{\text{sys}})$  with  $\text{Poly}(N_{\text{sys}}/|a_t|^2)$  in the requirements for  $\epsilon$ ,  $\delta$  and  $T$ . With this stated, QPE is BQP-complete [292], meaning that a black box that solves the QPE problem is as good an oracle as a quantum computer itself. Furthermore, a large fraction of the quantum algorithms with a proven quantum speedup rely on QPE as a subroutine [84, 85], making QPE one of the key building blocks for the field of quantum algorithms.

Various circuits for quantum phase estimation exist, with different qubit overhead and classical computation requirements. However, the fundamental building block of all such circuits is the notion of phase kickback [23]. The result of acting  $\mathbb{I} \oplus \hat{U}$  on the state  $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \sum_j a_j |\phi_j\rangle$  may be written as

$$\frac{1}{\sqrt{2}} \sum_j a_j |\phi_j\rangle (|0\rangle + e^{i\phi_j} |1\rangle). \quad (1.30)$$

This can be interpreted as the ancilla qubit making a weak measurement of the system register, and imprinting the result of the measurement on its phase  $\phi_j$ . One then wishes to turn this weak measurement into a strong separation of phases, and to read the phases out to a required accuracy. Traditionally this is done by performing  $\mathbb{I} \oplus \hat{U}^{2^n}$  on a register of qubits  $n = 1, \dots, N_{\text{anc}}$ . If each ancilla is initialized in the  $|+\rangle$  state, the resulting state is

$$\frac{1}{\sqrt{2^{N_{\text{anc}}} - 1}} \sum_{k=0}^{2^{N_{\text{anc}}} - 1} \sum_j a_j e^{ik\phi_j} |k\rangle |\phi_j\rangle. \quad (1.31)$$

Then, applying the quantum Fourier transform [84],

$$|k\rangle \rightarrow 2^{-N_{\text{anc}}/2} \sum_{l=1}^{2^{N_{\text{anc}}} - 1} e^{i \frac{2\pi kl}{2^{N_{\text{anc}}}}} |l\rangle, \quad (1.32)$$

creates destructive interference on states  $|l\rangle |\phi_j\rangle$  unless  $l - \phi_j \approx 0$ . This then copies a  $N_{\text{anc}}$ -bit approximation of  $\phi_j$  to the ancilla qubit, which may be read out. This readout projects the system register to the state  $|\phi_j\rangle$  (assuming that  $|l - \phi_{j'}|$  is sufficiently large for  $j \neq j'$ ). More recently, it has been realised that one does not need to perform the Fourier transform on the quantum computer. Instead, if one performs  $\mathbb{I} \oplus \hat{U}^k$  conditional on a single ancilla qubit at various  $k$ , and tomographs the resulting phase  $g(k) = \sum_j e^{ik\phi_j}$  on the ancilla qubit, one may treat this as a classical waveform and extract its dominant frequencies. This can be shown to perform with the same accuracy as its multi-ancilla counterpart [297, 299, 316], although the system register is never projected to an eigenstate  $|\phi_j\rangle$  in any experiment. In Chapter 8, we study this in some detail, and in particular demonstrate that this performance remains when we are unable to otherwise prepare the system in an eigenstate  $|\phi_0\rangle$ .

## 1. Introduction

### 1.4.2. Variational quantum eigensolvers

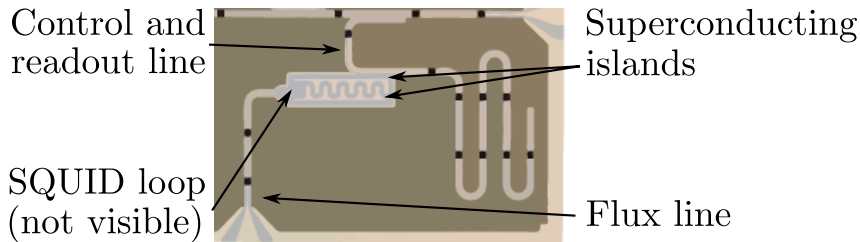
Though phase estimation and quantum error correction provide a means by which a quantum computer might eventually obtain an exponential advantage over classical computation, their large cost [263] has led to a hunt for lower-overhead algorithms to replace them. A class of such algorithms is that of variational quantum eigensolvers [86, 281], which aim to approximate the true ground state  $|\phi_0\rangle$  of an  $N_{\text{sys}}$ -qubit Hamiltonian  $\hat{H}$ . These algorithms consist of a starting state  $|\psi_0\rangle$ , and a unitary  $\hat{U}(\vec{\theta})$  parametrized by a set of continuously-tunable classical control parameters  $\vec{\theta}$ . Setting these parameters allows one to construct an approximation to  $|\phi_0\rangle$  and measure its energy, by defining

$$|\psi(\vec{\theta})\rangle := \hat{U}(\vec{\theta})|\psi_0\rangle, \quad E(\vec{\theta}) := \langle\psi(\vec{\theta})|\hat{H}|\psi(\vec{\theta})\rangle. \quad (1.33)$$

$E(\vec{\theta})$  may be measured by tomographing  $|\psi(\vec{\theta})\rangle$ . By tomographing, we mean writing  $\hat{H} = \sum_P H_P \hat{P}$  for Pauli operators  $\hat{P} \in \Pi^{N_{\text{sys}}}$ , repeatedly preparing  $|\psi(\vec{\theta})\rangle$  and measuring  $\hat{P}$ , which has a probability  $\frac{1}{2}(1 + \langle\hat{P}\rangle)$  of observing 0. One may then infer  $\langle\hat{P}\rangle$  from  $N$  measurements with error scaling as  $\frac{1}{\sqrt{N}}$ , and sum  $\langle\hat{H}\rangle = \sum_P H_P \langle\hat{P}\rangle$ . The final result  $E(\vec{\theta})$  is then returned to a classical control script, which minimizes  $E(\vec{\theta})$  by tuning  $\vec{\theta}$ . The resulting energy,  $\min_{\vec{\theta}} E(\vec{\theta})$  is then taken as an approximation to the ground state energy  $E_0 = \langle\phi_0|\hat{H}|\phi_0\rangle$ , and the state  $|\Psi(\vec{\theta})\rangle$  as an approximation to  $|\phi_0\rangle$ .

We are guaranteed that  $\min_{\vec{\theta}} E(\vec{\theta}) \geq \langle\phi_0|\hat{H}|\phi_0\rangle$  by the variational principle [86], but the error in the estimation is unbounded. At best, a variational quantum eigensolver explores a  $\dim(\vec{\theta})$ -dimensional submanifold of the entire Hilbert space. This takes  $\text{Poly}(\dim(\vec{\theta}))$  time, making choosing sufficient parameters to explore the entire Hilbert space unfeasible. Without this, we have no guarantee that  $|\phi_0\rangle$  lies in or nearby our chosen manifold, and one must appeal to some physical intuition (e.g. perturbative approaches [86]) that this would be the case. Given current short coherence times, an alternative approach is to choose those variational quantum eigensolvers that require the least quantum circuitry to implement per dimension of  $\vec{\theta}$  [87]. In Chapter 7, we suggest that Majorana hardware may be able to combine the best of both worlds, as physical operations on fermions are close in nature to the excitations required for perturbative approaches.





**Figure 1.6.:** Labeled photograph of a transmon qubit. The flux line adjusts the magnetic field running through the SQUID loop, which in turn changes the Josephson energy, allowing for the qubit frequency to be adjusted. Single qubit gates and readout may be performed via microwave pulses through the line labeled ‘control and readout’. Two-qubit gates are achieved by similar lines coupling multiple qubits (none shown here), which provide a constant two-qubit interaction. Fig taken from Ref. [182].

## 1.5. Quantum computing with superconducting qubits

In chapters 4,8,and 6, we have worked on quantum algorithms and quantum error correction for near-term devices, necessitating connection to experiment via simulation of real devices. One of the most promising and popular physical platforms for quantum computing is that of superconducting transmon systems [88]. In this section we briefly outline the construction of such a quantum device, and the leading source of errors during algorithm execution. More in-depth details are given in Ch. 4, and relevant numbers for state-of-the-art devices are also stated here.

### 1.5.1. Transmon architecture

A superconducting transmon system is constructed by lithographing two superconducting parallel plates on a substrate, and connecting them via a SQUID loop (see Fig. 1.6). A charge difference between the two plates creates a capacitive energy (with the energy scale  $E_C = e^2/2C$  determined by the capacitance  $C$  of the system). Similarly, a difference in the superconducting phase on the two plates creates an inductive energy (with the energy scale  $E_J = E_{J,\max} |\cos(\pi\Phi/\Phi_0)|$  set by the flux  $\Phi$  passing through the SQUID loop). Quantum-mechanically, charge and superconducting phase are conjugate variables (as the phase is obtained from differentiating the superconducting condensate wavefunction). One

## 1. Introduction

may write the system's Hamiltonian as [88]

$$\hat{H} = 4E_C(\hat{n} - n_g)^2 - E_J \cos(\hat{\varphi}), \quad (1.34)$$

where  $\hat{n}$  measures the charge difference between the two plates, and  $\hat{\varphi}$  measures the phase difference between the two plates. In the  $E_J \gg E_C$  limit, this approximates an anharmonic oscillator

$$\hat{H}_{\text{transmon}} = -\frac{\hbar^2}{2m} \partial_x^2 + \frac{k}{2} \hat{x}^2 \quad (1.35)$$

under the replacements

$$4E_C \rightarrow \frac{\hbar^2}{2m\lambda^2}, \quad (\hat{n} - n_g) \rightarrow i\lambda \partial_x, \quad E_J \rightarrow k\lambda^2, \quad \hat{\varphi} \rightarrow \hat{x}/\lambda \quad (1.36)$$

with the anharmonicity given by the higher-order terms in the expansion of  $\cos(\hat{\varphi})$ , and  $\lambda$  a unit of length (so that the dimensions match). The resulting energy gap  $E_1 - E_0$  between the ground  $|0\rangle$  and first excited state  $|1\rangle$  is then  $\sqrt{8E_C E_J}$  [88]. The anharmonicity reduces the energy of the second state  $E_2 - E_1 < E_1 - E_0$ , allowing for excitations between the first two energy levels without exciting higher levels  $|n\rangle$ .

Coupling a transmon qubit to a resonator allows for coherent control and readout via an applied photon field. The system of a photon mode of frequency  $\omega$  coupled to a transmon qubit may be described via a generalized Jaynes-Cummings model [88, 89]

$$\hat{H}_{\text{JC}} = \sum_n E_n |n\rangle \langle n| + \hbar\omega \hat{a}^\dagger \hat{a} + \frac{1}{2} \sum_n (g_n |n\rangle \langle n+1| \hat{a}^\dagger + g_n^* |n+1\rangle \langle n| \hat{a}), \quad (1.37)$$

where  $\hat{a}$  and  $\hat{a}^\dagger$  are the photon annihilation and creation operators, and the coupling constants  $g_n$  are determined by the interaction of the photon with the charge degree of freedom  $\hat{n}$ . When  $E_1 - E_0 \approx \omega$ , the system is in resonance, and if prepared in the state  $|0\rangle$  it will oscillate between absorbing and emitting a single photon from the readout line. As the photon number is in general not conserved, this oscillation does not alter the wavefunction of the line, and the operation can be restricted to a unitary operation on the transmon subspace. This allows for arbitrary rotations on the single qubit defined on the  $|0\rangle$  and  $|1\rangle$  states. When  $|E_1 - E_0 - \omega| \gg 0$ , the system is off-resonant. This causes the resonator to be dispersively coupled to the transmon; the  $|1\rangle$  and  $|0\rangle$  states cause a slight difference in energy of resonator photons. This energy difference

can be measured by the difference in phase accumulated by photons in the resonator (compared to a coherent reference signal). This in turn projects the qubit into a corresponding eigenstate, allowing for readout. Finally, a resonator coupling two qubits induces a second-order coupling between the qubits. Tuning two energy levels of the qubits into resonance then causes this (otherwise small) coupling to dominate, which can be used to entangle the pair. In particular, this is typically used to generate either a controlled-phase rotation [199], which, combined with arbitrary single-qubit rotations forms a universal gate-set. The above is then sufficient to satisfy the DiVincenzo criterion, so long as sufficiently low error rates can be obtained.

### 1.5.2. Sources of errors

The dominant source of error in superconducting transmon systems is decay due to qubit decay, described by a  $T_1$  channel in the Kraus operator formalism [23]

$$\hat{K}_1^{(T_1)} = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{e^{-t/T_1}} \end{pmatrix}, \quad \hat{K}_2^{(T_1)} = \begin{pmatrix} 0 & \sqrt{1 - e^{-t/T_1}} \\ 0 & 0 \end{pmatrix}, \quad (1.38)$$

and qubit dephasing, described by a  $T_2$  channel

$$\hat{K}_1^{(T_2)} = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{e^{-t/T_2}} \end{pmatrix}, \quad \hat{K}_2^{(T_2)} = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{1 - e^{-t/T_2}} \end{pmatrix}. \quad (1.39)$$

The former is driven by coupling to stray modes and impurities on the chip, decay into surrounding resonators, and external sources of radiation. The latter is mostly driven by fluctuations in the magnetic field  $\Phi$  through the SQUID loop, known as ‘flux noise’, which in turn change the Josephson energy  $E_J$ . These are mitigated somewhat by holding resting qubits at the spot  $\Phi = 0$ , such that  $\partial E_J / \partial \Phi = 0$ , making the qubit only second-order sensitive to flux noise [88]. However, during two qubit gates qubits are driven away from this ‘sweet-spot’ and become sensitive again. For example, a C-Phase gate is performed by tuning two qubits such that the  $|11\rangle$  and  $|02\rangle$  qubits are in resonance, which shifts the  $|11\rangle$  state via level repulsion, causing a phase to accumulate. Flux noise will cause the accumulated phase to differ from its target. Worse, such a gate can cause the  $|11\rangle$  state to be driven out of the computational subspace. To see this, note that a C-Phase gate can be considered a fully-reflective Landau-Zener process, which requires fine-tuning, and any fluctuations

## 1. Introduction

will generally break this fine-tuning. Such *leakage* errors are generally considered to be worse for quantum computers, and in particular for quantum error correction. However, schemes do exist to mitigate leakage such that fault-tolerant thresholds remain [90, 91].

A further range of errors also exist on transmon qubits that remain significant, but are generally smaller in magnitude than the above (though this is highly dependent on the exact chip used). Most single qubit gates have small residual control errors, though state-of-the-art single qubit gates are mostly limited by qubit decay during the duration of the gate [183]. Classical readout error in measurement is also present, though this is again usually smaller than the effect of decay. A more serious issue with readout is leftover photons in the readout resonator, which must be depleted via the Purcell effect [167] or depletion pulses [182] in order to operate on a transmon post-measurement. Various types of crosstalk between qubits also exists; measurements or operations directed at one qubit may affect another, and the residual coupling between any connected pair of qubits causes a slow coherent rotation over the course of an experiment. And finally, qubits can only be re-prepared in the ground state via either natural decay processes (i.e. waiting for times much longer than  $T_1$ ), or conditional feedback schemes [185], both of which leave some probability for a qubit to be erroneously excited. This residual excitation can be minimized by measuring qubits prior to an experiment and pre-selecting on a measurement of 0.

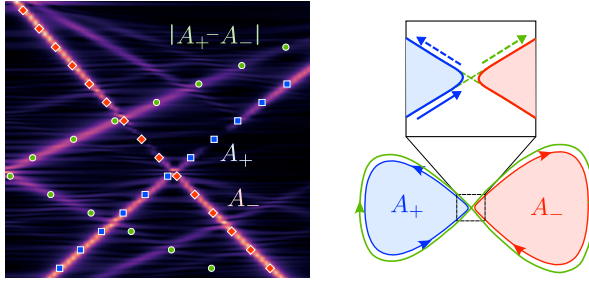
## 1.6. Outline of this thesis

### 1.6.1. Part I

In the first part of this thesis, we investigate transport properties of exotic and superconducting Weyl semimetals.

#### Chapter 2: Magnetic breakdown and Klein tunneling in a type-II Weyl semimetal

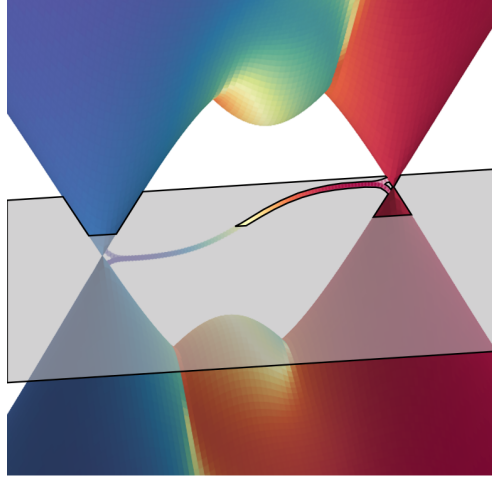
In Chapter 2, we predict a topologically protected signature of type-II Weyl cones in magnetic oscillation experiments. This signature arises from the existence of maximal figure-of-eight orbits around the hourglass-like Fermi surface of a type-II Weyl semimetal, which exist under surprisingly general conditions due to the non-orientability of the projective plane. In order to achieve these orbits, electrons must jump between electron and hole pockets, in a momentum-space analogue of Klein tunnelling. We numerically simulate these experiments on a toy lattice model of a type-II Weyl semimetal, and obtain close agreement between experiment and theory (shown in Fig. 1.7). We then finish with an investigation into the effect of additional details in the Fermi surface on the predicted effect, in particular the effect of multiple Weyl points.



**Figure 1.7.:** Simulated demonstration of Klein tunnelling in the momentum space of a type-II Weyl semimetal. Peaks in the displayed Fourier transform (right) correspond to maximal orbits in the Fermi surface of the crystal (left). The green circles mark a peak that can only be explained by electron tunnelling between the electron (blue square) and hole (red diamond) pockets in the Fermi surface, which requires tunnelling between the two, as shown. Figure adapted from Chapter 2; more details available there.

### Chapter 3: Superconductivity provides access to the chiral magnetic effect of an unpaired Weyl cone

In Chapter 3, we predict that the equilibrium chiral magnetic effect, which cannot exist in the presence of a  $U(1)$  symmetry [137–139], may re-emerge as a charge current in a superconductor when this symmetry is broken. This comes about due to the Bogoliubov-de Gennes Hamiltonian allowing for a single cone to appear twice in the Brillouin zone (as the electron and hole counterparts), as shown in Fig. 1.8. The resulting chiral current is renormalized by the charge on the cone, which is increased as its particle and hole halves are separated in the Brillouin zone. We again compare numerical simulations on a toy lattice model with analytic predictions, and find good agreement.



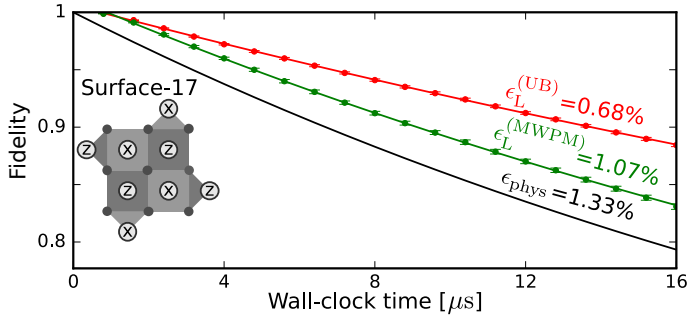
**Figure 1.8.:** Bandstructure of a Weyl superconductor in the  $k_z - k_y$  plane under an applied magnetic field and supercurrent, coloured by the charge on the Weyl cones. In the BdG formalism, only bands above the Fermi surface (gray plane) correspond to real excitations; bands below are lost in the BCS condensate. This system can be observed then to have only a single Weyl cone, split into conjugate particle and hole pairs. An applied magnetic field will then drive a charge current in equilibrium, allowing for observation of the chiral-magnetic effect. The particle current remains zero however, as the surface band replaces charged fermions with neutral Majorana particles flowing in the opposite direction. Figure adapted from the system studied in Chapter 3, more details available there.

### 1.6.2. Part II

In the second part of this thesis, we switch to investigating the performance of small quantum error correcting codes with realistic noise and to optimizing decoders for said experiments.

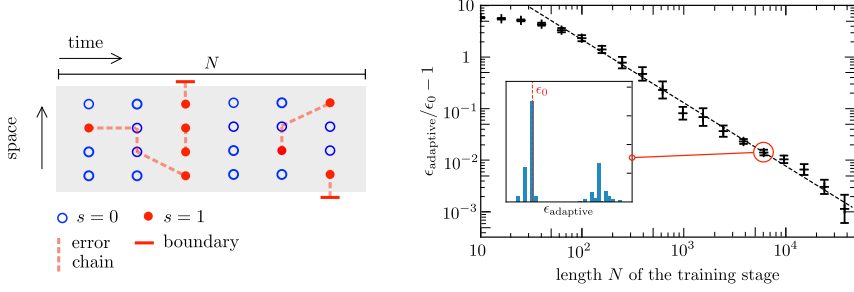
#### Chapter 4: Density-matrix simulation of small surface codes under current and projected experimental noise

In Chapter 4, we present a density-matrix simulation of superconducting qubits, and use this to determine the performance of a small error correcting code (the distance-3 code Surface-17). This simulation demonstrates that state-of-the-art superconducting quantum devices are sufficient to construct a logical qubit which outperforms its component qubits as a quantum memory, as shown in Fig. 1.9. We are able to use our simulation to optimize various design choices in the resulting code, and test the performance of the code as various noise parameters are increased or decreased.



**Figure 1.9.:** Simulation of performance of Surface-17 (drawn in inset) as a quantum memory, under realistic noise simulations. Over multiple 800 ns cycles (points), the system accumulates both uncorrectable and correctable errors. A maximum likelihood decoder (red points) successfully corrects all correctable errors, resulting in an upper bound of  $\epsilon_L^{(UB)} = 0.68\%$  for the logical error rate per cycle. However, this decoder is too computationally expensive for real-world use. The amount by which a faster decoder falls below this bound gives the decoder efficiency [212];  $\eta_D = \epsilon_L^{(UB)} / \epsilon_L^{(D)}$ . Despite being slightly inefficient, we observe that the MWPM decoder is sufficient to make the lifetime of the Surface-17 logical qubit longer than that of its physical components (black line). Figure taken from Chapter 4, more details available there.

## 1. Introduction



**Figure 1.10.:** Schematic and convergence of the adaptive weight estimation technique. (left) The estimator is fed stabilizer data (red filled and blue open circles) from  $N$  QEC cycles, which are generated by unseen chains of errors (red dashed lines). The estimator infers the rate at which each chain is generated, which gives sufficient information for minimum-weight perfect matching. (right) The resultant decoder converges to an optimal minimum-weight perfect matching decoder as  $N$  increases, and by  $N \approx 10^4$  rounds is effectively completely trained. Further details (including simulation details) can be found in Chapter 5.

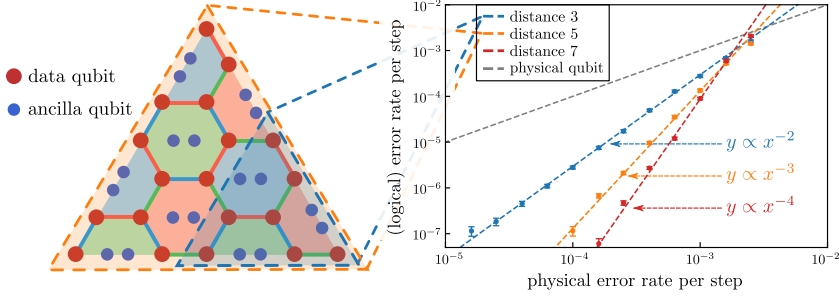
## Chapter 5: Adaptive weight estimator for quantum error correction in a time-dependent environment

In Chapter 5, we improve the performance of minimum-weight perfect matching decoders for surface and repetition codes by inferring the error rates of such codes from repeated syndrome measurement. We are able to derive an exact analytic formula for such inference, allowing for low-cost weight extraction for minimum-weight perfect matching. The resulting decoder converges to near-optimal decoder efficiency within only  $N = 10,000$  quantum error correcting cycles, as shown in Fig. 1.10, making this a very attractive technique for future use. We further investigate the performance of this decoder in the presence of fluctuating noise, and show that the convergence rate gives an order of magnitude for the ability of a decoder to adapt to said fluctuations.

## Chapter 6: Neural network decoder for topological color codes with circuit level noise

In Chapter 6, we investigate the performance of machine learning decoders for color codes, for which no high-efficiency algorithm was previously known. This follows a previous work on the Surface code [211], but is complicated by the need to account for hook errors in the color code, which





**Figure 1.11.:** Schematic and performance of the color code. (Left) The distance-5  $[6, 6, 6]$  color code. Pairs of (blue) ancilla qubits are used to make repeated flagged parity checks on neighbouring (red) data qubits. The equivalent distance-3 code may be obtained by choosing a smaller plaquette of the lattice (blue shaded region). (Right) Repeated stabilizer measurements and quantum error correction allow for the correction of all combinations of errors that do not span at least half the lattice, leading to a logical error rate exponentially small in the resulting code distance  $d$ . That this exponential decay scales as  $(d + 1)/2$  implies that the single-flag fault-tolerance scheme protects against hook errors.

we do via the measurement of a flagging circuit along the lines of [243]. We make critical improvements to both the decoder stability and flag circuit size; in particular observing that a distance- $d$  code may be corrected with a single flag qubit per stabilizer rather than  $(d - 1)/2$ . This is demonstrated by the observed  $p^{(d+1)/2}$  scaling in Fig. 1.11. We observe that with our new decoder, the resulting code sits at the threshold of outperforming its physical components as a memory, making it potentially attractive to near-term quantum experiments.

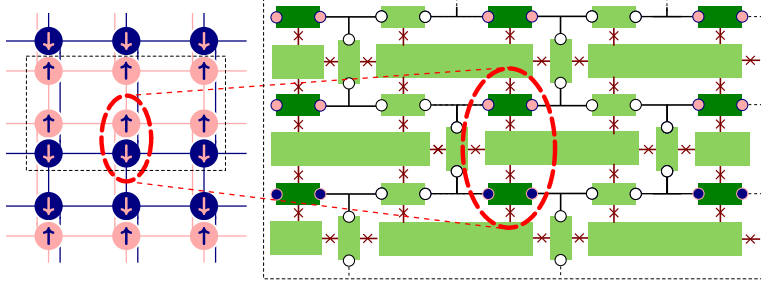
### 1.6.3. Part III

In the final part of this thesis, we switch to developing new quantum algorithms, primarily for the subject of digital quantum simulation, and with a focus on near-term devices.

## Chapter 7: Majorana-based fermionic quantum computation

In Chapter 7, we study the idea of using Majorana zero-modes to build fermions for quantum computing, instead of qubits. We deconstruct various digital quantum simulation circuits in terms of the evolution of Majorana

## 1. Introduction

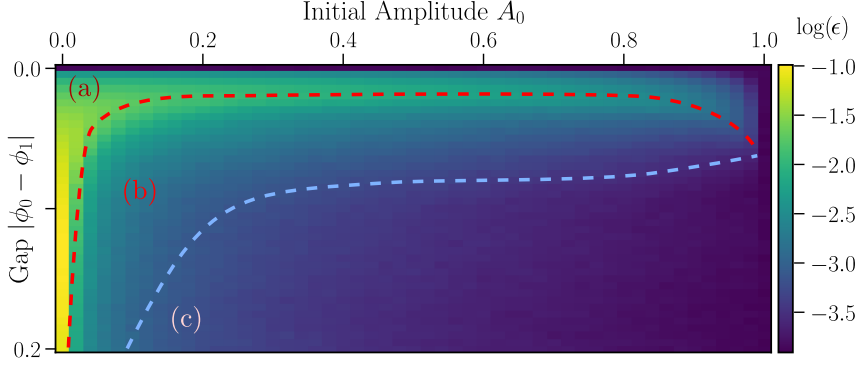


**Figure 1.12.:** Scheme to simulate the Hubbard model on a 2D square lattice using Majorana-based fermions for quantum computing. (Left) the 2-d lattice of the Hubbard model; each site contains two interacting spins, and electrons interact via a spin-conserving hopping term between neighbouring sites. (Right) this may be encoded on a tessellatable lattice of superconducting islands (green) connected by Josephson junctions (red lines) and nanowires (black lines); Majorana zero modes may be isolated on sites where a nanowire crosses a superconducting island (circles). Two Majoranas encode a fermion, however we pattern the square lattice sparsely (1/3 filling) on the computer architecture, using only the shaded sites. This allows free space for Majoranas to be shuttled around for computation, allowing for simulated time evolution in constant time. The Majorana hardware shown encodes only those spins circled in the black dashed lines; single pairs of spins are associated as per the red dashed lines.

zero-modes, and find a class of low-cost operations that can be used to achieve all such operations. We demonstrate that this would result in lower overhead for solving the electronic structure problem, and present schemes (see Fig. 1.12) for simulation of the Hubbard model in a constant time step (as a function of the system size). We further investigate magic state distillation in Majorana-based circuits, as will be needed for fault-tolerance.

## Chapter 8: Quantum phase estimation for noisy, small-scale experiments

Finally, as an outlook, in Chapter 8 we study the ability of single-ancilla variants of quantum phase estimation in realistic scenarios. Said variants are attractive due to lower overhead requirements, but have till now not been investigated in the situation where the phase estimation algorithm is performed on starting states that are superpositions of Hamiltonian eigenstates. We demonstrate that both Prony's method and Bayesian



**Figure 1.13.:** Error in quantum phase estimation of a two-eigenvalue system, given a short, noisy signal and using an estimator based on Prony’s method. This toy model may be parametrized in terms of the gap between the eigenvalues and the amplitude of the target eigenvalue  $\phi_0$ . The resulting diagram may be split into three regions. In region (a), the estimator cannot distinguish the two phases, and estimates a single phase in between; the error then comes from their separation. In region (b), the phases are separated enough to identify individually, but the estimation of each is marred by the other’s presence. In region (c), the phases are completely separated and do not affect the estimation of each other. As the signal is lengthened, or made less noisy, regions (a) and (b) retract, and the estimation precision improves.

techniques are able to resolve single eigenphases from these superpositions. We are able to derive or numerically observe scaling laws for both methods (see Fig. 1.13) in terms of various experimental and Hamiltonian parameters. We further investigate the effect of noise on these methods, and find methods for compensating simple noise methods by incorporating them into the estimation itself.



## Part I.

# Magnetotransport in topological semimetals



## 2. Magnetic breakdown and Klein tunneling in a type-II Weyl semimetal

### 2.1. Introduction

Weyl semimetals provide a condensed matter realization of massless relativistic fermions [50]. Their spectrum features a diaboloid-shaped surface in energy-momentum space that separates helical electron-like states (moving in the direction of the momentum) from hole-like states (moving opposite to the momentum) [92]. These “Weyl cones” are the three-dimensional analogue of the two-dimensional Dirac cones in graphene. The third spatial dimension provides a topological protection, by which the conical point (Weyl point) cannot be opened up unless two Weyl cones of opposite helicity are brought together in momentum space [93].

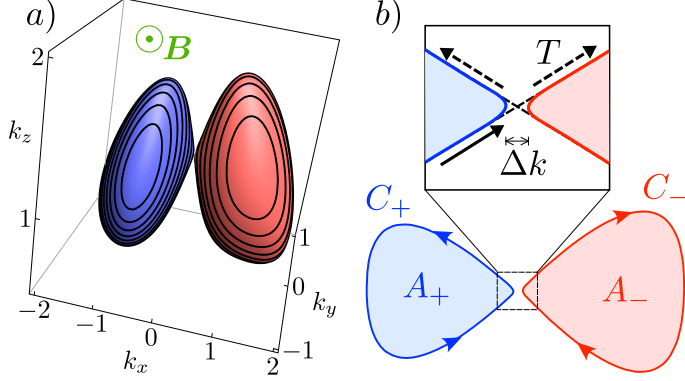
Although the Weyl point cannot be locally removed, the cones can be tilted and may even tip over [51–59]. For the relativistic Weyl cone such a distortion is forbidden by particle-hole symmetry, but that is not a fundamental symmetry in condensed matter. While in graphene the high symmetry of the honeycomb lattice keeps the cone upright, strain providing only a weak tilt [94], the tilting can be strong in 3D Weyl semimetals. This leads to a natural division of Weyl cones into two topologically distinct types [56]. In type I the cone is only weakly tilted so that the electron-like states and hole-like states occupy separate energy ranges, above or below the Weyl point. In type II the cone has tipped over so that electron and hole states coexist in energy. Many experimental realizations of a type-II Weyl semimetal have recently been reported [95–101].

In a magnetic field the coexisting electron and hole pockets of a type-II Weyl semimetal are coupled by tunneling through the Weyl point (Fig. 2.1).

---

The contents of this chapter have been published in T. E. O’Brien, M. Diez and C. W. J. Beenakker, *Phys. Rev. Lett.* **116** (23), 236401 (2016).

## 2. Magnetic breakdown and Klein tunneling in a type-II Weyl semimetal



**Figure 2.1.:** a) Fermi surface of a type-II Weyl semimetal, calculated from the model Hamiltonian (2.1), showing the electron and hole pockets touching at the Weyl point. Equi-energy contours in planes perpendicular to the magnetic field  $B$  are indicated. The magnetic quantum oscillations have a periodicity in  $1/B$  determined by the contour that encloses an extremal area. b) Intersection of the Fermi surface with a plane perpendicular to  $B$  that passes through the Weyl point. Electron and hole pockets are bounded by a contour  $C_{\pm}$  enclosing an area  $A_{\pm}$ . The semiclassical orbit of an electron follows the contour in the direction of the arrow. Tunneling between the pockets happens with a probability  $T$  that tends to unity when their minimal separation  $\Delta k \rightarrow 0$ . This magnetic breakdown is a manifestation of Klein tunneling in momentum space.

Here we investigate how this process, a momentum space manifestation of Klein tunneling [102], affects the magnetic quantum oscillations of the density of states (De Haas-Van Alphen effect), providing a unique thermodynamic signature of the topologically protected band structure (an alternative to proposed transport signatures [56, 103–105]). Because the quantum oscillations are governed by extremal cross-sections of the Fermi surface, one might wonder whether some high symmetry is required to align the extremal cross-section with the Weyl point, so that it becomes observable. Our analysis shows that a magnetic field axis for this alignment exists generically, because of the Möbius strip topology of the projective plane. We first consider Klein tunneling through a single type-II Weyl point, and then turn to pairs of Weyl cones of opposite helicity, which can be combined in topologically distinct ways [55] — with a qualitatively different dependence on the Klein tunneling probability.



## 2.2. Semiclassical quantization

To first order in momentum  $\mathbf{k}$ , the Hamiltonian of a Weyl cone has the generic form

$$H = \sum_{ij} v_{ij} k_i \sigma_j + a_{\text{tilt}} k_x \sigma_0, \quad (2.1)$$

in terms of Pauli matrices  $\sigma_i$ ,  $i \in \{x, y, z\}$  (unit matrix  $\sigma_0$ ). The eigenvalues lie on two hyperboloid sheets  $E_{\pm}$ ,

$$E_{\pm} = a_{\text{tilt}} k_x \pm \sqrt{\sum_{ijl} v_{il} v_{jl} k_i k_j}, \quad (2.2)$$

that touch at the Weyl point  $\mathbf{k} = 0$ .

For sufficiently small  $a_{\text{tilt}}$  the Fermi surface contains either electron-like states in  $E_+$  or hole-like states in  $E_-$ , depending on the sign of the Fermi energy. With increasing  $a_{\text{tilt}}$  the Weyl cone is tilted in the (arbitrarily chosen)  $x$ -direction, and when it tips over coexisting electron and hole states appear on the Fermi surface. This is the type-I to type-II Weyl semimetal transition [56].

The hyperboloid dispersion (2.2) only holds near the Weyl point. In the physical realizations of a type-II Weyl semimetal the Fermi surface closes away from the Weyl point, forming compact electron and hole pockets. A cross-section is defined by fixing an axis (unit vector  $\hat{\mathbf{n}}$ ) and choosing a coordinate  $q$  along that axis. The intersection of the Fermi surface with the plane  $\hat{\mathbf{n}} \cdot \mathbf{k} = q$  is an oriented contour  $C_{\pm}(q)$  enclosing the signed area  $A_{\pm}(q)$  (positive for  $C_+$  and negative for  $C_-$ ). The contours are the classical momentum-space orbits for a magnetic field  $B$  in the  $\hat{\mathbf{n}}$ -direction, the change in orientation between  $C_+$  and  $C_-$  resulting from the sign change of the effective mass in the electron and hole pockets.

Semiclassical quantization of the orbits produces Landau tubes [25], with quantized cross-sectional area

$$A_{\pm}(q) = 2\pi(n + \nu)eB/\hbar, \quad n = \pm 1, \pm 2. \quad (2.3)$$

The Maslov index  $\nu = 1/2$  for massive electrons, while  $\nu = 0$  for massless Weyl fermions [106]. The Landau tubes give rise to oscillations in the density of states periodic in  $1/B^*$ ,

$$\delta\rho/\rho_0 = \text{Re} \left\{ [-iA''_{\pm}(q_c)]^{-1/2} e^{2\pi i(F_{\pm}/B - \nu)} \right\}, \quad (2.4)$$

---

\*For a convex electron or hole pocket one has  $\text{sign } A''_{\pm}(q_c) = \mp 1$  so the factor  $[-iA''_{\pm}(q_c)]^{-1/2}$  in Eq. (2.4) contributes a phase shift  $\mp\pi/4$  to the magnetic oscillations.

## 2. Magnetic breakdown and Klein tunneling in a type-II Weyl semimetal

with frequency given by the Onsager relation [107, 108]

$$F_{\pm} = (\hbar/2\pi e)|\mathcal{A}_{\pm}|. \quad (2.5)$$

The extremal area  $\mathcal{A}_{\pm} = A_{\pm}(q_c)$  is the area at which the first derivative  $dA_{\pm}(q)/dq = 0$ . The contour enclosing the extremal area is denoted by  $\mathcal{C}_{\pm}$ .

### 2.3. Magnetic breakdown

The two sheets  $E_{\pm}$  of a type-II Weyl cone are coupled by quantum tunneling. This magnetic-field-induced tunneling between electron and hole pockets is the momentum space counterpart of Klein tunneling at a  $p$ - $n$  junction in graphene [109], and can be analyzed along the same lines [110].

The effect of a magnetic field  $B$  in, say, the  $y$ -direction, with vector potential  $\mathbf{A} = (Bz, 0, 0)$ , is accounted for by the substitution  $k_x \mapsto k_x + eBz$  (setting  $\hbar = 1$ ). In momentum representation, the Schrödinger equation  $H\psi = E\psi$  reads

$$iU_0 \frac{\partial \psi}{\partial k_z} = U(k_z)\psi, \quad U_0 = eB(\sum_j v_{xj}\sigma_j + a_{\text{tilt}}\sigma_0), \quad (2.6a)$$

$$U(k_z) = E\sigma_0 - \sum_{ij} v_{ij}k_i\sigma_j - a_{\text{tilt}}k_x\sigma_0. \quad (2.6b)$$

For  $a_{\text{tilt}} > (\sum_j v_{xj}^2)^{1/2}$  the matrix  $U_0$  is positive definite, so that it can be factorized as  $U_0 = VV^{\dagger}$  with invertible  $V$  and we may write

$$i\partial\psi/\partial k_z = V^{-1}U(k_z)(V^{\dagger})^{-1}\psi \equiv \mathcal{H}(k_z)\psi, \quad (2.7)$$

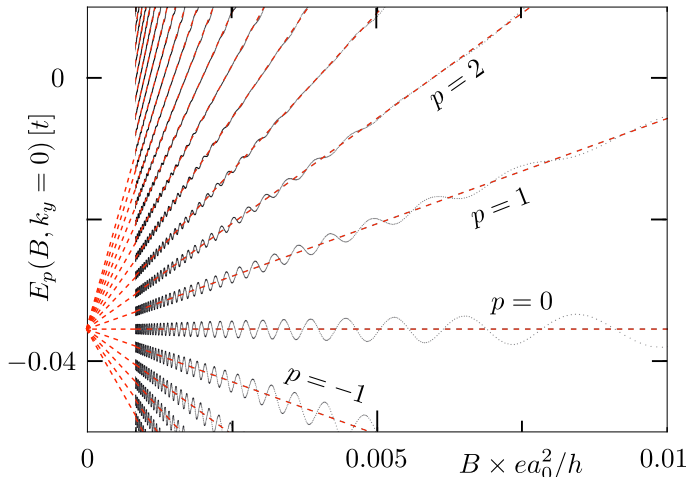
with  $\mathcal{H}(k_z) = \mathcal{H}_0 + \mathcal{H}_1 k_z$ . If we interpret  $k_z \equiv t$  as “time”, this looks like a Schrödinger equation for a spin-1/2 particle with “time”-dependent Hamiltonian  $\mathcal{H}(t)$ . Because the  $t$ -dependence of  $\mathcal{H}(t)$  is linear, we can use the Landau-Zener formula for the tunneling probability between the electron and hole pockets [111].

Quite generally, for a two-level system with time-dependent Hamiltonian

$$\mathcal{H}(t) = \begin{pmatrix} \alpha t + c & \gamma \\ \gamma^* & \beta t + c' \end{pmatrix}, \quad (2.8)$$

the Landau-Zener tunnel probability is

$$T = \exp(-2\pi|\gamma|^2|\alpha - \beta|^{-1}). \quad (2.9)$$



**Figure 2.2.:** Energy spectrum at  $k_y = 0$  of the type-II Weyl semimetal with Hamiltonian (2.11) (parameters  $t = 1$ ,  $t' = 2$ ,  $\mu = 3$ ,  $b = 1.2$ ,  $a_{\text{tilt}} = 1.7$ ,  $\xi = 0.08$ ). The black dotted curves are the exact numerical results, the red dashed lines form the semiclassical Landau fan (2.15) for tunnel-coupled electron and hole pockets. The individual pockets are responsible for the high-frequency oscillations superimposed on the fan.

The matrix (2.7) is of the form (2.8) in the basis where  $\mathcal{H}_1$  is diagonal, so in that basis we can read off the coefficients  $\alpha, \beta, \gamma$  needed to determine  $T$ .

For a specific example we consider the Hamiltonian (2.1) with  $v_{ij} = v_i \delta_{ij}$ , which for  $a_{\text{tilt}} > v_x$  represents a type-II Weyl cone. We find

$$\begin{aligned} T &= \exp \left( -\frac{\pi \hbar}{eB} \frac{v_x^2 E^2 + v_y^2 k_y^2 (a_{\text{tilt}}^2 - v_x^2)}{v_z (a_{\text{tilt}}^2 - v_x^2)^{3/2}} \right) \\ &= \exp \left( -\frac{\pi \hbar}{4eBv_z} (\Delta k)^2 (a_{\text{tilt}}^2 - v_x^2)^{1/2} \right), \end{aligned} \quad (2.10)$$

with  $\Delta k$  the minimal separation of the contours  $C_+$  and  $C_-$ . This has the general form of the interband tunnel probability in the theory of magnetic breakdown [108, 112, 113], with a breakdown field  $B_c \propto (\Delta k)^2$ . The characteristic feature of Klein tunneling is that the tunnel probability  $T \rightarrow 1$  and  $B_c \rightarrow 0$  at the conical point of the band structure — here a 3D Weyl point and a 2D Dirac point in Ref. 102.

To illustrate the effect of Klein tunneling between electron and hole pockets on the magnetic quantum oscillations in the density of states, we

## 2. Magnetic breakdown and Klein tunneling in a type-II Weyl semimetal

consider the model Hamiltonian [65]<sup>\*</sup>

$$\begin{aligned}
 H = & \tau_z(t' \sigma_x \sin k_x + t' \sigma_y \sin k_y) + t \tau_z \sigma_0 \sin k_z \\
 & + \tau_x \sigma_0 (\mu - t \cos k_x - t \cos k_y - t \cos k_z) \\
 & + b \tau_0 \sigma_z + [a_{\text{tilt}} \sin k_x + \xi(1 - \cos k_x)] \tau_0 \sigma_0.
 \end{aligned} \tag{2.11}$$

This is a tight-binding Hamiltonian on a cubic lattice (lattice constant  $a_0 = 1$ ), with a spin and orbital degree of freedom on each lattice site (Pauli matrices  $\sigma_i$  and  $\tau_i$ , respectively). The time-reversal symmetry breaking term  $b$  splits the Dirac cone into two Weyl cones separated along the  $z$ -axis. To produce a type-II Weyl semimetal we have added a tilting term  $a_{\text{tilt}}$  and a term  $\xi$  that breaks the symmetry between the electron and hole pockets.

As derived in App. A, near a Weyl point the effective low-energy Hamiltonian has the form (2.1) with diagonal velocity tensor  $v_{ij} = v_i \delta_{ij}$  given by

$$v_x = v_y = \frac{(2t - \mu)^2 - t^2 + b^2}{2b(2t - \mu)} t', \tag{2.12a}$$

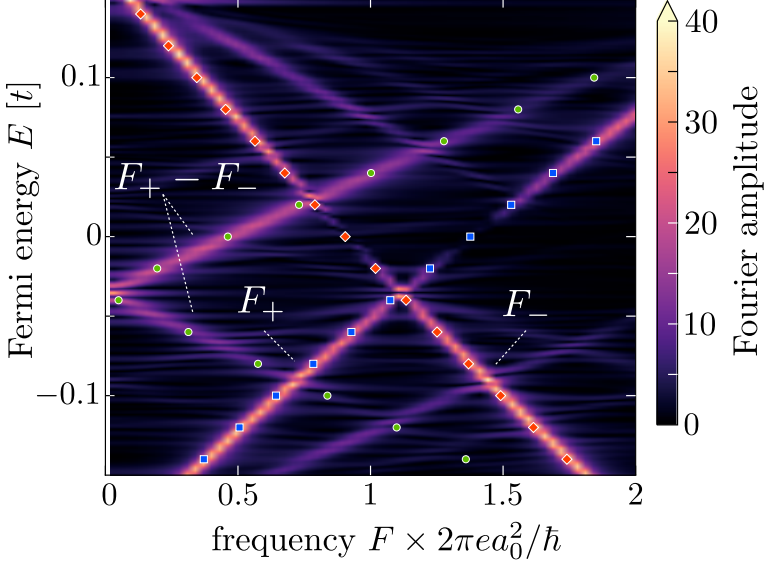
$$v_z = \frac{1}{2b} \sqrt{[(t - \mu)^2 - b^2][b^2 - (3t - \mu)^2]}. \tag{2.12b}$$

The Hamiltonian (2.11) retains a mirror symmetry in the  $x$ - $z$  plane (to be removed later on), which implies that for a magnetic field in the  $y$ -direction the areas  $A_{\pm}(k_y)$  are extremal for  $k_y = 0$ . By means of exact diagonalization<sup>†</sup> we have calculated the partial density of states  $\rho(E, B, k_y) = \sum_p \delta[E - E_p(B, k_y)]$  for  $k_y = 0$ , assuming that this gives the dominant contribution to the magnetic quantum oscillations. We choose the gauge  $\mathbf{A} = (0, 0, -Bx)$ , with a rational flux  $Ba_0^2 = 1/N \times h/e$  through a unit cell. The lattice has dimensions  $N \times NM$  in the  $x$ - $z$  plane ( $M \gg N \gg 1$ ), with periodic boundary conditions in both directions.

Fig. 2.2 shows the energy spectrum as a function of magnetic field and Fig. 2.3 shows the periodicity of the magnetic oscillations, extracted from a Fourier transform of the density of states. When the Fermi level is far from the Weyl point  $E = 0$ , the electron and hole pockets contribute separately with frequencies  $F_{\pm}$  set by the extremal areas  $\mathcal{A}_{\pm}$ . The slopes

<sup>\*</sup>The Hamiltonian (2.11) differs from that in this reference by the addition of the  $a_{\text{tilt}}$  and  $\xi$  terms, and also in the replacement of  $\tau_y \sigma_0 \sin k_z$  by  $\tau_z \sigma_0 \sin k_z$ . This last change was introduced to produce small electron and hole pockets that do not spread out over the entire Brillouin zone.

<sup>†</sup>To discretize the model Hamiltonian (2.11) we used the KWANT toolbox of [114].



**Figure 2.3.:** Fourier amplitudes of the magnetic quantum oscillations. The numerical data for the partial density of states  $\rho(E, k_y = 0)$  (smoothed with a Gaussian of width  $\Gamma = t/500$ ) is Fourier transformed over the field range  $B \lesssim 0.005 \hbar/ea_0^2$  ( $200 < N < 1500$ ). The fundamental frequencies from the electron and hole pockets are indicated by  $F_+$  and  $F_-$ , respectively (the first harmonics are also faintly visible). Klein tunneling between the pockets when the Fermi energy approaches the Weyl point ( $E = 0$ ) suppresses these high-frequency oscillations, introducing a new component at the difference frequency  $|F_+ - F_-|$ . The colored data points for  $F_\pm$  are the semiclassical prediction (2.5) from the extremal areas.

$dF_\pm/dE$  have opposite sign in the two pockets, signifying the opposite sign of the cyclotron effective mass

$$m_\pm = \frac{\hbar^2}{2\pi} \frac{d}{dE} |\mathcal{A}_\pm|. \quad (2.13)$$

Near the Weyl point a low-frequency component appears at the difference  $|F_+ - F_-|$ , and the individual high-frequency components  $F_\pm$  are suppressed. In a semiclassical description, the orbit responsible for the difference frequency is the “figure of eight” orbit formed by joining  $C_+$  to

## 2. Magnetic breakdown and Klein tunneling in a type-II Weyl semimetal

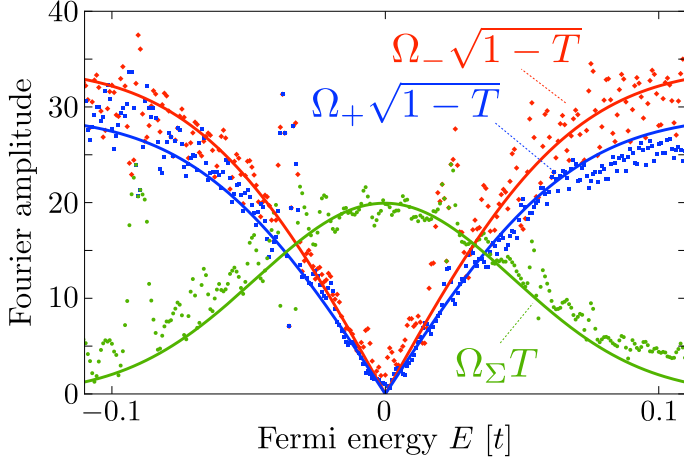
$C_-$  at the Weyl point (see Fig. 2.1b). The corresponding effective mass

$$m_\Sigma = \frac{\hbar^2}{2\pi} \frac{d}{dE} |\mathcal{A}_+ + \mathcal{A}_-| \quad (2.14)$$

governs the Landau fan in Fig. 2.2,

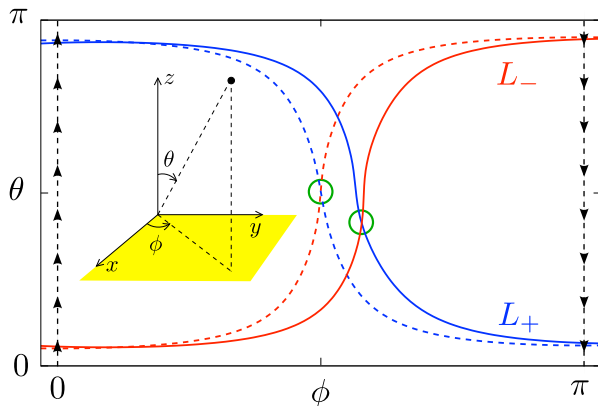
$$E_p(B) = E_p(0) + p \times \hbar e B / m_\Sigma. \quad (2.15)$$

Notice the absence of a  $1/2$  offset from the integer  $p$ , canceled by a Berry phase.



**Figure 2.4.:** Energy dependence of the Fourier amplitudes from Fig. 2.3. The curves are fits to  $\Omega_\pm \sqrt{1-T}$  and  $\Omega_\Sigma T$ , with the transmission probability  $T(E)$  calculated from Eq. (2.10) and energy-independent fit parameters  $\Omega_\pm, \Omega_\Sigma$ . When two frequency lines in Fig. 2.3 cross we cannot reliably determine the individual amplitudes — which explains some of the large scatter in the data points.

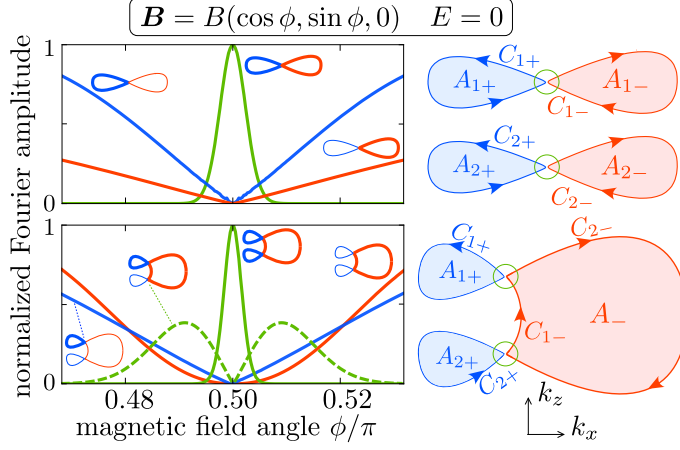
The tunnel probability (2.10) evaluates for our model parameters to  $T(E) = \exp[-0.52 N(E/t)^2]$ . The contribution of an orbit to the Fourier amplitude contains a factor  $t = \sqrt{T}$  for each transmission through the Weyl point and a factor  $r = \sqrt{1-T}$  for each reflection. In Fig. 2.4 we plot the peak heights of Fig. 2.3 as a function of energy. The solid lines are fits to  $\Omega_\pm \sqrt{1-T(E)}$  and  $\Omega_\Sigma T(E)$ , with energy-independent fit parameters  $\Omega_\pm, \Omega_\Sigma$ . We take for the inverse field strength  $N = 850$ , half-way the interval used in the Fourier transform. A good match to the predicted Gaussian  $T(E)$  is obtained.



**Figure 2.5.:** Magnetic field axis  $L_{\pm} = (\theta_{\pm}, \phi_{\pm})$  for which the extremal contour  $C_{\pm}$  at  $E = 0$  touches the Weyl point. The dashed curves correspond to the Hamiltonian (2.11) with the parameters of Fig. 2.1. For the solid curves we have broken the mirror symmetry by adding the term  $V_0 \tau_0 \sigma_0 \sin k_y$  with  $V_0 = 0.5$ . The intersection of  $L_+$  and  $L_-$  (encircled) is the special axis at which Klein tunneling between electron and hole pockets produces magnetic quantum oscillations with the difference frequency  $|F_+ - F_-|$ , suppressing both the electron and hole frequencies  $F_{\pm}$ . The intersection is protected by the topology of the Möbius strip (indicated by arrows, which show how the edges at  $\phi = 0, \pi$  should be glued with a twist).

The above analysis was simplified by the mirror symmetry in the  $x$ - $z$  plane, because we could immediately identify the special magnetic field axis for which the extremal contours  $C_{\pm}$  in the electron and hole pockets both touch the Weyl point when  $E \rightarrow 0$ , allowing for Klein tunneling. One might wonder how restrictive this alignment is — is it possible to find such a special axis in the absence of any symmetry? The answer is yes, as we demonstrate with the help of Fig. 2.5. At  $E = 0$  we plot the polar and azimuthal angles  $\theta_{\pm}, \phi_{\pm}$  of the magnetic field axis for which the extremal contour  $C_{\pm}$  touches the Weyl point. Because  $(\theta, \phi)$  and  $(\pi - \theta, \pi + \phi)$  represent the same axis, we may restrict  $\phi$  to the range  $[0, \pi]$  — half the usual range for spherical coordinates — identifying the end points  $(\theta, 0)$  and  $(\pi - \theta, \pi)$ . The  $(\theta, \phi)$  plane with these “twisted” periodic boundary conditions is the so-called projective plane  $\mathbb{P}_2$ , and has the topology of a Möbius strip.

If the loops  $L_+ = (\theta_+, \phi_+)$  and  $L_- = (\theta_-, \phi_-)$  both wind around the Möbius strip, as they do in Fig. 2.5, they must necessarily intersect because



**Figure 2.6.:** Left panels: Dependence on the orientation of the magnetic field of the amplitude of the magnetic quantum oscillations (normalized to unit maximal amplitude), for a fixed Fermi energy  $E_F = 0$ . Pairs of type-II Weyl points at  $E = 0$  with disconnected or connected Fermi surfaces are compared. The right panels show a cross-section through the electron and hole pockets. For each curve in the left panels the corresponding orbit is indicated. The calculations, detailed in App. C, are for the Hamiltonian (2.11) with parameters  $t = 1$ ,  $\mu = 3$ ,  $b = 1.2$ ,  $a_{\text{tilt}} = 1.7$  for all panels and  $t' = 2$ ,  $\xi = 0.08$  (top panels);  $t' = 1.7$ ,  $\xi = 0.24$  (bottom panels).

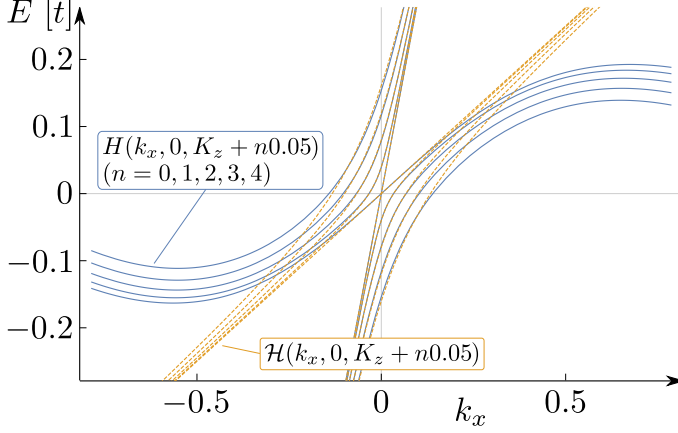
of the twist. The point of intersection is the special axis at which both  $\mathcal{C}_+$  and  $\mathcal{C}_-$  touch the Weyl point. In App. B we show that such non-contractible loops always exist if the Fermi surface is convex, independent of any symmetry requirement.

## 2.4. Conclusion

So far we considered Klein tunneling at a single Weyl point. A second Weyl point of opposite helicity necessarily exists in the Brillouin zone, and this allows for topologically distinct Fermi surfaces [55]. In Fig. 2.6 we illustrate how Klein tunneling can distinguish connected from disconnected pairs of type-II Weyl cones, by the qualitatively different dependence on the magnetic field orientation.

Experimentally, Klein tunneling through a type-II Weyl point can be detected in measurements of the De Haas-Van Alphen effect in the magnetic





**Figure 2.7.:** Dispersion close to a type-II Weyl point. Shown are five  $k_z$ -subbands at  $k_y = 0$ ; blue-solid lines show the dispersion of the four-band model (2.11); yellow-dashed lines show the corresponding low energy description (2.23). Parameters are the same as in Fig. 2.2.

susceptibility: If the magnetic axis is rotated towards the special alignment of Fig. 2.5, the high-frequency magnetic quantum oscillations from the electron and hole pockets would both be suppressed in favor of a low-frequency oscillation from the coupled orbits. The characteristic field for this magnetic breakdown would depend quadratically on the energy mismatch  $E$  between the Weyl point and the Fermi energy, with unit tunnel probability in the limit  $E \rightarrow 0$  as the defining signature of Klein tunneling in momentum space. With sufficient doping WTe<sub>2</sub> would produce disconnected type-II Weyl cones near the Fermi energy [56, 115]\*, while they are connected in undoped LaAlGe [96]. Klein tunneling is a powerful diagnostic for such topologically distinct Fermi surfaces.

---

\*The recent experimental study [115] of magnetic quantum oscillations in semimetallic WTe<sub>2</sub> reports the sudden appearance of a new frequency above a critical field. Since the new frequency is the sum rather than the difference of the low-field frequencies, it cannot be associated with Klein tunneling through a Weyl point.

## 2.A. Low-energy limit of the four-band model Hamiltonian of a type-II Weyl semimetal

The dispersion along the  $k_z$ -axis (for  $k_x = k_y = 0$ ) of the four-band Hamiltonian (2.11) is given by

$$E_{\pm}^{\text{high}} = \pm b \pm \sqrt{(2t - \mu)^2 + t^2 + 2t(2t - \mu) \cos k_z}, \quad (2.16)$$

$$E_{\pm}^{\text{low}} = \pm b \mp \sqrt{(2t - \mu)^2 + t^2 + 2t(2t - \mu) \cos k_z}. \quad (2.17)$$

For  $\mu > 2t$  the two low-energy bands  $E_{\pm}^{\text{low}}$  form a pair of Weyl cones located at

$$K_z = \pm \arccos \left( \frac{(2t - \mu)^2 + t^2 - b^2}{2t(\mu - 2t)} \right). \quad (2.18)$$

We wish to derive the corresponding low-energy Hamiltonian. Notice that for  $k_x = k_y = 0$  the Hamiltonian (2.11) commutes with  $\sigma_z$  and is thus block-diagonal. Each of the two blocks contains one low and one high energy band. At  $\mathbf{K} = (0, 0, K_z)$  the corresponding low energy eigenstates are given by

$$\Psi_+^{\text{low}} = \frac{1}{N_+} \left( \frac{(2t - \mu)(2b - 2t\sqrt{1 - \cos^2 K_z})}{(2t - \mu)^2 - t^2 + b^2}, 1, 0, 0 \right), \quad (2.19)$$

$$\Psi_-^{\text{low}} = \frac{1}{N_-} \left( 0, 0, -\frac{(2t - \mu)(2b + 2t\sqrt{1 - \cos^2 K_z})}{(2t - \mu)^2 - t^2 + b^2}, 1 \right). \quad (2.20)$$

We expand the four-band Hamiltonian in the basis of these eigenstates:

$$\begin{pmatrix} \langle \Psi_+^{\text{low}} | H | \Psi_+^{\text{low}} \rangle & \langle \Psi_+^{\text{low}} | H | \Psi_-^{\text{low}} \rangle & \langle \Psi_+^{\text{low}} | H | \Psi_+^{\text{high}} \rangle & \langle \Psi_+^{\text{low}} | H | \Psi_-^{\text{high}} \rangle \\ \langle \Psi_-^{\text{low}} | H | \Psi_+^{\text{low}} \rangle & \langle \Psi_-^{\text{low}} | H | \Psi_-^{\text{low}} \rangle & \langle \Psi_-^{\text{low}} | H | \Psi_+^{\text{high}} \rangle & \langle \Psi_-^{\text{low}} | H | \Psi_-^{\text{high}} \rangle \\ \hline \langle \Psi_+^{\text{high}} | H | \Psi_+^{\text{low}} \rangle & \langle \Psi_+^{\text{high}} | H | \Psi_-^{\text{low}} \rangle & \langle \Psi_+^{\text{high}} | H | \Psi_+^{\text{high}} \rangle & \langle \Psi_+^{\text{high}} | H | \Psi_-^{\text{high}} \rangle \\ \langle \Psi_-^{\text{high}} | H | \Psi_+^{\text{low}} \rangle & \langle \Psi_-^{\text{high}} | H | \Psi_-^{\text{low}} \rangle & \langle \Psi_-^{\text{high}} | H | \Psi_+^{\text{high}} \rangle & \langle \Psi_-^{\text{high}} | H | \Psi_-^{\text{high}} \rangle \end{pmatrix} = \left( \begin{array}{c|c} H_{\text{low}} & V_{\text{high,low}} \\ \hline V_{\text{high,low}}^\dagger & H_{\text{high}} \end{array} \right). \quad (2.21)$$

At  $\mathbf{K}$  the high and low energy blocks are uncoupled ( $V_{\text{high,low}}(0, 0, K_z) =$

0). Close to the Weyl point we have

$$\mathcal{H} \approx H_{\text{low}} + V_{\text{high,low}}^\dagger (H_{\text{high}})^{-1} V_{\text{high,low}}. \quad (2.22)$$

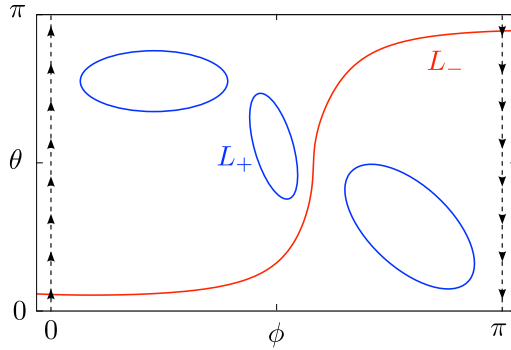
Thus, to linear order in the deviation from  $\mathbf{K}$  we can neglect this coupling and simply linearize  $H_{\text{low}}$ . After some algebra we find the corresponding low-energy Weyl Hamiltonian,

$$\mathcal{H} = a_{\text{tilt}} k_x \tilde{\sigma}_0 - v_x k_x \tilde{\sigma}_x + v_y k_y \tilde{\sigma}_y + v_z (k_z - K_z) \tilde{\sigma}_z. \quad (2.23)$$

The matrices  $\tilde{\sigma}_{0,x,y,z}$  are the identity and the Pauli matrices in the basis of  $|\Psi_\pm^{\text{low}}\rangle$ . The anisotropic velocity components were given in the main text, Eq. (2.12).

Fig. 2.7 shows a comparison between the type-II Weyl cone of the four-band model and its effective low-energy description.

## 2.B. Topological protection of the special magnetic field axis for Klein tunneling between electron and hole pockets



**Figure 2.8.:** Same as Fig. 2.5, but now the incontractible loop  $L_+$  is replaced by a set of contractible loops, containing the entire set of magnetic field axes with extremal contours in the electron pocket that touch the Weyl point. This arrangement would avoid the topological protection of the intersection of incontractible loops in a Möbius strip, but we show by contradiction that it cannot happen in a convex electron pocket.

The topology of the Möbius strip (the projective plane  $\mathbb{P}_2$ ) protects the

intersection of two incontractible loops, ensuring the existence of a special magnetic field axis where the extremal contours  $\mathcal{C}_{\pm}$  in the electron and hole pockets both touch the Weyl point at  $E = 0$ . This is the arrangement shown in Fig. 2.5. Contractible loops can avoid the intersection, as they do in Fig. 2.8. For convex electron and hole Fermi surfaces the existence of incontractable loops is guaranteed by the following argument.

Consider the full set  $\mathcal{S}_+$  of magnetic field axes for which the extremal contour  $\mathcal{C}_+$  in the electron pocket touches the Weyl point. If this set would consist only of contractible loops, then we would be able to pass an incontractible loop  $L$  through  $\mathbb{P}_2$  that avoids  $\mathcal{S}_+$ . We will now see that this leads to a contradiction.

For a convex Fermi surface each field axis  $\hat{\mathbf{n}}$  on  $L$  is associated with a unique extremal contour  $C(\hat{\mathbf{n}})$ . By construction, the contour  $C(\hat{\mathbf{n}})$  lies in a plane normal to  $\hat{\mathbf{n}}$ . The direction  $\hat{\mathbf{n}}$  defines whether the Weyl point lies above or below this plane. Inversion of the axis produces the same extremal contour and therefore the same normal plane, with “above” and “below” interchanged. As we follow the incontractible loop  $L$  from polar angle  $\phi = 0$  to  $\phi = \pi$ , the field axis is inverted, so at some axis  $\hat{\mathbf{n}}_0$  on  $L$  the Weyl point must move from above to below the plane. As motion of the plane is continuous, this can only happen if the Weyl point actually lies on  $C(\hat{\mathbf{n}}_0) \in L$ . This would mean that  $C(\hat{\mathbf{n}}_0) \in \mathcal{S}_+$ , which we had excluded by the construction of  $L$ .

The same argument can be applied to the hole pocket, and we conclude that for both the (convex) electron and hole pockets there must exist incontractible loops  $L_{\pm}$  of field axes with extremal contours that touch the Weyl point.

## 2.C. Klein tunneling for pairs of connected type-II Weyl points

The curves in Fig. 2.6 are calculated as follows. The Fermi level is fixed at the energy  $E = 0$  of the Weyl points and the magnetic field  $\mathbf{B}$  is rotated in the  $x$ - $y$  plane, staying close to the  $y$ -axis (angles  $\theta = 0$ ,  $|\phi/\pi - 1/2| \ll 1$ ). We assume that the dominant  $\phi$ -dependence of the amplitude of the magnetic quantum oscillations is then given by the Klein tunneling probability.

For a given field orientation  $\hat{\mathbf{n}}$  we define  $T(q)$  as the Klein tunneling probability between electron and hole pockets at  $E = 0$  and  $\hat{\mathbf{n}} \cdot \mathbf{k} = q$ . Because of the symmetry of our band structure, both Weyl points have

### 2.C. Klein tunneling for pairs of connected type-II Weyl points

the same  $T$ . We then take a planar cross-section of the Fermi surface at a momentum  $q$  parallel to the field and select one of the contours indicated in the left panels of Fig. 2.6. The contour encloses a signed area  $A(q)$  and we determine the  $q_c$  at which the area is extremal,  $A'(q_c) = 0$ . We calculate  $T_c = T(q_c)$  using the general Landau-Zener formula (2.9). Finally, we follow the contour for one period, collecting a factor  $\sqrt{T_c}$  for each transmission through a Weyl point and a factor  $\sqrt{1 - T_c}$  for each reflection. The product of these factors is plotted in Fig. 2.6 (left panels) as a function of the field orientation  $\phi$ .



### 3. Superconductivity provides access to the chiral magnetic effect of an unpaired Weyl cone

#### 3.1. Introduction

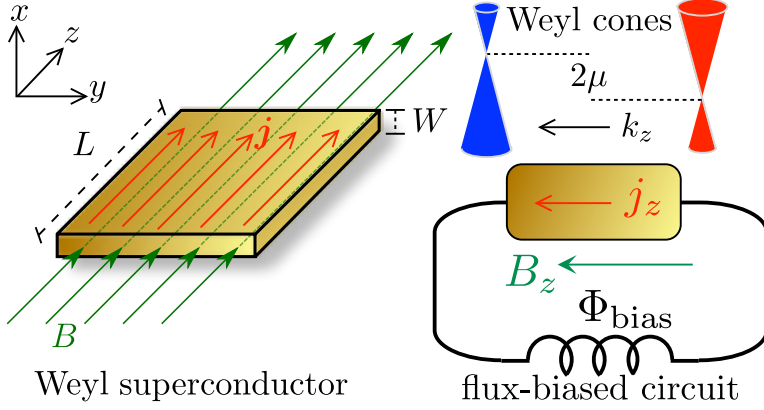
Massless spin-1/2 particles, so-called Weyl fermions, remain unobserved as elementary particles, but they have now been realized as quasiparticles in a variety of crystals known as Weyl semimetals [116–120]. Weyl fermions appear in pairs of left-handed and right-handed chirality, occupying a pair of cones in the Brillouin zone. The pairing is enforced by the chiral anomaly [64]: A magnetic field induces a current of electrons in a Weyl cone, flowing along the field lines in the chiral zeroth Landau level. The current in the Weyl cone of one chirality has to be canceled by a current in the Weyl cone of opposite chirality, to ensure zero net current in equilibrium. The generation of an electrical current density  $\mathbf{j}$  along an applied magnetic field  $\mathbf{B}$ , the so-called chiral magnetic effect (CME) [63, 121], has been observed as a dynamic, nonequilibrium phenomenon [122–126] — but it cannot be realised in equilibrium because of the fermion doubling [65, 127–136].

Here we present a method by which single-cone physics may be accessed in a superconducting Weyl semimetal, allowing for observation of the CME in equilibrium. The geometry is shown in Fig. 3.1. Application of a flux bias gaps out all but a single particle-hole conjugate pair of Weyl cones, of a single chirality  $\pm$  set by the sign of the flux bias. At nonzero chemical potential  $\mu$ , one of the two Weyl points sinks in the Cooper pair sea, the chiral anomaly is no longer cancelled, and we find an equilibrium response

---

The contents of this chapter have been published in T.E. O’Brien, C.W.J. Beenakker and Ī. Adagideli, Phys. Rev. Lett. **118** (20), 207701 (2017).

### 3. Chiral magnetic effect of an unpaired Weyl cone



**Figure 3.1.:** Left panel: Slab of a Weyl superconductor subject to a magnetic field  $B$  in the plane of the slab (thickness  $W$  less than the London penetration depth). The equilibrium chiral magnetic effect manifests itself as a current response  $\partial j / \partial B = \pm \kappa (e/h)^2 \mu$  along the field lines, with  $\kappa$  a charge renormalization factor and  $\mu$  the equilibrium chemical potential. The right panel shows the flux-biased measurement circuit and the charge-conjugate pair of Weyl cones responsible for the effect, of a single chirality  $\pm$  determined by the sign of the flux bias.

$\partial j / \partial B = \pm (e^* e / h^2) \mu$ , with  $e^*$  the charge expectation value at the Weyl point.

We stress that the CME in a superconductor is not in violation of thermodynamics, which only demands a vanishing heat current in equilibrium. Indeed, in previous work on magnetically induced currents [137–139] it was shown that the fundamental principles of Onsager symmetry and gauge invariance forbid a linear relation between  $\mathbf{j}$  and  $\mathbf{B}$  in equilibrium. However, in a superconductor the gauge symmetry is broken at a fixed phase of the order parameter, opening the door for the CME.

## 3.2. Pathway to single-cone physics

We first explain the mechanism by which a superconductor provides access to single-cone physics. A pair of Weyl cones at momenta  $\pm \mathbf{k}_0$  of opposite chirality has Hamiltonian [93]

$$\mathcal{H} = \frac{1}{2} v_F \sum_{\mathbf{k}} [\psi_{\mathbf{k}}^\dagger (\mathbf{k} - \mathbf{k}_0) \cdot \boldsymbol{\sigma} \psi_{\mathbf{k}} - \phi_{\mathbf{k}}^\dagger (\mathbf{k} + \mathbf{k}_0) \cdot \boldsymbol{\sigma} \phi_{\mathbf{k}}], \quad (3.1)$$



### 3.3. Model Hamiltonian of a Weyl superconductor

where  $\mathbf{k} \cdot \boldsymbol{\sigma} = k_x \sigma_x + k_y \sigma_y + k_z \sigma_z$  is the sum over Pauli matrices acting on the spinor operators  $\psi$  and  $\phi$  of left-handed and right-handed Weyl fermions. The Fermi velocity is  $v_F$  and we set  $\hbar \equiv 1$  (but keep  $\hbar$  in the formula for the CME).

If  $\mathcal{H}$  would be the Bogoliubov-De Gennes (BdG) Hamiltonian of a superconductor, particle-hole symmetry would require that  $\phi_{\mathbf{k}} = \sigma_y \psi_{-\mathbf{k}}^\dagger$ . With the help of the matrix identity  $\sigma_y \sigma_\alpha \sigma_y = -\sigma_\alpha^*$  and the anticommutator  $\psi \sigma_\alpha^* \psi^\dagger = -\psi^\dagger \sigma_\alpha \psi$  we rewrite Eq. (3.1) as

$$\begin{aligned} \mathcal{H} &= \frac{1}{2} v_F \sum_{\mathbf{k}} [\psi_{\mathbf{k}}^\dagger (\mathbf{k} - \mathbf{k}_0) \cdot \boldsymbol{\sigma} \psi_{\mathbf{k}} - \psi_{-\mathbf{k}}^\dagger (\mathbf{k} + \mathbf{k}_0) \cdot \boldsymbol{\sigma} \psi_{-\mathbf{k}}] \\ &= v_F \sum_{\mathbf{k}} \psi_{\mathbf{k}}^\dagger (\mathbf{k} - \mathbf{k}_0) \cdot \boldsymbol{\sigma} \psi_{\mathbf{k}}, \end{aligned} \quad (3.2)$$

producing a single-cone Hamiltonian. If we then, hypothetically, impose a magnetic field  $\mathbf{B} = \nabla \times \mathbf{A}$  via  $\mathbf{k} \mapsto \mathbf{k} - e\mathbf{A}$ , the zeroth Landau level carries a current density  $\mathbf{j} = (e/h)^2 \mu \mathbf{B}$  in an energy interval  $\mu$ . This is the chiral anomaly of an unpaired Weyl cone [64].

### 3.3. Model Hamiltonian of a Weyl superconductor

As a minimal model for single-cone physics we consider the BdG Hamiltonian [140]

$$\mathcal{H} = \sum_{\mathbf{k}} \Psi_{\mathbf{k}}^\dagger H(\mathbf{k}) \Psi_{\mathbf{k}}, \quad \Psi_{\mathbf{k}} = (\psi_{\mathbf{k}}, \sigma_y \psi_{-\mathbf{k}}^\dagger), \quad (3.3a)$$

$$H(\mathbf{k}) = \begin{pmatrix} H_0(\mathbf{k} - e\mathbf{A}) & \Delta_0 \\ \Delta_0^* & -\sigma_y H_0^*(-\mathbf{k} - e\mathbf{A}) \sigma_y \end{pmatrix}, \quad (3.3b)$$

$$\begin{aligned} H_0(\mathbf{k}) &= \sum_{\alpha} \tau_z \sigma_{\alpha} \sin k_{\alpha} + \tau_0 (\beta \sigma_z - \mu \sigma_0) + m_{\mathbf{k}} \tau_x \sigma_0, \\ m_{\mathbf{k}} &= m_0 + \sum_{\alpha} (1 - \cos k_{\alpha}). \end{aligned} \quad (3.3c)$$

This is a tight-binding model on a simple cubic lattice (lattice constant  $a_0 \equiv 1$ , nearest-neighbor hopping energy  $t_0 \equiv 1$ , electron charge  $+e$ ). The Pauli matrices  $\tau_{\alpha}$  and  $\sigma_{\alpha}$ , with  $\alpha \in \{x, y, z\}$ , act respectively on the orbital and spin degree of freedom. (The corresponding unit matrices are  $\tau_0$  and  $\sigma_0$ .) Time-reversal symmetry is broken by a magnetization  $\beta$  in the  $z$ -direction,  $\mu$  is the chemical potential,  $\mathbf{A}$  the vector potential, and  $\Delta_0$  is the  $s$ -wave pair potential.

The single-electron Hamiltonian  $H_0$  in the upper-left block of  $H$  is the four-band model [65, 141] of a Weyl semimetal formed from a topological

### 3. Chiral magnetic effect of an unpaired Weyl cone

insulator in the  $\text{Bi}_2\text{Se}_3$  family, layered in the  $x$ - $y$  plane. For a small mass term  $m_0 < \beta$  it has a pair of Weyl cones centered at  $(0, 0, \pm\sqrt{\beta^2 - m_0^2})$ , displaced in the  $k_z$ -direction by the magnetization. (We retain inversion symmetry, so the Weyl points line up at the same energy.) A coupling of this pair of electron Weyl cones to the pair of particle-hole conjugate Weyl cones in the lower-right block of  $H$  is introduced by the pair potential, which may be realized by alternating the layers of topological insulator with a conventional BCS superconductor [142, 143]. (Intrinsic superconducting order in a doped Weyl semimetal, with more unconventional pair potentials, is an alternative possibility [144–153].) The superconductor does not gap out the Weyl cones if  $\Delta_0 < \sqrt{\beta^2 - m_0^2}$ .

### 3.4. Flux bias into the single-cone regime

As explained by Meng and Balents [142], a Weyl superconductor has topologically distinct phases characterized by the number  $\mathcal{N} \in \{2, 1, 0\}$  of ungapped particle-hole conjugate pairs of Weyl cones. We propose to tune through the phase transitions in an externally controllable way by means of a flux bias, as shown in the circuit of Fig. 3.1. For a real  $\Delta_0 > 0$  the flux bias  $\Phi_{\text{bias}}$  enters in the Hamiltonian via the vector potential component  $A_z = \Phi_{\text{bias}}/L \equiv \Lambda/e$ . The  $\Phi_{\text{bias}}$ -dependent band structure is shown in Fig. 3.2, calculated [114] in a slab geometry with hard-wall boundaries at  $x = \pm W/2$  and periodic boundary conditions at  $y = \pm W'/2$  (sending  $W' \rightarrow \infty$ ).

The two pairs of particle-hole conjugate Weyl cones are centered at  $(0, 0, K_{\pm})$  and  $(0, 0, -K_{\pm})$ , with

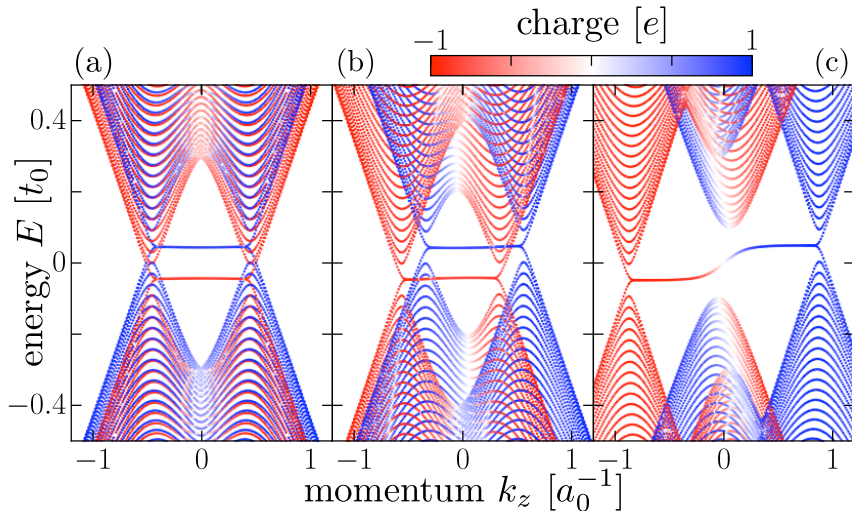
$$K_{\pm}^2 = (\sqrt{\beta^2 - m_0^2} \pm \Lambda)^2 - \Delta_0^2. \quad (3.4)$$

We have assumed  $\Lambda, K_{\pm} \ll 1$ , so the Weyl cones are near the center of the Brillouin zone. A cone is gapped when  $K_{\pm}$  becomes imaginary, hence the  $\mathcal{N} = 1$  phase is entered with increasing  $\Lambda > 0$  when

$$\sqrt{\beta^2 - m_0^2} + \Lambda > \Delta_0 > |\sqrt{\beta^2 - m_0^2} - \Lambda|. \quad (3.5)$$

This is the regime in which we can observe the CME of an unpaired Weyl cone, as we will show in the following.

### 3.5. Magnetic response of an unpaired Weyl cone



**Figure 3.2.:** Effect of a flux bias on the band structure of a Weyl superconductor. The plots are calculated from the Hamiltonian (3.3) in the slab geometry of Fig. 3.1 (parameters:  $m_0 = 0$ ,  $\Delta_0 = 0.2$ ,  $\beta = 0.5$ ,  $\mu = -0.05$ ,  $k_y = 0$ ,  $W = 100$ ,  $B_z = 0$ ). The color scale indicates the charge expectation value, to distinguish electron-like and hole-like cones. As the flux bias is increased from  $\Lambda = 0$  in panel (a), to  $\Lambda = 0.1$  and  $0.4$  in panels (b) and (c), one electron-hole pair of Weyl cones merges and is gapped by the pair potential. What remains in panel (c) is a single pair of charge-conjugate Weyl cones, connected by a surface Fermi arc. This is the phase that supports a chiral magnetic effect in equilibrium.

### 3.5. Magnetic response of an unpaired Weyl cone

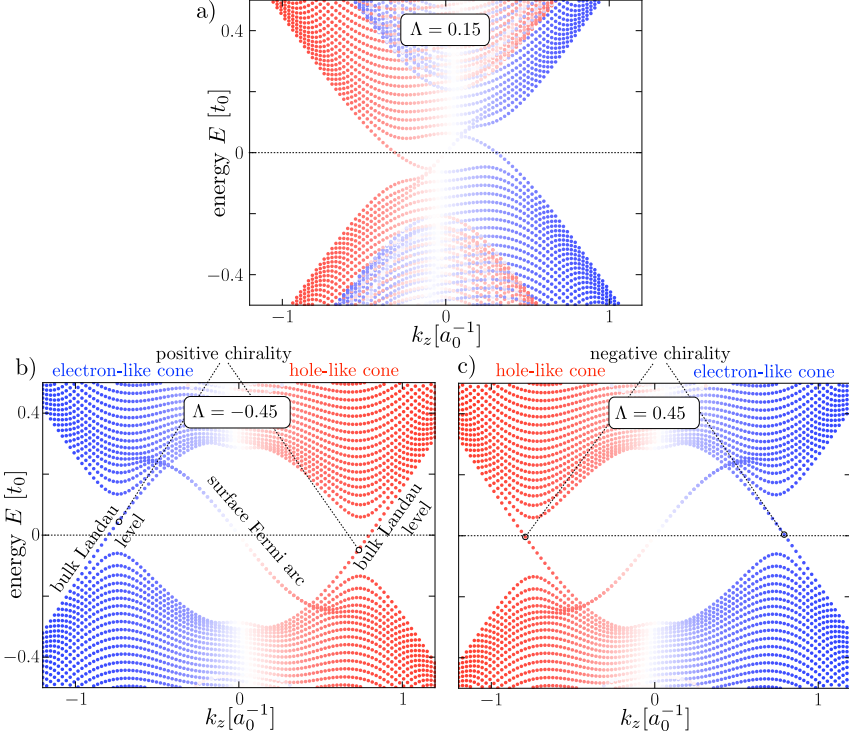
We assume that the slab is thinner than the London penetration depth, so that we can impose an unscreened magnetic field  $B_z$  in the  $z$ -direction\*. The vector potential including the flux bias is  $\mathbf{A} = (0, xB_z, \Lambda/e)$ . To explain in the simplest terms how single-cone physics emerges we linearize in  $\mathbf{k}$  and  $\mathbf{A}$  and set  $m_0 = 0$ , so the mass term  $m_{\mathbf{k}}$  can be ignored. (A discussion of the nonlinearities may be found in App. 3.A.)

The Hamiltonian (3.3) is approximately block-diagonalized by the Bo-

---

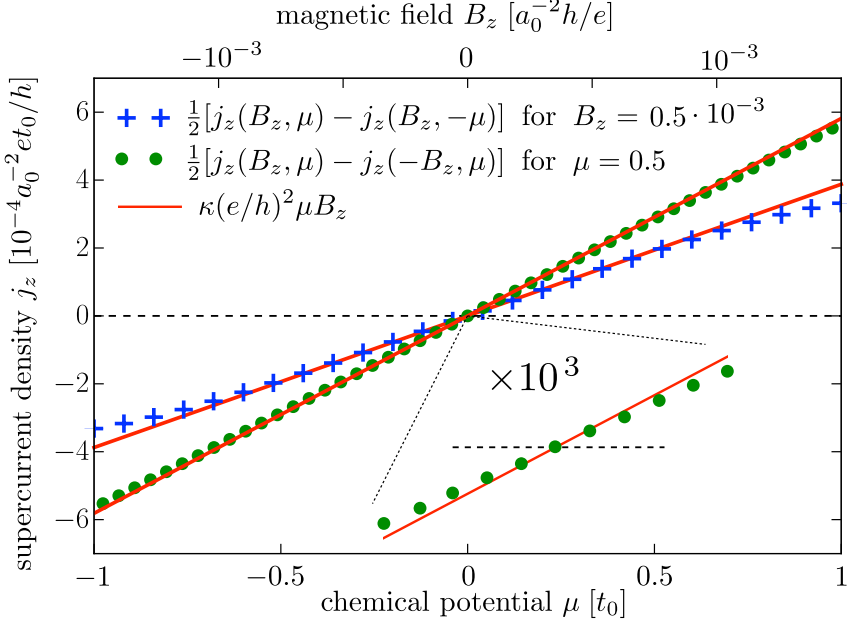
\*For magnetic lengths  $l_m = \sqrt{\hbar/eB}$  below the thickness  $W$  of the slab, the parallel magnetic field  $B$  may induce vortices in the order parameter. We take a uniform order parameter in our analysis, however, the numerical data in Fig. 3.4 shows that our results extend down to the lowest fields with  $l_m \gg W$ , when vortex formation is suppressed.

### 3. Chiral magnetic effect of an unpaired Weyl cone



**Figure 3.3.:** Chirality switch of a pair of charge-conjugate Weyl cones, induced by a sign change of the flux bias  $\Lambda = 0.15$ ,  $-0.45$ , and  $0.45$  in panels a, b, and c, respectively. All other parameters are the same in each panel:  $m_0 = 0$ ,  $\Delta_0 = 0.6$ ,  $\beta = 0.5$ ,  $W = 100$ ,  $k_y = 0$ ,  $\mu = -0.05$ , and  $B_z = 0.001 a_0^{-2} h/e$ . The charge color scale of the band structure is as in Fig. 3.2. Particles in the zeroth Landau level propagate through the bulk in *the same direction* both in the electron-like cone and in the hole-like cone, as determined by the chirality  $\chi = -\text{sign } \Lambda^*$ . A net charge current appears in equilibrium because  $\mu < 0$ , so there is an excess of electron-like states at  $E > 0$ . [States at  $E < 0$  do not contribute to the equilibrium current (3.11).] The particle current is cancelled by the Fermi arc that connects the charge-conjugate Weyl cones. The Fermi arc carries an approximately neutral current, hence the charge current in the chiral Landau level is not much affected by the counterflow of particles in the Fermi arc.

### 3.5. Magnetic response of an unpaired Weyl cone



**Figure 3.4.:** Data points: numerical calculation of the equilibrium supercurrent in the flux-biased circuit of Fig. 3.1. The parameters are  $m_0 = 0$ ,  $\Delta_0 = 0.6$ ,  $\beta = 0.5$ ,  $\Lambda = 0.45$ ,  $W = 100$ ,  $k_B T = 0.01$ ; the green data points are for a fixed  $\mu$  with variation of  $B_z$  and the blue points for a fixed  $B_z$  with variation of  $\mu$ . The data is antisymmetrized as indicated, to eliminate the background supercurrent from the flux bias. The solid curves are the analytical prediction (3.10), with  $\kappa = 0.775$  following directly from Eq. (3.9) (no fit parameters). The  $B_z$ -dependent data is also shown with a zoom-in to very small magnetic fields, down to  $10^{-7} a_0^{-2} h/e$ , to demonstrate that the linear  $B_z$ -dependence continues when  $l_m > W$ .

goliubov transformation

$$\begin{aligned} \tilde{\psi}_{\mathbf{k}} &= \cos(\theta_k/2) \psi_{\mathbf{k}} + i \sin(\theta_k/2) \tau_z \sigma_x \psi_{-\mathbf{k}}^\dagger, \\ \tilde{H} &= U^\dagger H U, \quad U = \exp\left(\frac{i}{2} \theta_k \nu_y \tau_z \sigma_z\right), \end{aligned} \quad (3.6)$$

where the Pauli matrix  $\nu_\alpha$  acts on the particle-hole degree of freedom. If

---

\*The chirality  $\chi$  of a Weyl cone determines the sign of the dispersion of the zeroth Landau level in a magnetic field:  $\text{sign}(dE/dk_z) = \chi \text{sign}(B_z)$ . In the flux-biased Weyl superconductor  $\chi = -\text{sign} \Lambda$ , as one can see in this figure.

### 3. Chiral magnetic effect of an unpaired Weyl cone

we choose the  $k_z$ -dependent angle  $\theta_k$  such that

$$\begin{aligned}\cos \theta_k &= -(\sin k_z)/\Delta_k, \quad \sin \theta_k = \Delta_0/\Delta_k, \\ \Delta_k &= \sqrt{\Delta_0^2 + \sin^2 k_z},\end{aligned}\tag{3.7}$$

the gapless particle-hole conjugate Weyl points at  $k_z^2 = K_+^2 \approx 2\Delta_0(\beta + \Lambda - \Delta_0) \ll 1$  are predominantly contained in the  $(\nu, \tau) = (-, -)$  block of  $\tilde{H}$ . Projection onto this block gives the low-energy Hamiltonian

$$\tilde{\mathcal{H}} = \sum_{\mathbf{k}} \tilde{\psi}_{\mathbf{k}}^\dagger [\sum_{\alpha} v_{\alpha} (\delta k_{\alpha} - q_{\alpha} A_{\alpha}) \sigma_{\alpha} - q_0 \mu \sigma_0] \tilde{\psi}_{\mathbf{k}},\tag{3.8}$$

where  $\mathbf{k} = (0, 0, K_+) + \delta\mathbf{k}$ ,  $\mathbf{v} = (1, 1, -\kappa)$ ,  $q_0 = \kappa$ ,  $\mathbf{q} = (\kappa e, \kappa e, e/\kappa)$ , and

$$\kappa \approx K_+ / \sqrt{\Delta_0^2 + K_+^2} = \sqrt{1 - \Delta_0^2 / (\beta + \Lambda)^2}.\tag{3.9}$$

Eq. (3.8) represents a single-cone Hamiltonian of the form (3.2), with a renormalized velocity  $v_{\alpha}$  and charge  $q_{\alpha}$ . As a consequence, the CME formula for the equilibrium current density  $j_z$  is renormalized into\*

$$\frac{\partial j_z}{\partial B_z} = \frac{q_y q_z}{h^2} q_0 \mu = \frac{e^* e}{h^2} \mu, \quad e^* = \kappa e.\tag{3.10}$$

The renormalization of  $\mathbf{v}$  does not enter because the CME is independent of the Fermi velocity. One can understand why the product  $e^* e$  appears rather than the more intuitive  $(e^*)^2$ , by noticing that  $\partial j_z / \partial B_z$  changes sign upon inversion of the momentum — hence only odd powers of  $\kappa \propto K_+$  are permitted.

### 3.6. Consistency of a nonzero equilibrium electrical current and vanishing particle current

For thermodynamic consistency, to avoid heat transport at zero temperature, the CME should not produce a particle current in the superconductor. The flow of charge  $e^*$  particles in the  $z$ -direction should therefore be cancelled by a charge-neutral counterflow. This counterflow is provided by

---

\*A more formal derivation of the effective charge formula (3.10) for the equilibrium CME is given in App. C

the surface Fermi arc, as illustrated in Fig. 3.3. The Fermi arc is the band of surface states connecting the Weyl cones [154, 155], to ensure that the chirality of the zeroth Landau level does not produce an excess number of left-movers over right-movers. In a Weyl superconductor one can distinguish a trivial or nontrivial connectivity, depending on whether the Fermi arc connects cones of the same or of opposite charge [140, 156]. Here the connectivity is necessarily nontrivial, because there is only a single pair of charge-conjugate Weyl cones. As a consequence, the Fermi arc is approximately charge neutral near the Fermi level (near  $E = 0$ ), so it can cancel the particle current without cancelling the charge current<sup>\*†</sup>.

### 3.7. Numerical simulation

We have tested these analytical considerations in a numerical simulation of the model Hamiltonian (3.3), in the slab geometry of Fig. 3.1. At temperature  $T$  the equilibrium current is given by [159]

$$I_z = \frac{1}{2} \sum_{n,m} \int \frac{dk_z}{2\pi} \tanh\left(\frac{E_{nm}}{2k_B T}\right) \Theta(E_{nm}) \frac{\partial E_{nm}}{\partial A_z}, \quad (3.11)$$

where  $\Theta(E)$  is the unit step function and the prefactor of  $1/2$  takes care of a double counting in the BdG Hamiltonian  $H$ . The eigenvalues  $E_{nm}(k_z)$  of  $H$  are labeled by a pair of mode indices  $n, m$  for motion in the  $x$ - $y$  plane transverse to the current. In Fig. 3.4 we show results for the current density  $j_z = I_z / WW'$  in the  $T = 0$  limit, including a small thermal broadening in the numerics to improve the stability of the calculation.

We see that the numerical data is well described by the analytical result (3.10), with charge renormalization factor  $\kappa = 0.775$  from Eq. (3.9). That analytical formula was derived upon linearization in  $\mathbf{k}$  and  $\mathbf{A}$ . A more accurate calculation<sup>‡</sup> that includes the nonlinear terms in the BdG Hamiltonian gives  $\kappa = 0.750$ , so the simple formula (3.9) is quite accurate.

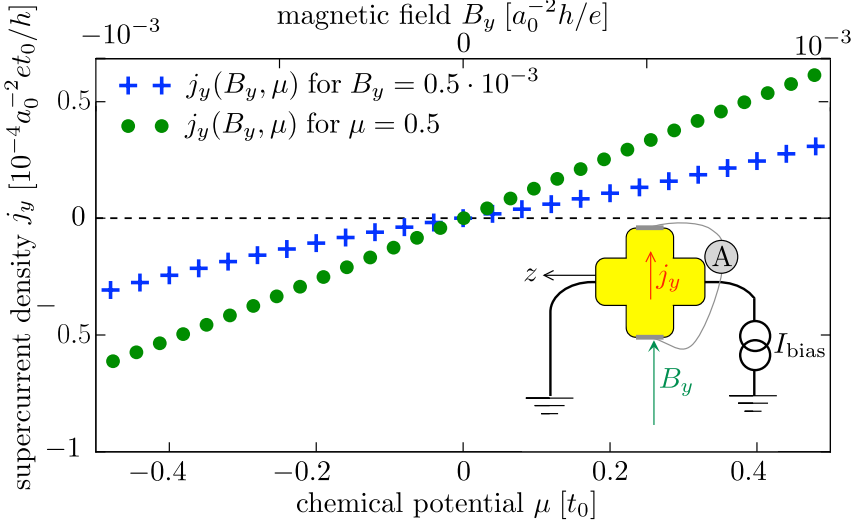
---

<sup>\*</sup>A cancellation of a particle current in the bulk by a particle current at the surface is possible without superconductivity, but then also the charge current is cancelled. For such a spatial separation of counter-propagating particle currents in the normal state see [157, 158].

<sup>†</sup>See App. B for a detailed calculation of the approximately neutral current carried by the surface Fermi arc.

<sup>‡</sup>Details of the calculation of the charge renormalization factor, including all nonlinearities in  $\mathbf{k}$  and  $\mathbf{A}$ , are given in App. A.

### 3. Chiral magnetic effect of an unpaired Weyl cone



**Figure 3.5.:** Same as Fig. 3.4 in the current-biased circuit show in the inset. No antisymmetrization of the data is needed because the measured current is perpendicular to the current bias.

## 3.8. Extensions

We mention extension of our findings that may help to observe the equilibrium CME in an experiment. A first extension is to smaller flux biases in the  $\mathcal{N} = 2$  regime, when two pairs of charge-conjugate cones remain gapless. The supercurrent is then given by

$$\frac{\partial j_z}{\partial B_z} = (\kappa_+ - \kappa_-) \frac{e^2}{h^2} \mu, \quad \kappa_{\pm} = \sqrt{1 - \Delta_0^2 / (\beta \pm \Lambda)^2}, \quad (3.12)$$

so the CME can be observed without fully gapping out one pair of cones.

A second extension is to a current-biased, rather than flux-biased circuit, with the applied magnetic field  $B_y$  perpendicular to the current bias  $j_0$  in the  $z$ -direction. The current bias then drives the Weyl superconductor into the  $\mathcal{N} = 1$  phase via the vector potential component  $A_z = \mu_0 \lambda^2 j_0 \equiv \Lambda/e$ , with  $\lambda$  the London penetration depth [159]. The analytical theory for this alternative configuration is more complicated, and not given here, but numerical results are shown in Fig. 3.5. While the effect is smaller than in the flux-biased configuration, it is not superimposed on a large background supercurrent so it might be more easily observed.



A third extension concerns the inclusion of disorder. Our analysis is simplified by the assumption of a clean slab, without disorder. We expect that the chirality of the zeroth Landau level will protect the equilibrium CME from degradation by impurity scattering, in much the same way as the nonequilibrium CME is protected.

### 3.9. Conclusion

We have shown how the chiral anomaly of an unpaired Weyl cone can be accessed in equilibrium in a superconducting Weyl semimetal. A flux bias drives the system in a state with a single charge-conjugate pair of Weyl cones, that responds to an applied magnetic field as a single species of Weyl fermions. The cancellation of the chiral magnetic effect (CME) for left-handed and right-handed Weyl fermions is removed, resulting in an equilibrium current along the field lines. The predicted size of the induced current is the same as that of the nonequilibrium CME, up to a charge renormalization of order unity, and since that dynamical effect has been observed [122–126] the static counterpart should be observable as well — perhaps even more easily because decoherence and relaxation play no role.

In closing we note that the chiral anomaly in a crystal was originally proposed [64] as a condensed matter realization of an effect from relativistic quantum mechanics, and has since been an inspiration in particle physics and cosmology [160–163]. The doorway to single-cone physics that we have opened here might well play a similar role.

### 3.A. Charge renormalization in a superconducting Weyl cone

We develop an effective low-energy description of the BdG Hamiltonian (3.3), to determine the charge renormalization factors that govern the equilibrium CME. In the main text we gave a simplified description, linearized in  $\mathbf{k}$  and  $\mathbf{A}$ , valid if the Weyl points are near the center of the Brillouin zone. Here we retain the nonlinear terms to obtain more accurate expressions valid throughout the Brillouin zone. As it turns out, our final result (3.34) for the charge renormalization factor is within a few percent of the simple formula (3.9) for the parameters in the simulation of Fig. 3.4.

In this Appendix 3.A we focus on the bulk spectrum, the surface states are considered in the next Appendix 3.B.

### 3. Chiral magnetic effect of an unpaired Weyl cone

#### 3.A.1. Block diagonalization

For a real pair potential  $\Delta_0$  and including the flux bias  $A_z = \Lambda/e$  by the substitution  $k_z \mapsto k_z - \Lambda\nu_z$ , the BdG Hamiltonian is

$$\begin{aligned} H = & \nu_z \tau_z (\sigma_x \sin k_x + \sigma_y \sin k_y + \sigma_z \sin k_z \cos \Lambda) \\ & + m_{\mathbf{k}} \nu_z \tau_x \sigma_0 - \mu \nu_z \tau_0 \sigma_0 + \beta \nu_0 \tau_0 \sigma_z + \Delta_0 \nu_x \tau_0 \sigma_0 \\ & - \nu_0 \tau_z \sigma_z \cos k_z \sin \Lambda - \nu_0 \tau_x \sigma_0 \sin k_z \sin \Lambda, \end{aligned} \quad (3.13)$$

$$m_{\mathbf{k}} = m_0 + (3 - \cos k_x - \cos k_y - \cos k_z \cos \Lambda). \quad (3.14)$$

The  $8 \times 8$  matrix  $H$  is constructed from the tensor product  $\nu_\alpha \tau_\beta \sigma_\gamma \equiv \nu_\alpha \otimes \tau_\beta \otimes \sigma_\gamma$  of the Pauli matrices  $\nu_\alpha$ ,  $\tau_\beta$ ,  $\sigma_\gamma$ , acting respectively on the particle-hole, orbital, and spin degree of freedom.

Adapting the block-diagonalization procedure of Ref. 140, we carry out a sequence of  $k_z$ -dependent unitary transformations,

$$\tilde{H} = U_3^\dagger U_2^\dagger U_1^\dagger H U_1 U_2 U_3, \quad (3.15a)$$

$$\begin{aligned} U_1 &= \exp\left(-\frac{1}{2} i k_z \nu_0 \tau_y \sigma_z\right), \quad U_2 = \exp\left(\frac{1}{2} i \theta \nu_y \tau_z \sigma_z\right), \\ U_3 &= \exp\left(\frac{1}{2} i (\phi_0 \nu_0 + \phi_z \nu_z) \tau_y \sigma_z\right), \end{aligned} \quad (3.15b)$$

where the angles  $\theta, \phi_0, \phi_z$  are determined by

$$\cos \theta = \frac{u_k}{M_0}, \quad \sin \theta = \frac{\Delta_0}{M_0}, \quad (3.16a)$$

$$\cos(\phi_0 \pm \phi_z) = \frac{M_0 \pm \sin \Lambda}{M_\pm}, \quad (3.16b)$$

$$\sin(\phi_0 \pm \phi_z) = \frac{v_k}{M_\pm}, \quad (3.16c)$$

$$u_k = -m_{\mathbf{k}} \sin k_z - \sin k_z \cos k_z \cos \Lambda, \quad (3.16d)$$

$$v_k = m_{\mathbf{k}} \cos k_z - \sin^2 k_z \cos \Lambda, \quad (3.16e)$$

$$M_0 = \sqrt{\Delta_0^2 + u_k^2}, \quad (3.16f)$$

$$M_\pm = \sqrt{(M_0 \pm \sin \Lambda)^2 + v_k^2}. \quad (3.16g)$$

### 3.A. Charge renormalization in a superconducting Weyl cone

We thus arrive at a transformed Hamiltonian,

$$\begin{aligned}\tilde{H} = & \nu_z \tau_z (\sigma_x \sin k_x + \sigma_y \sin k_y) + \beta \nu_0 \tau_0 \sigma_z \\ & - \nu_z \tau_z \sigma_z \sqrt{(M_0 + \nu_z \sin \Lambda)^2 + v_k^2} - \mu \cos \theta \nu_z \tau_0 \sigma_0 \\ & - \mu \sin \theta \cos \phi_0 \nu_x \tau_z \sigma_z - \mu \sin \theta \sin \phi_0 \nu_x \tau_x \sigma_0,\end{aligned}\quad (3.17)$$

that for small  $\mu$  is predominantly block-diagonal in the  $\nu$  and  $\tau$  degree of freedom.

We focus on the parameter range  $M_- < \beta < M_+$  where two of the four Weyl cones are gapped by the phase bias  $\Lambda$ , leaving one gapless particle-hole conjugate pair. The effective low-energy Hamiltonian  $H_{\text{eff}}$  is then obtained by projecting  $\tilde{H}$  onto the  $\nu_z = -1$ ,  $\tau_z = -1$  band,

$$H_{\text{eff}} = \sigma_x \sin k_x + \sigma_y \sin k_y + (\beta - M_-) \sigma_z + \mu \sigma_0 \cos \theta. \quad (3.18)$$

The two Weyl points are at the momenta  $\pm \mathbf{K} = (0, 0, \pm K_z)$  where  $M_- = \beta$ . Near one of the Weyl points, to first order in  $\delta \mathbf{k} = \mathbf{k} - \mathbf{K}$ , the effective Hamiltonian represents an anisotropic Weyl cone:

$$H_{\mathbf{K}} = \sum_{\alpha} v_{\alpha} \delta k_{\alpha} \sigma_{\alpha} + \mu \sigma_0 \cos \theta, \quad (3.19)$$

with effective velocity  $\mathbf{v} = (1, 1, -\partial M_- / \partial k_z)$  evaluated at  $\mathbf{k} = \mathbf{K}$ .

#### 3.A.2. Current and charge operators

The electrical current operator

$$\mathbf{j} = - \lim_{\mathbf{a} \rightarrow 0} \frac{\partial}{\partial \mathbf{a}} H(\mathbf{k} - e \nu_z \mathbf{a}) \quad (3.20)$$

associated with the BdG Hamiltonian (3.13) has components

$$j_x = e \nu_0 \tau_z \sigma_x \cos k_x + e \nu_0 \tau_x \sigma_0 \sin k_x, \quad (3.21a)$$

$$j_y = e \nu_0 \tau_z \sigma_y \cos k_y + e \nu_0 \tau_x \sigma_0 \sin k_y, \quad (3.21b)$$

$$\begin{aligned}j_z = & e \nu_0 \tau_z \sigma_z \cos k_z \cos \Lambda + e \nu_0 \tau_x \sigma_0 \sin k_z \cos \Lambda \\ & + e \nu_z \tau_z \sigma_z \sin k_z \sin \Lambda - e \nu_z \tau_x \sigma_0 \cos k_z \sin \Lambda.\end{aligned}\quad (3.21c)$$

The unitary transformation (3.15) maps this into

$$\tilde{j}_{\alpha} = U_3^{\dagger} U_2^{\dagger} U_1^{\dagger} j_{\alpha} U_1 U_2 U_3, \quad (3.22)$$

### 3. Chiral magnetic effect of an unpaired Weyl cone

resulting in

$$\begin{aligned}
\tilde{J}_x &= e\nu_0\tau_z\sigma_x \cos k_x \cos \theta \\
&- e\nu_0\tau_z\sigma_z \sin k_x [\cos k_z \cos \theta \sin(\phi_0 + \nu_z\phi_z) - \sin k_z \cos(\phi_0 + \nu_z\phi_z)] \\
&+ e\nu_0\tau_x\sigma_0 \sin k_x [\cos k_z \cos \theta \cos(\phi_0 + \nu_z\phi_z) + \sin k_z \sin(\phi_0 + \nu_z\phi_z)] \\
&- e\nu_x\tau_0\sigma_0 \sin k_x \cos k_z \sin \theta \sin \phi_z + e\nu_y\tau_0\sigma_y \cos k_x \sin \theta \cos \phi_0 \\
&+ e\nu_y\tau_y\sigma_z \sin k_x \cos k_z \sin \theta \cos \phi_z - e\nu_y\tau_y\sigma_x \cos k_x \sin \theta \sin \phi_0, \quad (3.23a)
\end{aligned}$$

$$\begin{aligned}
\tilde{J}_y &= e\nu_0\tau_z\sigma_y \cos k_y \cos \theta \\
&- e\nu_0\tau_z\sigma_z \sin k_y [\cos k_z \cos \theta \sin(\phi_0 + \nu_z\phi_z) - \sin k_z \cos(\phi_0 + \nu_z\phi_z)] \\
&+ e\nu_0\tau_x\sigma_0 \sin k_y [\cos k_z \cos \theta \cos(\phi_0 + \nu_z\phi_z) + \sin k_z \sin(\phi_0 + \nu_z\phi_z)] \\
&- e\nu_x\tau_0\sigma_0 \sin k_y \cos k_z \sin \theta \sin \phi_z - e\nu_y\tau_0\sigma_x \cos k_y \sin \theta \cos \phi_0 \\
&+ e\nu_y\tau_y\sigma_z \sin k_y \cos k_z \sin \theta \cos \phi_z - e\nu_y\tau_y\sigma_y \cos k_y \sin \theta \sin \phi_0, \quad (3.23b)
\end{aligned}$$

$$\tilde{J}_z = e\nu_0\tau_z\sigma_z \cos(\Lambda + \phi_0 + \nu_z\phi_z) + e\nu_0\tau_x\sigma_0 \sin(\Lambda + \phi_0 + \nu_z\phi_z). \quad (3.23c)$$

Upon projection onto the  $\nu_z = -1$ ,  $\tau_z = -1$  band we thus arrive at

$$\begin{aligned}
\tilde{J}_x &= e\sigma_z \sin k_x (\cos k_z \cos \theta \sin \phi_- - \sin k_z \cos \phi_-) \\
&- e\sigma_x \cos k_x \cos \theta \quad (3.24a)
\end{aligned}$$

$$\begin{aligned}
\tilde{J}_y &= e\sigma_z \sin k_y (\cos k_z \cos \theta \sin \phi_- - \sin k_z \cos \phi_-) \\
&- e\sigma_y \cos k_y \cos \theta, \quad (3.24b)
\end{aligned}$$

$$\tilde{J}_z = -e\sigma_z \cos(\Lambda + \phi_-) = e\sigma_z \partial M_- / \partial \Lambda. \quad (3.24c)$$

We have abbreviated  $\phi_- \equiv \phi_0 - \phi_z$ .

The corresponding charge operator is simply

$$Q = -e\partial H_{\text{eff}}/\partial\mu = -e\sigma_0 \cos \theta, \quad (3.25)$$

resulting in a charge expectation value

$$\begin{aligned}
\langle Q \rangle &= -e \cos \theta \\
&= \frac{e(3 + m_0 - \cos k_x - \cos k_y) \sin k_z}{\sqrt{\Delta_0^2 + (3 + m_0 - \cos k_x - \cos k_y) \sin^2 k_z}} \quad (3.26)
\end{aligned}$$

of the gapless quasiparticles. The charge changes sign as we move from

### 3.A. Charge renormalization in a superconducting Weyl cone

one Weyl cone at  $\mathbf{K}$  to its particle-hole conjugate at  $-\mathbf{K}$ .

Notice that  $\langle Q \rangle$  is independent of  $A_z = e\Lambda$ . We will make use of this later on to explain why the off-shell contributions to the CME can be neglected [see Eq. (3.56)].

#### 3.A.3. Effective Hamiltonian in the zeroth Landau level

To apply the effective low-energy Hamiltonian (3.18) to the zeroth Landau level we include the vector potential  $\mathbf{A}$  from an applied magnetic field to first order,

$$H_{\text{eff}}(\mathbf{A}) = \sigma_x \sin k_x + \sigma_y \sin k_y + (\beta - M_-)\sigma_z + \mu\sigma_0 \cos \theta - \sum_{\alpha} \tilde{J}_{\alpha} A_{\alpha}. \quad (3.27)$$

We take the vector potential  $\mathbf{A} = (0, B_z x, 0)$  for a magnetic field  $B_z$  in the  $z$ -direction and linearize with respect to  $k_x$ . This linearization also eliminates  $k_x$  from the mass term  $m_{\mathbf{k}}$ , which would otherwise interfere with the  $x$ -dependent  $\mathbf{A}$  when we perform the unitary transformations (3.15). We thus obtain

$$H_{\text{eff}} = \sigma_x k_x + \sigma_y \sin k_y + (\beta - M_-)\sigma_z + \mu\sigma_0 \cos \theta - eB_z x (V_y \sigma_y + V_z \sigma_z), \quad (3.28a)$$

$$V_y = -\cos k_y \cos \theta, \quad (3.28b)$$

$$V_z = \sin k_y (\cos k_z \cos \theta \sin \phi_- - \sin k_z \cos \phi_-). \quad (3.28c)$$

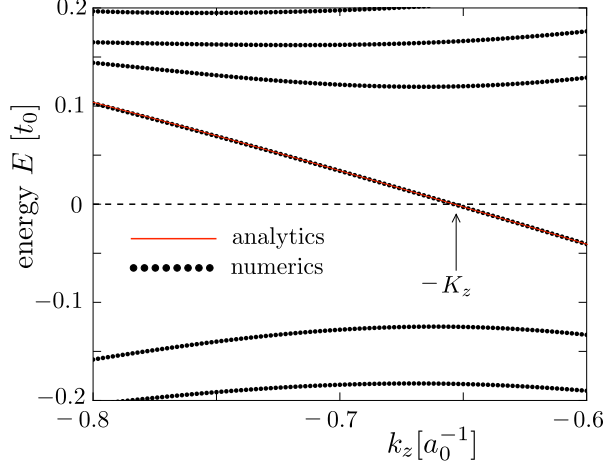
The  $x$  and  $k_x = -i\partial/\partial x$  dependent parts of the Hamiltonian govern the decay of the wave function when  $x \rightarrow \pm\infty$ , according to

$$\begin{aligned} \partial\psi/\partial x &= i\sigma_x eB_z x (V_y \sigma_y + V_z \sigma_z) \psi, \\ \Rightarrow \psi(x) &\propto \exp\left(-\frac{1}{2}eB_z x^2 \sqrt{V_y^2 + V_z^2}\right) |V\rangle, \end{aligned} \quad (3.29a)$$

$$|V\rangle = \begin{pmatrix} V_y + \sqrt{V_y^2 + V_z^2} \\ -iV_z \end{pmatrix}. \quad (3.29b)$$

The energy  $E_0(k_y, k_z)$  of the zeroth Landau level then follows upon

### 3. Chiral magnetic effect of an unpaired Weyl cone



**Figure 3.6.:** Data points: Numerical results for the band structure of the Weyl superconductor near the hole-like Weyl point at  $-K_z$ , showing the first few Landau levels in a magnetic field  $B_z = 5 \cdot 10^{-4} a_0^{-2} \hbar/e$  (other parameters  $m_0 = 0$ ,  $\Delta_0 = 0.6$ ,  $\beta = 0.5$ ,  $\Lambda = 0.45$ ,  $W = 50$ ,  $k_y = 0$ ,  $\mu = 0$ ). Red curve: Analytical result (3.30) in the chiral zeroth Landau level, plotted without any fit parameters.

projection of  $H_{\text{eff}}$  onto  $|V\rangle$ ,

$$\begin{aligned} E_0(k_y, k_z) &= \frac{\langle V | H_{\text{eff}} | V \rangle}{\langle V | V \rangle} \\ &= \frac{(\beta - M_-)V_y - V_z \sin k_y}{\sqrt{V_y^2 + V_z^2}} + \mu \cos \theta. \end{aligned} \quad (3.30)$$

Near each of the two Weyl points at  $\mathbf{k} = (0, 0, \pm K_z) + \delta \mathbf{k}$  this reduces to the dispersion

$$\begin{aligned} E_{\pm}(k_z) &= v_0 \delta k_z - q_{\pm} \mu + \mathcal{O}(\delta k^2), \\ q_{\pm} &= -\cos \theta|_{K_z}, \quad v_0 = -\left. \frac{\partial M_-}{\partial k_z} \right|_{K_z}. \end{aligned} \quad (3.31)$$

of a zeroth Landau level that propagates chirally (unidirectionally) in the  $z$ -direction with the same velocity  $v_0$  and opposite charge  $q_{\pm}$ .

In Fig. 3.6 we compare the dispersion (3.30) in the zeroth Landau

level, derived from the effective low-energy Hamiltonian (3.27), with the numerical result from the full Hamiltonian (3.13). The agreement is very good without any adjustable parameters, giving confidence in the reliability of the low-energy description.

### 3.A.4. Renormalized charge for the CME

To make contact with the single-cone Hamiltonian (3.8) from the main text, we seek the charge and velocity renormalization near the Weyl point at  $\mathbf{K}$ . The current and charge operators (3.24) and (3.25) enter into the effective Hamiltonian (3.19) as

$$H_{\mathbf{K}} = \sum_{\alpha} v_{\alpha} (\delta k_{\alpha} - q_{\alpha} A_{\alpha}) \sigma_{\alpha} - q_0 \mu \sigma_0, \quad (3.32a)$$

$$\left. \begin{aligned} \mathbf{v} &= (1, 1, -\partial M_- / \partial k_z), \\ q_0 &= -\cos \theta, \\ \mathbf{q} &= -e (\cos \theta, \cos \theta, 1 / \cos \theta), \end{aligned} \right\} \text{ at } \mathbf{k} = \mathbf{K}. \quad (3.32b)$$

We have linearized in the momentum  $\delta \mathbf{k} = \mathbf{k} - \mathbf{K}$  and vector potential  $\mathbf{A}$  and we have used the fact that

$$\frac{\partial M_- / \partial \Lambda}{\partial M_- / \partial k_z} = \frac{1}{\cos \theta}. \quad (3.33)$$

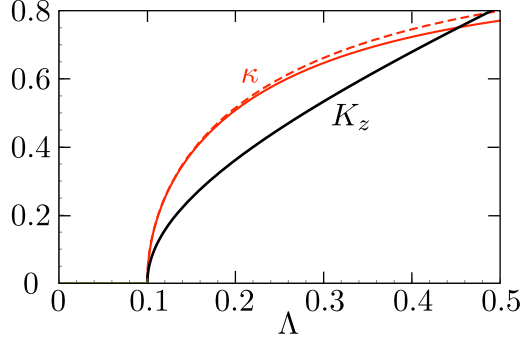
From Eq. (3.10) we find the contribution from the zeroth Landau level to the equilibrium supercurrent density,

$$\begin{aligned} \frac{\partial j_z}{\partial B_z} &= \frac{q_0 q_y q_z}{h^2} \mu = \kappa \frac{e^2}{h^2} \mu, \\ \kappa &= -\cos \theta \big|_{\mathbf{k}=\mathbf{K}} = \frac{(1 + m_0) \sin K_z}{\sqrt{\Delta_0^2 + (1 + m_0)^2 \sin^2 K_z}}, \end{aligned} \quad (3.34)$$

with  $\mathbf{K} = (0, 0, K_z)$  determined by the equation  $M_- = \beta$ .

For the parameter values of Fig. 3.4 we find  $K_z = 0.747$ , resulting in the charge renormalization factor  $\kappa = 0.750$ . The formula (3.9) from the linearized theory in the main text gives  $\kappa = 0.775$  for the same parameter values. It is remarkable how accurate that simple formula is, see Fig. 3.7, even when  $K_z$  is not much smaller than unity.

### 3. Chiral magnetic effect of an unpaired Weyl cone



**Figure 3.7.:** Black curve: Momentum  $K_z$  of the Weyl point as a function of the flux bias  $\Lambda$ , calculated from the solution of  $M_- = \beta$  for the parameters  $m_0 = 0$ ,  $\Delta_0 = 0.6$ ,  $\beta = 0.5$ . Red curves: The corresponding charge renormalization factor  $\kappa$ , from Eq. (3.34) (solid curve) and from the small- $K_z$  approximation (3.9) (dashed curve). The curves terminate at the value  $\Lambda = \Delta_0 - \beta = 0.1$  where a gap opens in the Weyl cone and the solution to  $M_- = \beta$  becomes imaginary.

## 3.B. Surface Fermi arc

In App. 3.A we gave a low-energy description of the bulk Weyl cones. We now turn to the surface states, to derive the dispersion relation shown in Fig. 3.3 of the main text and to demonstrate that the Fermi arc carries an approximately neutral current along the surface.

### 3.B.1. Boundary condition

In the slab geometry of Fig. 3.1 the Weyl superconductor is confined to the inner region  $|x| < W/2$  by an infinite mass in the outer region  $|x| > W/2$ . The requirement of a decaying wave function in the outer region, where  $m_0 \rightarrow \infty$ , implies that the wave function at the interfaces satisfies

$$(1 \pm \nu_0 \tau_y \sigma_x) \psi(\pm W/2) = 0. \quad (3.35)$$



The unitary transformation (3.15) changes this boundary condition into

$$(1 \pm U_b)\tilde{\psi}(\pm W/2) = 0, \quad (3.36)$$

$$\begin{aligned} U_b &= U_3^\dagger U_2^\dagger U_1^\dagger \nu_0 \tau_y \sigma_x U_1 U_2 U_3 \\ &= \nu_0 \tau_0 \sigma_y [\cos k_z \sin(\phi_0 + \nu_z \phi_z) - \cos \theta \sin k_z \cos(\phi_0 + \nu_z \phi_z)] \\ &\quad + \nu_0 \tau_y \sigma_x [\cos k_z \cos(\phi_0 + \nu_z \phi_z) + \cos \theta \sin k_z \sin(\phi_0 + \nu_z \phi_z)] \\ &\quad + \nu_y \tau_z \sigma_x \sin \theta \sin k_z \cos \phi_z + \nu_x \tau_x \sigma_y \sin \theta \sin k_z \sin \phi_z, \end{aligned} \quad (3.37)$$

for the transformed wave function  $\tilde{\psi} = U_3^\dagger U_2^\dagger U_1^\dagger \psi$ .

For later use we note that the two matrices  $U_0 = \nu_z \tau_z \sigma_y$  and  $U_b$  commute, so they can be jointly diagonalized. Each matrix has eigenvalues  $\pm 1$ , we seek the eigenspace where both eigenvalues have same sign. The two orthonormal eigenvectors  $u_1$  and  $u_2$  with eigenvalue  $-1$  are given by

$$u_1 = \frac{1}{2} Z_0^{-2} (iZ_1, Z_1, -iZ_2, Z_2, 0, 0, iZ_4, Z_4), \quad (3.38a)$$

$$\begin{aligned} u_2 &= \frac{1}{2} Z_0^{-2} (i \cos \phi_z Z_4, \cos \phi_z Z_4, i \sin \phi_z Z_4, \\ &\quad - \sin \phi_z Z_4, -iZ_0, Z_0, -iZ_3, -Z_3), \end{aligned} \quad (3.38b)$$

$$Z_0 = 1 - \cos k_z \sin \phi_- + \sin k_z \cos \theta \cos \phi_-, \quad (3.38c)$$

$$Z_1 = \sin \phi_0 \sin k_z \cos \theta + \cos \phi_0 \cos k_z + \sin \phi_z, \quad (3.38d)$$

$$Z_2 = \cos \phi_0 \sin k_z \cos \theta - \sin \phi_0 \cos k_z + \cos \phi_z, \quad (3.38e)$$

$$Z_3 = \cos k_z \cos \phi_- + \sin k_z \cos \theta \sin \phi_-, \quad (3.38f)$$

$$Z_4 = \sin k_z \sin \theta. \quad (3.38g)$$

The eigenspace with eigenvalue  $+1$  of  $U_0$  and  $U_b$  is spanned by  $u_3 = \nu_0 \tau_0 \sigma_z u_1$  and  $u_4 = \nu_0 \tau_0 \sigma_z u_2$ .

### 3.B.2. Construction of the surface state

For  $M_- < \beta < M_+$  there is only one pair of gapless Weyl cones, so there is a single low-energy surface state connecting them. We assume that  $W$  is sufficiently large that we can treat the two surfaces at  $x = \pm W/2$  independently. Let us consider the surface state  $\tilde{\psi}$  at  $x = W/2$ . It should be a solution of  $\tilde{H}\tilde{\psi} = E\tilde{\psi}$  that decays for  $x \rightarrow -\infty$  and that satisfies the boundary condition  $U_b\tilde{\psi} = -\tilde{\psi}$  at  $x = W/2$ .

We first solve this matching problem to zeroth order in  $\mu$ , when the

### 3. Chiral magnetic effect of an unpaired Weyl cone

Hamiltonian (3.17) reduces to

$$\begin{aligned}\tilde{H}_0 = & \nu_z \tau_z (\sigma_x \sin k_x + \sigma_y \sin k_y) + \beta \nu_0 \tau_0 \sigma_z \\ & - \nu_z \tau_z \sigma_z \sqrt{(M_0 + \nu_z \sin \Lambda)^2 + v_k^2}.\end{aligned}\quad (3.39)$$

We linearize in  $k_x = -i\partial/\partial x$  and obtain the solution of  $\tilde{H}_0 \tilde{\psi} = E_0 \tilde{\psi}$  in the form

$$\begin{aligned}\tilde{\psi}(x) = & \exp \left[ i \delta x \nu_z \tau_z \sigma_x (E_0 - \nu_z \tau_z \sigma_y \sin k_y - \beta \nu_0 \tau_0 \sigma_z \right. \\ & \left. + \nu_z \tau_z \sigma_z \sqrt{(M_0 + \nu_z \sin \Lambda)^2 + v_k^2}) \right] \tilde{\psi}(W/2),\end{aligned}\quad (3.40)$$

abbreviating  $\delta x = x - W/2$ .

For  $E_0 = -\sin k_y$  the solution (3.40) that decays for  $\delta x \rightarrow -\infty$  is an eigenvector of  $U_0 = \nu_z \tau_z \sigma_y$  with eigenvalue  $-1$ :

$$\begin{aligned}\tilde{\psi}(x) = & \left( 0, 0, -iC_1 e^{(\beta+M_+)\delta x}, C_1 e^{(\beta+M_+)\delta x}, \right. \\ & \left. -iC_2 e^{(\beta+M_-)\delta x}, C_2 e^{(\beta+M_-)\delta x}, \right. \\ & \left. iC_3 e^{(\beta-M_-)\delta x}, C_3 e^{(\beta-M_-)\delta x} \right).\end{aligned}\quad (3.41)$$

To satisfy the boundary condition at  $x = W/2$ , the coefficients  $C_1, C_2, C_3$  should be chosen such that  $\tilde{\psi}(W/2) = (0, 0, -iC_1, C_1, -iC_2, C_2, iC_3, C_3)$  is a superposition of the eigenvectors  $u_1$  and  $u_2$  in Eq. (3.38). This results in

$$\begin{aligned}C_1 = & Z_1 Z_4 \sin \phi_z + Z_2 Z_4 \cos \phi_z, \quad C_2 = -Z_0 Z_1, \\ C_3 = & Z_1 Z_3 + Z_4^2 \cos \phi_z,\end{aligned}\quad (3.42)$$

up to an overall normalization constant.

### 3.B.3. Surface dispersion relation

We now add to the zeroth order energy  $E_0 = -\sin k_y$  the contribution  $\delta E_\mu$  from the chemical potential in first order perturbation theory,

$$\delta E_\mu = \frac{\langle \tilde{\psi} | \delta \tilde{H} | \tilde{\psi} \rangle}{\langle \tilde{\psi} | \tilde{\psi} \rangle}, \quad (3.43)$$

$$\begin{aligned} \delta \tilde{H} = \tilde{H} - \tilde{H}_0 = & -\mu \cos \theta \nu_z \tau_0 \sigma_0 - \mu \sin \theta \cos \phi_0 \nu_x \tau_z \sigma_z \\ & - \mu \sin \theta \sin \phi_0 \nu_x \tau_x \sigma_0. \end{aligned} \quad (3.44)$$

Two of the three  $\mu$ -dependent terms in  $\delta \tilde{H}$  mix the  $\nu = \pm 1$  bands in the bulk. The small parameter that governs the  $\nu$ -band mixing is  $\delta_{\text{mix}} = (\beta - M_-)/(\beta + M_+)$ . If we neglect this mixing and project both  $\tilde{\psi}$  and  $\tilde{H}$  onto the  $\nu = -1$  band, we have simply

$$\delta E_\mu = \mu \cos \theta. \quad (3.45)$$

In the same way we include to first order the contribution  $\delta E_B$  from the magnetic field with vector potential  $A_y = B_z x$ ,

$$\delta E_B = -B_z \frac{\langle \tilde{\psi} | x \tilde{j}_y | \tilde{\psi} \rangle}{\langle \tilde{\psi} | \tilde{\psi} \rangle} = -\frac{1}{2} W e B_z \cos k_y \cos \theta, \quad (3.46)$$

where we have projected  $\tilde{\psi}$  and  $\tilde{j}_y$  onto the  $\nu = -1$  band and taken the large- $W$  limit of the expectation value.

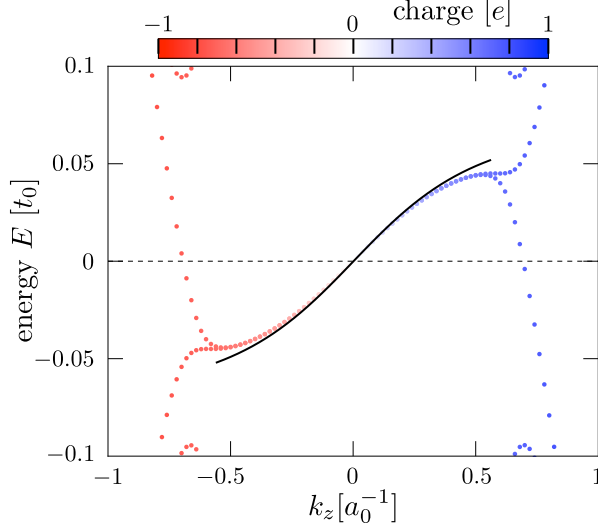
Collecting results we thus obtain the dispersion relation  $E_{\text{surface}}(k_y, k_z)$  for the surface Fermi arc,

$$\begin{aligned} E_{\text{surface}} = & -\sin k_y - \left( \frac{1}{2} W e B_z \cos k_y - \mu \right) \cos \theta \\ = & -\sin k_y \\ & + \frac{(\frac{1}{2} W e B_z \cos k_y - \mu)(2 + m_0 - \cos k_y) \sin k_z}{\sqrt{\Delta_0^2 + (2 + m_0 - \cos k_y) \sin^2 k_z}}. \end{aligned} \quad (3.47)$$

This is for the surface at  $x = W/2$ . For the opposite surface at  $x = -W/2$  we should substitute  $k_y \mapsto -k_y$ .

From Eq. (3.47) we calculate the expectation values of the charge  $\langle Q \rangle$

### 3. Chiral magnetic effect of an unpaired Weyl cone



**Figure 3.8.:** Data points: Dispersion of the surface states connecting the electron-like and hole-like zeroth Landau levels, for the same parameters as Fig. 3.6. The color scale gives the charge expectation value. The black curve is the analytical dispersion (3.47) of the surface Fermi arc.

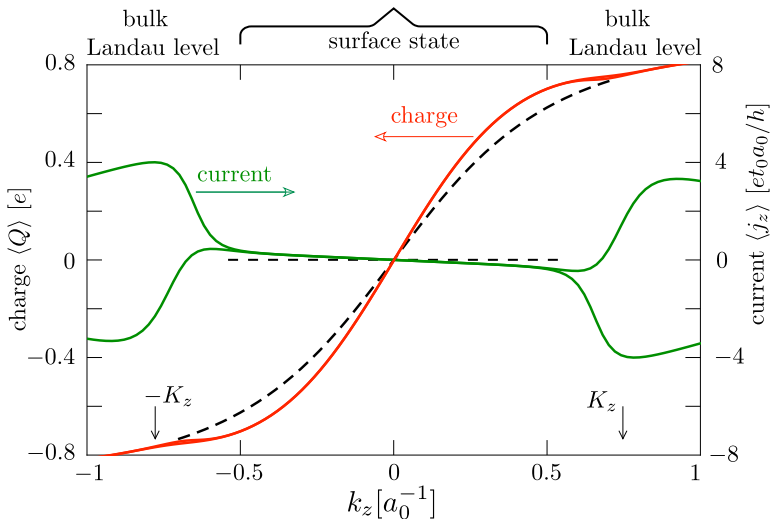
and the electrical current  $\langle j_z \rangle$  of the surface state,

$$\begin{aligned} \langle Q \rangle &= -e \frac{\partial E_{\text{surface}}}{\partial \mu} = -e \cos \theta, \\ \langle j_z \rangle &= -e \frac{\partial E_{\text{surface}}}{\partial \Lambda} = 0, \end{aligned} \tag{3.48}$$

the same on both surfaces. The Fermi arc transports no charge in the  $z$ -direction — up to corrections of order  $\delta_{\text{mix}}$  from the band mixing. The approximately neutral current in a Fermi arc explains why the calculation of the CME including only the chiral Landau level in the bulk agrees so well with the numerics in Fig. 3.5.

In Figs. 3.8 and 3.9 we compare these analytical results for the surface dispersion, charge, and current with the numerical data. The agreement is quite satisfactory, without any adjustable parameter.

### 3.C. Derivation of the renormalized-charge formula for the CME



**Figure 3.9.:** Solid curves: Expectation value of charge  $Q$  (red, left axis) and electrical current  $j_z$  (green, right axis), for the same parameters as Fig. 3.8. The black dashed curves are the analytical result (3.48) for the surface state. The electrical current is predominantly carried by the bulk Landau level, while the surface Fermi arc carries an approximately neutral current.

## 3.C. Derivation of the renormalized-charge formula for the CME

Equation (3.10) in the main text for the equilibrium CME in a superconductor has the form expected for a single Weyl cone, modified by charge renormalization. We give a derivation of this formula.

### 3.C.1. On-shell and off-shell contributions

The equilibrium supercurrent

$$I_z = \frac{1}{2} \sum_{n,m} \int \frac{dk_z}{2\pi} \Theta(E) \tanh\left(\frac{E}{2k_B T}\right) \frac{\partial E}{\partial A_z} \quad (3.49)$$

is not a Fermi surface property, but contains contributions over a range of energies  $E = E_{nm}(k_z) > 0$  even in the limit that the temperature  $T$  goes to zero. For the CME we seek a contribution to  $I_z$  that is linear in

### 3. Chiral magnetic effect of an unpaired Weyl cone

the chemical potential  $\mu$ , measured relative to the Weyl points. As we will now show, the derivative  $\partial I/\partial\mu$  in the limit  $\mu \rightarrow 0$  has predominantly Fermi-surface (*on-shell*) contributions, which at  $T = 0$  can be written as a sum over propagating modes at the Fermi energy  $E = 0$ .

Using particle-hole symmetry (relating states at energy  $\pm E$  carrying opposite current  $\pm\partial E/\partial A_z$ ) we rewrite Eq. (3.49) as an integral over all states of positive and negative energies,

$$I_z = -\frac{1}{2} \sum_{n,m} \int \frac{dk_z}{2\pi} f(E) \frac{\partial E}{\partial A_z}, \quad (3.50)$$

weighted by the Fermi function

$$f(E) = \left(1 + e^{E/k_B T}\right)^{-1} = \frac{1}{2} - \frac{1}{2} \tanh(E/2k_B T). \quad (3.51)$$

The derivative of the energy in Eq. (3.50) gives the expectation value of the electrical current operator  $j_z = -\partial H/\partial A_z$  in the state with energy  $E$ ,

$$\langle j_z \rangle_E = -\langle \partial H/\partial A_z \rangle_E = -\partial E/\partial A_z, \quad (3.52)$$

according to the Hellmann-Feynman theorem. Two other expectation values that we need are those of the velocity operator  $v_z = \partial H/\partial k_z$  and the charge operator  $Q = -e\partial H/\partial\mu$ , given by

$$\langle v_z \rangle_E = \partial E/\partial k_z, \quad \langle Q \rangle_E = -e\partial E/\partial\mu. \quad (3.53)$$

We take the derivative with respect to  $\mu$  of Eq. (3.50):

$$\frac{\partial I_z}{\partial\mu} = \mathcal{J}_{\text{on-shell}} + \mathcal{J}_{\text{off-shell}}, \quad (3.54)$$

$$\mathcal{J}_{\text{on-shell}} = -\frac{1}{2e} \sum_{n,m} \int \frac{dk_z}{2\pi} f'(E) \langle Q \rangle_E \langle j_z \rangle_E, \quad (3.55)$$

$$\mathcal{J}_{\text{off-shell}} = \frac{1}{2e} \sum_{n,m} \int \frac{dk_z}{2\pi} f(E) \frac{\partial}{\partial A_z} \langle Q \rangle_E. \quad (3.56)$$

At low temperatures, when  $-f'(E) \rightarrow \delta(E)$  becomes a delta function, the on-shell contribution  $\mathcal{J}_{\text{on-shell}}$  involves only Fermi surface properties. It is helpful to rewrite it as a sum over modes at the Fermi energy. For that purpose we replace the integration over  $k_z$  by an energy integration

### 3.C. Derivation of the renormalized-charge formula for the CME

weighted with the density of states:

$$\mathcal{J}_{\text{on-shell}} = -\frac{1}{4\pi e} \sum_{n,m} \int_{-\infty}^{\infty} dE f'(E) \left| \frac{\partial E}{\partial k_z} \right|^{-1} \langle Q \rangle_E \langle j_z \rangle_E. \quad (3.57)$$

This equation may be written in a more suggestive form by defining a *vector charge*

$$\mathbf{Q} = (Q_x, Q_y, Q_z), \quad \text{with} \quad Q_\alpha(E) \equiv \frac{\langle j_\alpha \rangle_E}{\langle v_\alpha \rangle_E}, \quad (3.58)$$

which may be different from the average (scalar) charge  $Q_0 \equiv \langle Q \rangle_E$  because the average of the product of charge and velocity may differ from the product of the averages. (For example, the coherent superposition of a right-moving electron and a left-moving hole has zero average charge and zero average velocity, but nonzero average electrical current.) We finally arrive at

$$\begin{aligned} \mathcal{J}_{\text{on-shell}} = & -\frac{1}{4\pi e} \sum_{n,m} \int_{-\infty}^{\infty} dE f'(E) \\ & \times Q_0(E) Q_z(E) (\text{sign} \langle v_z \rangle_E). \end{aligned} \quad (3.59)$$

At zero temperature a sum over modes remains,

$$\mathcal{J}_{\text{on-shell}} = \frac{1}{2} \frac{e}{h} \sum_{n,m} \frac{Q_0 Q_z}{e^2} (\text{sign} \langle v_z \rangle) \Big|_{E_{nm}=0}, \quad (3.60)$$

where we have restored the units of  $\hbar = h/2\pi$ . The subscript  $n, m$  labels the mode indices of a propagating mode in the  $z$ -direction at the Fermi energy ( $E = 0$ ).

#### 3.C.2. Application to the zeroth Landau level

We evaluate Eq. (3.60) for the effective Hamiltonian (3.32) in the zeroth Landau level near the Weyl point at  $\mathbf{K}$  and its charge-conjugate at  $-\mathbf{K}$ . The two Weyl points have opposite sign of both the scalar charge  $Q_0 = -e \cos \theta$  and the vector charge  $Q_z = -e/\cos \theta$ , and the same sign  $\langle v_z \rangle = \chi (\text{sign} B_z)$ , so their contributions add. The Landau level degeneracy is

$$\mathcal{N} = \frac{1}{h} W W' |B_z Q_y| = \frac{e}{h} W W' |B_z \cos \theta|, \quad (3.61)$$

### 3. Chiral magnetic effect of an unpaired Weyl cone

Substitution into Eq. (3.60), times two for two Weyl points, gives the on-shell contribution to the zero-temperature equilibrium current,

$$\mathcal{J}_{\text{on-shell}} = \frac{e}{h} \mathcal{N} \frac{Q_0 Q_z}{e^2} \chi(\text{sign } B_z) = W W' \frac{e^2}{h^2} \kappa \chi B_z, \quad (3.62)$$

with charge renormalization factor

$$\kappa = \lim_{\mathbf{k} \rightarrow \mathbf{K}} |\cos \theta|. \quad (3.63)$$

This confirms Eq. (3.10) in the main text (where we took a positive chirality  $\chi$ ), provided that we can neglect 1) contributions from the surface states; and 2) off-shell contributions from the bulk states. A numerical demonstration that these contributions can be neglected is provided in Fig. 3.5, where the full expression (3.49) is evaluated in a slab geometry. Analytical justification comes from the effective low-energy Hamiltonian, which shows that 1)  $\partial E / \partial A_z = e \partial E / \partial \Lambda$  vanishes on the surface in view of Eq. (3.48); and 2)  $\partial \langle Q \rangle_E / \partial A_z$  vanishes in the bulk in view of Eq. (3.26).



## Part II.

# Topological codes and quantum error correction



## 4. Density-matrix simulation of small surface codes under current and projected experimental noise

### 4.1. Introduction

Recent experimental demonstrations of small quantum simulations [164–166] and quantum error correction (QEC) [167–170] position superconducting circuits for targeting quantum supremacy [171] and quantum fault tolerance [172], two outstanding challenges for all quantum information processing platforms. On the theoretical side, much modeling of QEC codes has been made to determine fault-tolerance threshold rates in various models [173–175] with different error decoders [176–178]. However, the need for computational efficiency has constrained many previous studies to oversimplified noise models, such as depolarizing and bit-flip noise channels. This discrepancy between theoretical descriptions and experimental reality compromises the ability to predict the performance of near-term QEC implementations, and offers limited guidance to the experimentalist through the maze of parameter choices and trade-offs. In the planar circuit quantum electrodynamics (cQED) [179] architecture, the major contributions to error are transmon qubit relaxation, dephasing from flux noise and resonator photons leftover from measurement, and leakage from the computational space, none of which are well-approximated by depolarizing or bit-flip channels. Simulations with more complex error models are now essential to accurately pinpoint the leading contributions to the logical error rate in the small-distance surface codes [173, 176, 180] currently pursued by several groups worldwide.

---

The contents of this chapter have been published in T. E. O’Brien, B. Tarasinski and L. DiCarlo, npj Quant. Inf. **3**, 27 (2017)

#### 4. Density-matrix simulation of small surface codes

In this chapter, we perform a density-matrix simulation of the distance-3 surface code named Surface-17, using the concrete quantum circuit recently proposed in [181] and the measured performance of current experimental multi-transmon cQED platforms [182–185]. For this purpose, we have developed an open-source density-matrix simulation package named `quantumsim` \*. We use `quantumsim` to extract the logical error rate per QEC cycle,  $\epsilon_L$ . This metric allows us to optimize and trade off between QEC cycle parameters, assess the merits of feedback control, predict gains from future improvements in physical qubit performance, and quantify decoder performance. We compare an algorithmic decoder using minimum-weight perfect matching (MWPM) with homemade weight calculation to a simple look-up table (LT) decoder, and weigh both against an upper bound (UB) for decoder performance obtainable from the density-matrix simulation. Finally, we make a low-order approximation to extend our predictions to the distance-5 Surface-49. The combination of results for Surface-17 and -49 allows us to make statements about code scaling and to predict the code size and physical qubit performance required to achieve break-even points for memory and computational performance.

## 4.2. Results

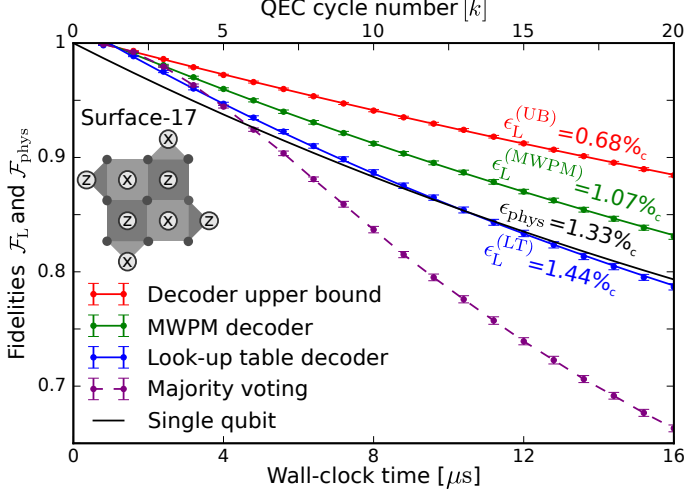
### 4.2.1. Error rates for Surface-17 under current experimental conditions

To quantify the performance of the logical qubit, we first define a test experiment to simulate. Inspired by the recent experimental demonstration of distance-3 and -5 repetition codes [167], we first focus on the performance of the logical qubit as a quantum memory. Specifically, we quantify the ability to hold a logical  $|0\rangle$  state, by initializing this state, holding it for  $k \in \{1, \dots, 20\}$  cycles, performing error correction, and determining a final logical state (see Fig. 4.6 for details). The logical fidelity  $\mathcal{F}_L[k]$  is then given by the probability to match the initial state. We observe identical results when using  $|1\rangle$  or  $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$  in place of  $|0\rangle$ .

We base our error model for the physical qubits on current typical experimental performance for transmons in planar cQED, using parameters from the literature and in-house results (e.g., gate-set tomography measurements). These are summarized in Table 4.1, and further detailed in Table 4.2. We focus on the QEC cycle proposed in [181], which pipelines the execution of  $X$ - and  $Z$ -type stabilizer measurements. Each

---

\*Please visit <https://github.com/brianzi/quantumsim>



**Figure 4.1.:** Logical fidelity  $\mathcal{F}_L[k]$  of Surface-17 with current experimental parameters (Tables 4.1 and 4.2), simulated with quantumsim as described in Fig. 4.6. The results from a MWPM decoder (green) and an implementation of the LT decoder of [176] (blue) are compared to the decoder upper bound (red). The labeled error rate is obtained from the best fit to Eq. (4.2) (also plotted). A further comparison is given to majority voting (purple, dashed), which ignores the outcome of individual stabilizer measurements, and to the fidelity  $\mathcal{F}_{\text{phys}}$  of a single transmon (black) [Eq. (4.1)]. Error bars (2 s.d.) are obtained by bootstrapping.

stabilizer measurement consists of three parts: a coherent step (duration  $\tau_c = 2\tau_{g,1Q} + 4\tau_{g,2Q}$ ), measurement ( $\tau_m$ ), and photon depletion from readout resonators ( $\tau_d$ ), making the QEC cycle time  $\tau_{\text{cycle}} = \tau_c + \tau_m + \tau_d$ .

Simulating this concrete quantum circuit with the listed parameters using quantumsim, we predict  $\mathcal{F}_L[k]$  of Surface-17 (Fig. 4.1). We show  $\mathcal{F}_L[k]$  for both a homemade MWPM decoder (green, described in App. 4.F), and an implementation of the LT decoder of [176] (blue, described in App. 4.G). To isolate decoder performance, we can compare the achieved fidelity to an upper bound extractable from the density-matrix simulation (red, described in Sec. 4.4.1). To assess the benefit of QEC, we also compare to a single decohering transmon, whose fidelity is calculated by averaging over the six cardinal points of the Bloch sphere:

$$\mathcal{F}_{\text{phys}}(t) = \frac{1}{6} \left( 1 + e^{-t/T_1} \right) + \frac{1}{3} \left( 1 + e^{-t(1/2T_1 + 1/T_\phi)} \right). \quad (4.1)$$

#### 4. Density-matrix simulation of small surface codes

The observation of  $\mathcal{F}_L[k] > \mathcal{F}_{\text{phys}}(k\tau_{\text{cycle}})$  for large  $k$  would constitute a demonstration of QEC beyond the quantum memory break-even point [170]. Equivalently, one can extract a logical error rate  $\epsilon_L$  from a best fit to  $\mathcal{F}_L[k]$  (as derived in Sec. 4.4.1 as the probability of an odd number of errors occurring),

$$\mathcal{F}_L[k] = \frac{1}{2}[1 + (1 - 2\epsilon_L)^{k-k_0}]. \quad (4.2)$$

Here,  $k_0$  and  $\epsilon_L$  are the parameters to be fit. We compare  $\epsilon_L$  to the physical error rate

$$\epsilon_{\text{phys}} = -\tau_{\text{cycle}} \left. \frac{d\mathcal{F}_{\text{phys}}(t)}{dt} \right|_{t=0} = \frac{\tau_{\text{cycle}}}{3T_1} + \frac{\tau_{\text{cycle}}}{3T_\phi}. \quad (4.3)$$

We observe  $\epsilon_L = 1.44\%$  for the LT decoder,  $\epsilon_L = 1.07\%$  for the MWPM decoder, and  $\epsilon_L = 0.68\%$  at the decoder upper bound ( $\% = \%$  per cycle). The latter two fall below  $\epsilon_{\text{phys}} = 1.33\%$ . Defining the decoder efficiency  $\eta_d = \epsilon_L^{(\text{UB})}/\epsilon_L$ , we find  $\eta_d^{(\text{LT})} = 0.47$  and  $\eta_d^{(\text{MWPM})} = 0.64$ .

We can also compare the multi-cycle error correction to majority voting, in which the state declaration is based solely on the output of the final data qubit measurements (ancilla measurements are ignored). Majority voting corrects any single data qubit error (over the entire experiment), and thus exhibits a quadratic decay for small  $k$  \*. A decoder should also be able to correct (at least) a single error, and thus should produce the same behavior at low  $k$ , delaying the onset of exponential decay in  $\mathcal{F}_L[k]$ . In fact, a good test for the performance of a MWPM decoder is to ensure it can outperform the majority vote at short timescales, as suboptimal configuration will prevent this (as seen for the look-up table decoder).

With the baseline for current performance established, we next investigate  $\epsilon_L$  improvements that may be achieved by two means. First, we consider modifications to the QEC cycle at fixed physical performance. Afterwards, we consider the effect of improving physical qubit  $T_1$  and  $T_\phi$ .

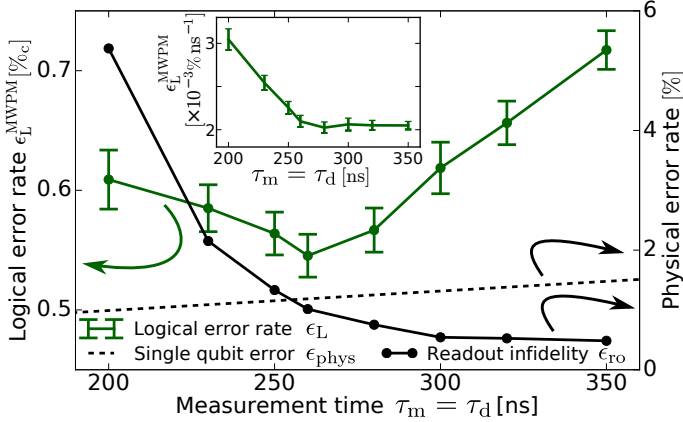
#### 4.2.2. Optimization of logical error rates with current experimental conditions

Error sources in current cQED setups derive primarily from transmon decoherence, as opposed to gate and measurement errors produced by control electronics. Thus, a path to reducing  $\epsilon_L$  may be to decrease

---

\*A distance- $d$  code with majority voting alone should exhibit a  $(d+1)/2$ -order decay

$\tau_{\text{cycle}}$ . Currently, the cycle is dominated by  $\tau_m + \tau_d$ . At fixed readout power, reducing  $\tau_m$  and  $\tau_d$  will reduce  $\tau_{\text{cycle}}$  at the cost of increased readout infidelity  $\epsilon_{\text{RO}}$  (described in Sec. 4.4.2). We explore this trade-off in Fig. 4.2, using a linear-dispersive readout model [186], keeping  $\tau_m = \tau_d$  and assuming no leftover photons. Because of the latter,  $\epsilon_L^{(\text{MWPM})}$  reduces from 1.07 %<sub>c</sub> (Fig. 4.1) to 0.62 %<sub>c</sub> at  $\tau_m = 300$  ns. The minimum  $\epsilon_L^{(\text{MWPM})} = 0.55$  %<sub>c</sub> is achieved at around  $\tau_m = 260$  ns. This is perhaps counterintuitive, as  $\epsilon_{\text{phys}}$  reduces only 0.13 %<sub>c</sub> while  $\epsilon_{\text{RO}}$  increases 0.5 %. However, it reflects the different sensitivity of the code to different types of errors. Indeed,  $\epsilon_L^{(\text{MWPM})}$  is smaller for  $\tau_m = 200$  ns than for  $\tau_m = 300$  ns, even though  $\epsilon_{\text{RO}}$  increases to 5 %. It is interesting to note that the optimal  $\tau_m$  for quantum memory, which minimizes logical error per unit time, rather than per cycle, is  $\tau_m = 280$  ns (Fig. 4.2 inset). This shows that different cycle parameters might be optimal for computation and memory applications.



**Figure 4.2.:** Optimization of the logical error rate (per cycle) of Surface-17 as a function of measurement-and-depletion time [182]. Changes in the underlying physical error rates are shown as well. Decreasing the measurement time causes an increase in the readout infidelity (solid black curve with dots), whilst decreasing the single qubit decay from  $T_1$  and  $T_2$  (black dashed curve) for all qubits. The logical rate with an MWPM decoder (green curve) is minimized when these error rates are appropriately balanced. The logical error rate is calculated from the best fit of Eq. (4.2). Error bars (2 s.d.) are obtained by bootstrapping ( $N = 10,000$  runs). Inset: Logical error rate per unit time, instead of per cycle.

#### 4. Density-matrix simulation of small surface codes

Next, we consider the possibility to reduce  $\epsilon_L$  using feedback control. Since  $T_1$  only affects qubits in the excited state, the error rate of ancillas in Surface-17 is roughly two times higher when in the excited state. The unmodified syndrome extraction circuit flips the ancilla if the corresponding stabilizer value is -1, and since ancillas are not reset between cycles, they will spend significant amounts of time in the excited state. Thus, we consider using feedback to hold each ancilla in the ground state as much as possible. We do not consider feedback on data qubits, as the highly entangled logical states are equally susceptible to  $T_1$ .

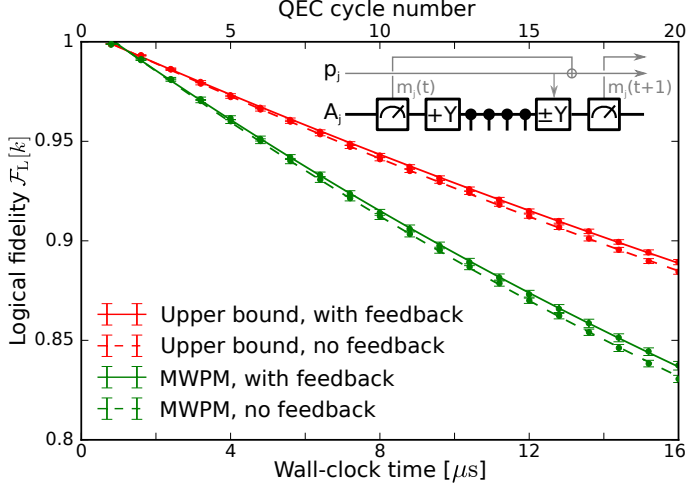
The feedback scheme (Inset of Fig. 3) consists of replacing the  $R_y(\pi/2)$  gate at the end of the coherent step with a  $R_y(-\pi/2)$  gate for some of the ancillas, depending on a classical control bit  $p$  for each ancilla. This bit  $p$  represents an estimate of the stabilizer value, and the ancilla is held in the ground state whenever this estimate is correct (i.e. in the absence of errors). Figure 4.3 shows the effect of this feedback on the logical fidelity, both for the MWPM decoder and the decoder upper bound. We observe  $\epsilon_L$  improve only 0.05 %<sub>c</sub> in both cases. Future experiments might opt not to pursue these small gains in view of the technical challenges added by feedback control.

##### 4.2.3. Projected improvement with advances in quantum hardware

We now estimate the performance increase that may result from improving the transmon relaxation and dephasing times via materials and filtering improvements. To model this, we return to  $\tau_{\text{cycle}} = 800$  ns, and adjust  $T_1$  values with both  $T_\phi = 2T_1$  (common in experiment) and  $T_\phi = \infty$  (all white-noise dephasing eliminated). We retain the same rates for coherent errors, readout infidelity, and photon-induced dephasing as in Fig. 4.1. Figure 4.4 shows the extracted  $\epsilon_L$  and  $\epsilon_{\text{phys}}$  over the  $T_1$  range covered. For the MWPM decoder (upper bound) and  $T_\phi = 2T_1$ , the memory figure of merit  $\gamma_m = \epsilon_{\text{phys}}/\epsilon_L$  increases from 1.3 (2) at  $T_1 = 30$   $\mu\text{s}$  to 2 (5) at 100  $\mu\text{s}$ . Completely eliminating white-noise dephasing will increase  $\gamma_m$  by 10% with MWPM and 30% at the upper bound.

A key question for any QEC code is how  $\epsilon_L$  scales with code distance  $d$ . Computing power limitations preclude similar density-matrix simulations of the  $d = 5$  surface code Surface-49. However, we can approximate the error rate by summing up all lowest-order error chains (as calculated for the MWPM decoder), and deciding individually whether or not these would be corrected by a MWPM decoder (see App. 4.H for details). Figure 4.5 shows the lowest-order approximation to the logical error rates of Surface-17



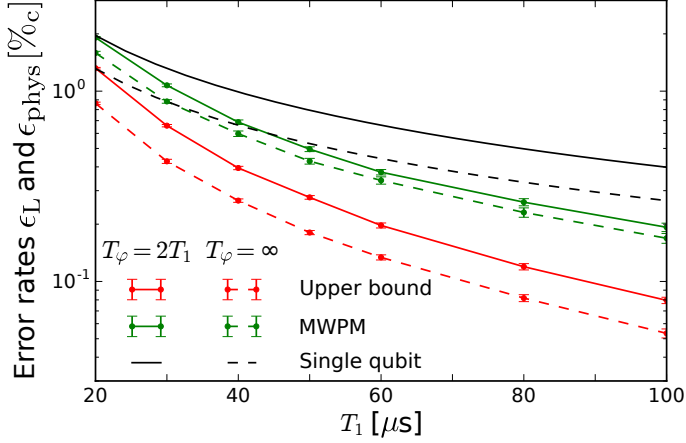


**Figure 4.3.:** Logical fidelity of Surface-17 with (solid) and without (dashed) an additional feedback scheme. The performance of a MWPM decoder (green) is compared to the decoder upper bound (red). Curves are fits of Eq. (4.2) to the data, and error bars (2 s.d.) are given by bootstrapping, with each point averaged over 10,000 runs. Inset: Method for implementing the feedback scheme. For each ancilla qubit  $A_j$ , we store a parity bit  $p_j$ , which decides the sign of the  $R_y(\pi/2)$  rotation at the end of each coherent step. The time  $A_j$  spends in the ground state is maximized when  $p_j$  is updated each cycle  $t$  by XORing with the measurement result from cycle  $t - 1$ , after the rotation of cycle  $t$  has been performed.

and -49 over a range of  $T_1 = T_\phi/2$ . Comparing the Surface-17 lowest-order approximation to the quantumsim result shows good agreement and validates the approximation. We observe a lower  $\epsilon_L$  for Surface-49 than for -17, indicating quantum fault tolerance over the  $T_1$  range covered. The fault-tolerance figure of merit defined in [172],  $\Lambda_t = \epsilon_L^{(17)}/\epsilon_L^{(49)}$ , increases from 2 to 4 as  $T_1$  grows from 30 to 100  $\mu\text{s}$ .

As a rough metric of computational performance, we offer to compare  $\epsilon_L$  (per cycle) to the error accrued by a physical qubit idling over  $\tau_{g,1Q}$ . We define a metric for computation performance,  $\gamma_c = (\epsilon_{\text{phys}}\tau_{g,1Q})/(\epsilon_L\tau_{\text{cycle}})$  and  $\gamma_c = 1$  as a computational break-even point. Clearly, using the QEC cycle parameters of Table 4.1 and even with  $T_1$  improvements, neither Surface-17 nor -49 can break-even computationally. However, including the readout acceleration recently demonstrated in [185], which allows  $\tau_m = \tau_d = 100$  ns and  $\tau_{\text{cycle}} = 400$  ns, Surface-49 can cross  $\gamma_c = 1$  by

#### 4. Density-matrix simulation of small surface codes



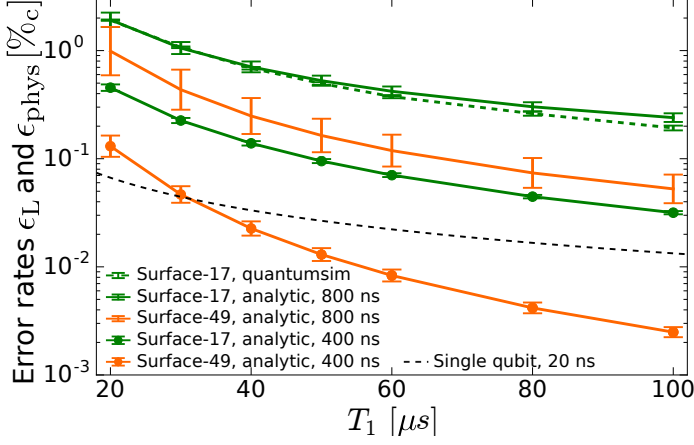
**Figure 4.4.:**  $T_1$  dependence of the Surface-17 logical error rate (MWPM and UB) and the physical error rate. We either fix  $T_\phi = 2T_1$  (solid) or  $T_\phi = \infty$  (dashed). Logical error rates are extracted from a best fit of Eq. (4.2) to  $\mathcal{F}_L[k]$  over  $k = 1, \dots, 20$  QEC cycles, averaged over  $N = 50,000$  runs. Error bars (2 s.d.) are calculated by bootstrapping.

$T_1 = 40 \mu\text{s}$ . In view of first reports of  $T_1$  up to  $80 \mu\text{s}$  emerging for planar transmons [187, 188], this important milestone may be within grasp.

### 4.3. Discussion

#### 4.3.1. Computational figure of merit

We note that our metric of computational power is not rigorous, due to the different gate sets available to physical and logical qubits. Logical qubits can execute multiple logical  $X$  and  $Z$  gates within one QEC cycle, but require a few cycles for two-qubit and Hadamard gates (using the proposals of [175, 180]), and state distillation over many cycles to perform non-Clifford gates. As such, this metric is merely a rough benchmark for computational competitiveness of the QEC code. However, given the amount by which all distance-3 logical fidelities fall above this metric, we find it unlikely that these codes will outperform a physical qubit by any fair comparison in the near future.



**Figure 4.5.:** Analytic approximation of  $\epsilon_L$  for Surface-17 (green) and Surface-49 (orange) using a MWPM decoder. Details of the calculation of points and error bars are given in App. 4.H. All plots assume  $T_\phi = 2T_1$ , and  $\tau_{\text{cycle}} = 800$  ns (crosses) or 400 ns (dots). Numerical results for Surface-17 with  $\tau_{\text{cycle}} = 800$  ns are also plotted for comparison (green, dashed). The physical-qubit computation metric is given as the error incurred by a single qubit over the resting time of a single-qubit gate (black, dashed).

### 4.3.2. Decoder performance

A practical question facing quantum error correction is how best to balance the trade-off between decoder complexity and performance. Past proposals for surface-code computation via lattice surgery [180] require the decoder to provide an up-to-date estimate of the Pauli error on physical qubits during each logical  $T$  gate. Because tracking Pauli errors through a non-Clifford gate is inefficient, however implemented, equivalent requirements will hold for any QEC code [66]. A decoder is thus required to process ancilla measurements from one cycle within the next (on average). This presents a considerable challenge for transmon-cQED implementations, as  $\tau_{\text{cycle}} < 1 \mu\text{s}$ . This short time makes the use of computationally intensive decoding schemes difficult, even if they provide lower  $\epsilon_L$ .

The leading strategy for decoding the surface code is MWPM using the blossom algorithm of Edmonds [173, 177, 189]. Although this algorithm is challenging to implement, it scales linearly in code distance [189]. The algorithm requires a set of weights (representing the probability that two

#### 4. Density-matrix simulation of small surface codes

given error signals are connected by a chain of errors) as input. An important practical question (see App. 4.F) is whether these weights can be calculated on the fly, or must be precalculated and stored. On-the-fly weight calculation is more flexible. For example, it can take into account the difference in error rates between an ancilla measured in the ground and in the excited state. The main weakness of MWPM is the inability to explicitly detect  $Y$  errors. In fact, in App. 4.C we see that MWPM is nearly perfect in the absence of  $Y$  errors. The decoder efficiency  $\eta_d$  may significantly increase by extending MWPM to account for correlations between detected  $X$  and  $Z$  errors originating from  $Y$  errors [190, 191].

If computational limitations preclude a MWPM decoder from keeping up with  $\tau_{\text{cycle}}$ , the look-up table decoder may provide a straightforward solution for Surface-17. However, at current physical performance, the  $\eta_d$  reduction will make Surface-17 barely miss memory break-even (Fig. 4.1). Furthermore, memory requirements make look-up table decoding already impractical for Surface-49. Evidently, real-time algorithmic decoding by MWPM or improved variants is an important research direction already at low code distance.

##### 4.3.3. Other observations

The simulation results allow some further observations. Although we have focused on superconducting qubits, we surmise that the following statements are fairly general.

We observe that small quasi-static qubit errors are suppressed by the repeated measurement. In our simulations, the  $1/f$  flux noise producing 0.01 radians of phase error per flux pulse on a qubit has a diamond norm approximately equal to the  $T_1$  noise, but a trace distance 100 times smaller. As the flux noise increases  $\epsilon_L$  by only 0.01 %, it appears  $\epsilon_L$  is dependent on the trace distance rather than the diamond norm of the underlying noise components. Quasi-static qubit errors can then be easily suppressed, but will also easily poison an experiment if unchecked.

We further observe that above a certain value, ancilla and measurement errors have a diminished effect on  $\epsilon_L$ . In our error model, the leading sources of error for a distance  $d$  code are chains of  $(d-1)/2$  data qubit errors plus either a single ancilla qubit error or readout error, which together present the same syndrome as a chain of  $(d+1)/2$  data qubit errors. An optimal decoder decides which of these chains is more likely, at which point the less-likely chain will be wrongly corrected, completing a logical error. This implies that if readout infidelity ( $\epsilon_{\text{RO}}$ ) or the ancilla error rate ( $\epsilon_{\text{anc}}$ ) is below the data qubit ( $\epsilon_{\text{phys}}$ ) error rate,  $\epsilon_L \propto (\epsilon_{\text{anc}} + \epsilon_{\text{RO}})\epsilon_{\text{phys}}^{(d-1)/2}$ .

However, if  $\epsilon_{\text{RO}} (\epsilon_{\text{anc}}) > \epsilon_{\text{phys}}$ ,  $\epsilon_{\text{L}}$  becomes independent of  $\epsilon_{\text{RO}} (\epsilon_{\text{anc}})$ , to lowest order. This can be seen in Fig. 4.2, where the error rate is almost constant as  $\epsilon_{\text{RO}}$  exponentially increases. This approximation breaks down with large enough  $\epsilon_{\text{anc}}$  and  $\epsilon_{\text{RO}}$ , but presents a counterintuitive point for experimental design;  $\epsilon_{\text{L}}$  becomes less sensitive to measurement and ancilla errors as these error get worse.

A final, interesting point for future surface-code computation is shown in Fig. 4.2: the optimal cycle parameters for logical error rates per cycle and per unit time are not the same. This implies that logical qubits functioning as a quantum memory should be treated differently to those being used for computation. This idea can be extended further: at any point in time, a large quantum computer performing a computation will have a set  $S_m$  of memory qubits which are storing part of a large entangled state, whilst a set  $S_c$  of computation qubits containing the rest of the state undergo operations. To minimize the probability of a logical error occurring on qubits within both  $S_c$  and  $S_m$ , the cycle time of the qubits in  $S_c$  can be reduced to minimize the rest time of qubits in  $S_m$ . As a simple example, consider a single computational qubit  $q_c$  and a single memory qubit  $q_m$  sharing entanglement. Operating all qubits at  $\tau_{\text{cycle}} = 720$  ns to minimize  $\epsilon_{\text{L}}$  would lead to a 1.09% error rate for the two qubits combined. However, shortening the  $\tau_{\text{cycle}}$  of  $q_c$  reduces the time over which  $q_m$  decays. If  $q_c$  operates at  $\tau_{\text{cycle}} = 600$  ns, the average error per computational cycle drops to 1.06%, as  $q_m$  completes only 5 cycles for every 6 on  $q_c$ . Although this is only a meager improvement, one can imagine that when many more qubits are resting than performing computation, the relative gain will be quite significant.

#### 4.3.4. Effects not taken into account

Although we have attempted to be thorough in the detailing of the circuit, we have neglected certain effects. We have used a simple model for C-Z gate errors as we lack data from experimental tomography (e.g. one obtained from two-qubit gate-set tomography [192]). Most importantly, we have neglected leakage, where a transmon is excited out of the two lowest energy states, i.e., out of the computational subspace. Previous experiments have reduced the leakage probability per C-Z gate to  $\sim 0.3\%$  [193], and per single-qubit gate to  $\sim 0.001\%$  [194]. Schemes have also been developed to reduce the accumulation of leakage [195]. Extending quantumsim to include and investigate leakage is a next target. However, the representation of the additional quantum state can increase the simulation effort significantly [by a factor of  $(9/4)^{10} \approx 3000$ ]. To still achieve this goal, some further

#### 4. Density-matrix simulation of small surface codes

approximations or modifications to the simulation will be necessary. Future simulations will also investigate the effect of spread in qubit parameters, both in space (i.e., variation of physical error rates between qubits) and time (e.g.,  $T_1$  fluctuations), and cross-talk effects such as residual couplings between nearest and next-nearest neighbor transmons, qubit cross-driving, and qubit dephasing by measurement pulses targeting other qubits.

### 4.4. Methods

#### 4.4.1. Simulated experimental procedure

##### Surface-17 basics

A QEC code can be defined by listing the data qubits and the stabilizer measurements that are repeatedly performed upon them [68]. In this way, Surface-17 is defined by a  $3 \times 3$  grid of data qubits  $\{D_0, \dots, D_8\}$ . In order to stabilize a single logical qubit,  $9 - 1 = 8$  commuting measurements are performed. The stabilizers are the weight-two and weight-four  $X$ - and  $Z$ -type parity operators  $X_2X_1$ ,  $Z_3Z_0$ ,  $X_4X_3X_1X_0$ ,  $Z_5Z_4Z_2Z_1$ ,  $Z_7Z_6Z_4Z_3$ ,  $X_8X_7X_5X_4$ ,  $Z_8Z_5$ , and  $X_7X_6$ , where  $X_j$  ( $Z_j$ ) denotes the  $X$  ( $Z$ ) Pauli operator acting on data qubit  $D_j$ . Their measurement is realized indirectly using nearest-neighbor interactions between data and ancilla qubits arranged in a square lattices, followed by ancilla measurements [Fig. 4.6(a)]. This leads to a total of 17 physical qubits when a separate ancilla is used for each individual measurement. We follow the circuit realization of this code described in [181], for which we give a schematic description in Fig. 4.6(b) (see App. 4.A for a full circuit diagram).

In an experimental realization of this circuit, qubits will regularly accumulate errors. Multiple errors that occur within a short period of time (e.g., one cycle) form error ‘chains’ that spread across the surface. Errors on single qubits, or correlated errors within a small subregion of Surface-17, fail to commute with the stabilizer measurements, creating error signals that allow diagnosis and correction of the error via a decoder. However, errors that spread across more than half the surface in a short enough period of time are misdiagnosed, causing an error on the logical qubit when wrongly corrected [173]. The rate at which these logical errors arise is the main focus of this chapter.

### Protocol for measurement of logical error rates

As the performance measure of Surface-17, we study the fidelity of the logical qubit as a quantum memory. We describe our protocol with an example ‘run’ in Fig. 4.6. We initialize all qubits in  $|0\rangle$  and perform  $k = 1, 2, \dots, 20$  QEC cycles [Fig. 4.6(b)]. Although this initial state is not a stabilizer eigenstate, the first QEC cycle projects the system into one of the 16 overlapping eigenstates within the  $+1$  eigenspace for  $Z$  stabilizers, which form the logical  $|0\rangle$  state [173]. This implies that, in the absence of errors, the first measurement of the  $Z$  stabilizers will be  $+1$ , whilst that of the  $X$  stabilizers will be random. In the following cycles, ancilla measurements of each run [Fig. 4.6(c)] are processed using a classical decoding algorithm. The decoder computes a Pauli update after each QEC cycle [Fig. 4.6(d)]. This is a best estimate of the Pauli operators that must be applied to the data qubits to transform the logical qubit back to the logical  $|0\rangle$  state. The run ends with a final measurement of all data qubits in the computational basis. From this 9-bit outcome, a logical measurement result is declared [Fig. 4.6(e)]. First, the four  $Z$ -type parities are calculated from the 9 data-qubit measurement outcomes and presented to the decoder as a final set of parity measurements. This ensures that the final computed Pauli update will transform the measurement results into a set that measures  $+1$  for all  $Z$  stabilizers. This results in one of 32 final measurements, from which the value of a logical  $Z$  operator can be calculated to give the measurement result (any choice of logical operator gives the same result). The logical fidelity  $\mathcal{F}_L[k]$  after  $k$  QEC cycles is defined as the probability of this declared result matching the initial  $+1$  state.

At long times and with low error rates, Surface codes have a constant logical error rate  $\epsilon_L$ . The fidelity  $\mathcal{F}_L[k]$  is obtained by counting the probability of an odd number of errors having occurred in total (as two  $\sigma_x$  errors cancel)\* [183]:

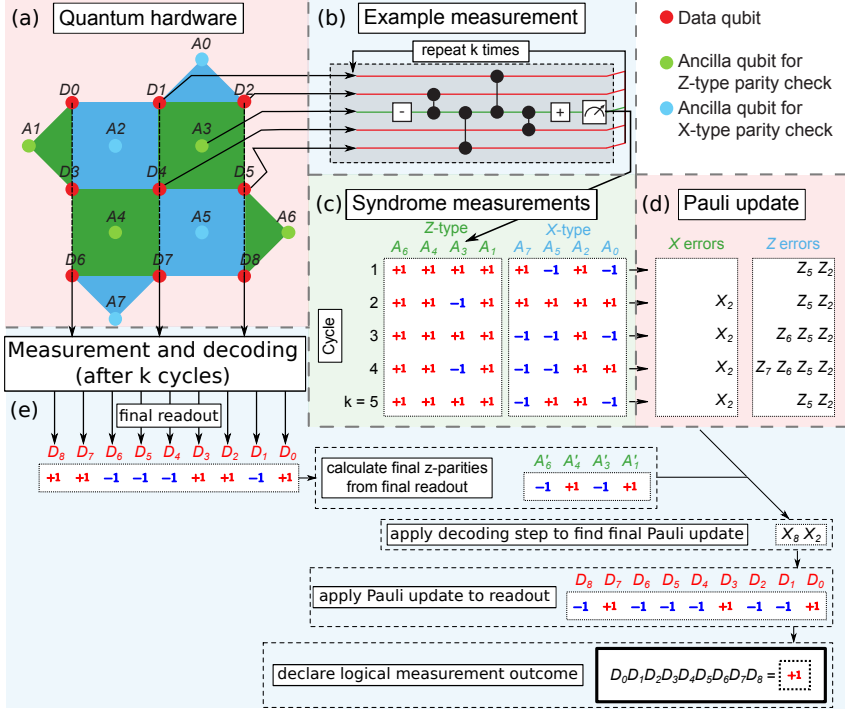
$$\mathcal{F}_L[k] = 1 - \sum_{l \text{ odd}} \binom{k}{l} \epsilon_L^l (1 - \epsilon_L)^{k-l}. \quad (4.4)$$

Here, the combinatorial factor counts the number of combinations of  $l$  errors in  $k$  rounds, given an  $\epsilon_L$  chance of error per round. This can be

---

\*We thank Barbara Terhal for providing this derivation.

#### 4. Density-matrix simulation of small surface codes



**Figure 4.6.:** Schematic overview of the simulated experiment. (a) 17 qubits are arranged in a surface code layout (legend top-right). The red data qubits are initialized in the ground state  $|0\rangle$ , and projected into an eigenstate of the measured X- (blue) and Z- (green) type stabilizer operators. (b) A section of the quantum circuit depicting the four-bit parity measurement implemented by the  $A_3$  ancilla qubit (+/- refer to  $R_y(\pm\pi/2)$  single-qubit rotations). The ancilla qubit (green line, middle) is entangled with the four data qubits (red lines) to measure  $Z_1 Z_2 Z_4 Z_5$ . Ancillas are not reset between cycles. Instead, the implementation relies on the quantum non-demolition nature of measurements. The stabilizer is then the product of the ancilla measurement results of successive cycles. This circuit is performed for all ancillas and repeated  $k$  times before a final measurement of all (data and ancilla) qubits. (c) All syndrome measurements of the  $k$  cycles are processed by the decoder. (d) After each cycle, the decoder updates its internal state to represent the most likely set of errors that occurred. (e) After the final measurement, the decoder uses the readout from the data qubits, along with previous syndrome measurements, to declare a final logical state. To this end, the decoder processes the Z-stabilizers obtained directly from the data qubits, finalizing its prediction of most likely errors. The logical parity is then determined as the product of all data qubit parities ( $\prod_{j=0}^8 D_j$ ) once the declared errors are corrected. The logical fidelity  $\mathcal{F}_L$  is the probability that this declaration is the same as the initial state ( $|0\rangle$ ).



simplified to

$$\begin{aligned}
\mathcal{F}_L[k] &= 1 - \frac{1}{2} \sum_l \binom{k}{l} \epsilon_L^l (1 - \epsilon_L)^{k-l} (1 - (-1)^l) \\
&= 1 - \frac{1}{2} [(1 - \epsilon_L + \epsilon_L)^k - (1 - \epsilon_L - \epsilon_L)^k] \\
&= \frac{1}{2} [1 + (1 - 2\epsilon_L)^k].
\end{aligned} \tag{4.5}$$

However, at small  $k$ , the decay is dominated by the majority vote, for which  $\epsilon_L \propto (k\epsilon_{\text{phys}})^{(d+1)/2}$ . For example, for all the Surface-17 decay curves, we observe a quadratic error rate at small  $k$ , as opposed to the linear slope predicted by Eq. (4.5). In order to correct for this, we shift the above equation in  $k$  by a free parameter  $k_0$ , resulting in Eq. (4.2). This function fits well to data with  $k \geq 3$  in all plots, and thus allows accurate determination of  $\epsilon_L$ .

### The quantumsim simulation package

Quantumsim performs calculations on density matrices utilizing a graphics processing unit in a standard desktop computer. Ancillas are measured at the end of each cycle, and thus not entangled with the rest of the system. As such, it is possible to obtain the effect of the QEC cycle on the system without explicitly representing the density matrix of all 17 qubits simultaneously. The simulation is set up as follows: the density matrix of the nine data qubits is allocated in memory with all qubits initialized to  $|0\rangle$ . One- and two-qubit gates are applied to the density matrix as completely positive, trace preserving maps represented by Pauli transfer matrices. When a gate involving an ancilla qubit must be performed, the density matrix of the system is dynamically enlarged to include that one ancilla.

Qubit measurements are simulated as projective and following the Born rule, with projection probabilities given by the squared overlap of the input state with the measurement basis states. In order to capture empirical measurement errors, we implement a black-box measurement model (Sec. 4.4.2) by sandwiching the measurement between idling processes. The measurement projects the system to a product state of the ancilla and the projected sub-block of the density matrix. We can therefore remove the ancilla from the density matrix and only store its state right after projection, and continue the calculation with the partial density matrix of the other qubits. Making use of the specific arrangement of the

#### 4. Density-matrix simulation of small surface codes

interactions between ancillas and data qubits in Surface-17, it is possible to apply all operations to the density matrix in such an order (shown in App. 4.A) that the total size of the density matrix never exceeds  $2^{10} \times 2^{10}$  (nine data qubits plus one ancilla), which allows relatively fast simulation. We emphasize that with the choice of error model in this chapter, this approach gives the same result as a full simulation on a 17-qubit density matrix. Only the introduction of residual entangling interactions between data and ancilla qubits (which we do not consider in this chapter) would make the latter necessary. On our hardware (see App. 4.E), simulating one QEC cycle of Surface-17 with `quantumsim` takes 25 ms.

We highlight an important advantage of doing density-matrix calculations with `quantumsim`. We do not perform projective measurements of the data qubits. Instead, after each cycle, we extract the diagonal of the data-qubit density matrix, which represents the probability distribution if a final measurement were performed. We leave the density matrix undisturbed and continue simulation up to  $k = 20$ . This is a very useful property of the density-matrix approach, because having a probability distribution of all final readout events greatly reduces sampling noise.

Our measurement model includes a declaration error probability (see Sec. 4.4.2), where the projected state of the ancilla after measurement is not the state reported to the decoder. Before decoding, we thus apply errors to the outcomes of the ancilla projections, and smear the probability distribution of the data qubit measurement. To then determine the fidelity averaged over this probability distribution, we present all 16 possible final  $Z$ -type parities to the decoder. This results in 16 different final Pauli updates, allowing us to determine correctness of the decoder for all 512 possible measurement outcomes. These are then averaged over the simulated probability distribution. This produces good results after about  $\sim 10^4$  simulated runs.

A second highlight of `quantumsim` is the possibility to quantify the sub-optimality of the decoder. The fidelity of the logical qubit obtained in these numerical simulations is a combination of the error rates of the physical qubits and the approximations made by the decoder. Full density-matrix simulations make it possible to disentangle these two contributions. Namely, the fidelity is obtained by assigning correctness to each of the 512 possible readouts according to 16 outputs of the decoder, and summing the corresponding probabilities accordingly. If the probabilities are known, it is easy to determine the 16 results that a decoder should output in order to maximize fidelity (i.e., the output of the best-possible decoder). This allows placing a decoder upper bound  $\mathcal{F}_L^{\max}$  on logical fidelity as limited by the physical qubits independent of the decoder. Conversely, it also allows

quantifying sub-optimality in the decoder used. In fact, we can make the following reverse statement: if our measurement model did not include a declaration error, then we could use the simulation to find the final density matrix of the system conditioned on a syndrome measurement. From this, the simulation could output exactly the 16 results that give  $\mathcal{F}_L^{\max}$ , so that `quantumsim` could thus be used as a maximum-likelihood decoder. In this situation,  $\mathcal{F}_L^{\max}$  would not only be an upper bound, but indeed the performance of the best-possible decoder. However, as we add the declaration errors after simulation, we can only refer to  $\mathcal{F}_L^{\max}$  as the decoder upper bound.

#### 4.4.2. Error models

We now describe the error model used in the simulations. Our motivation for the development of this error model is to provide a limited number of free parameters to study, whilst remaining as close to known experimental data as possible. As such, we have taken well-established theoretical models as a base, and used experimental tomography to provide fixed parameters for observed noise beyond these models. The parameters of the error model are provided in App. 4.B.

Parameter	Symbol	Value	Reference
Qubit relaxation time	$T_1$	30 $\mu\text{s}$	[182]
Qubit dephasing time (white noise)	$T_\phi$	60 $\mu\text{s}$	[182, 184]
Single-qubit gate time	$\tau_{\text{g},1\text{Q}}$	20 ns	[182, 184]
Two-qubit gate time	$\tau_{\text{g},2\text{Q}}$	40 ns	[168]
Coherent step time	$\tau_{\text{c}}$	200 ns	[181]
Measurement time	$\tau_{\text{m}}$	300 ns	[182]
Depletion time	$\tau_{\text{d}}$	300 ns	[182]
Fast measurement time	$\tau_{\text{m}}^{(\text{fast})}$	100 ns	[185]
Fast depletion time	$\tau_{\text{d}}^{(\text{fast})}$	100 ns	[185]

**Table 4.1.:** Standard simulation parameters: Summary of standard times used in all density-matrix simulations, unless otherwise indicated. The two-qubit gate is a conditional phase gate (C-Z). Other error rates and parameters are given in Table 4.2.

#### 4. Density-matrix simulation of small surface codes

##### Idling qubits

While idling for a time  $\tau$ , a transmon in  $|1\rangle$  can relax to  $|0\rangle$ . Furthermore, a transmon in superposition can acquire random quantum phase shifts between  $|0\rangle$  and  $|1\rangle$  due to  $1/f$  noise sources (e.g., flux noise) and broadband ones (e.g., photon shot noise [196] and quasiparticle tunneling [197]). These combined effects can be parametrized by probabilities  $p_1 = \exp(-\tau/T_1)$  for relaxation, and  $p_\phi = \exp(-\tau/T_\phi)$  for pure dephasing. The combined effects of relaxation and pure dephasing lead to decay of the off-diagonal elements of the qubit density matrix. We model dephasing from broadband sources in this way, taking for  $T_\phi$  the value extracted from the decay time  $T_2$  of standard echo experiments:

$$\frac{1}{T_2} = \frac{1}{T_\phi} + \frac{1}{2T_1}. \quad (4.6)$$

We model  $1/f$  sources differently, as discussed below.

##### Dephasing from photon noise

The dominant broadband dephasing source is the shot noise due to photons in the readout resonator. This dephasing is present whenever the coupled qubit is brought into superposition before the readout resonator has returned to the vacuum state following the last measurement. This leads to an additional, time-dependent pure dephasing (rates given in Table 4.2).

##### One-qubit Y rotations

We model  $y$ -axis rotations as instantaneous rotations sandwiched by idling periods of duration  $\tau_{g,1Q}/2$ . The errors in the instantaneous gates are modeled from process matrices measured by gate-set tomography [192, 198] in a recent experiment [183]. In this experiment, the GST analysis of single-qubit gates also showed that the errors can mostly be attributed to Markovian noise. For simplicity, we thus model these errors as Markovian.

##### Dephasing of flux-pulsed qubits

During the coherent step, transmons are repeatedly moved in frequency away from their sweetspot using flux pulses, either to implement a C-Z gate or to avoid one. Away from the sweetspot, transmons become first-order sensitive to flux noise, which causes an additional random phase shift. As this noise typically has a  $1/f$  power spectrum, the largest contribution

comes from low-frequency components that are essentially static for a single run, but fluctuating between different runs. In our simulation, we approximate the effect of this noise through ensemble averaging, with quasi-static phase error added to a transmon whenever it is flux pulsed. Gaussian phase errors with the variance (calculated in App. 4.B.4) are drawn independently for each qubit and for each run.

### C-Z gate error

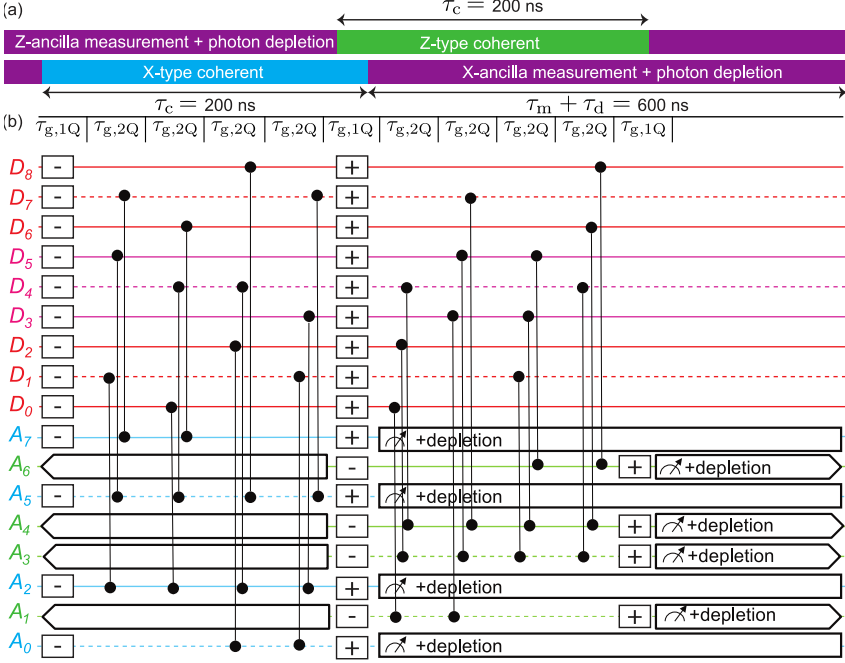
The C-Z gate is achieved by flux pulsing a transmon into the  $|11\rangle \leftrightarrow |02\rangle$  avoided crossing with another, where the 2 denotes the second-excited state of the fluxed transmon. Holding the transmons here for  $\tau_{g,2Q}$  causes the probability amplitudes of  $|01\rangle$  and  $|11\rangle$  to acquire phases [199]. Careful tuning allows the phase  $\phi_{01}$  acquired by  $|01\rangle$  (the single-qubit phase  $\phi_{1Q}$ ) to be an even multiple of  $2\pi$ , and the phase  $\phi_{11}$  acquired by  $|11\rangle$  to be  $\pi$  extra. This extra phase acquired by  $|11\rangle$  is the two-qubit phase  $\phi_{2Q}$ . Single- and two-qubit phases are affected by flux noise because the qubit is first-order sensitive during the gate. Previously, we discussed the single-qubit phase error. In App. 4.B.5, we calculate the corresponding two-qubit phase error  $\delta\phi_{2Q}$ . Our full (but simplistic) model of the C-Z gate consists of an instantaneous C-Z gate with single-qubit phase error  $\delta\phi_{1Q}$  and two-qubit phase error  $\delta\phi_{2Q} = \delta\phi_{1Q}/2$ , sandwiched by idling intervals of duration  $\tau_{g,2Q}/2$ .

### Measurement

We model qubit measurement with a black-box description using parameters obtained from experiment. This description consists of the eight probabilities for transitions from an input state  $|i\rangle \in \{|0\rangle, |1\rangle\}$  into pairs  $(m, |o\rangle)$  of measurement outcome  $m \in \{+1, -1\}$  and final state  $|o\rangle \in \{|0\rangle, |1\rangle\}$ . By final state we mean the qubit state following the photon-depletion period. Input superposition states in the computational bases are first projected to  $|0\rangle$  and  $|1\rangle$  following the Born rule. The probability tree (the butterfly) is then used to obtain an output pair  $(m, |o\rangle)$ . These experimental parameters can be described by a six-parameter model (described in detail in App. 4.B.6), consisting of periods of enhanced noise before and after a point at which the qubit is perfectly projected, and two probabilities  $\epsilon_{RO}^{[i]}$  for wrongly declaring the result of this projective measurement. In App. 4.B.6, a scheme for measuring these butterfly parameters and mapping them to the six-parameter model is described. In experiment, we find that the readout errors  $\epsilon_{RO}^{[i]}$  are almost independent of the qubit state  $|i\rangle$ , and so we

#### 4. Density-matrix simulation of small surface codes

describe them with a single readout error parameter  $\epsilon_{\text{RO}}$  in this chapter.



**Figure 4.7.:** The quantum circuit for Surface-17 syndrome measurement used in all simulations. (a) Outline of the timing of the standard circuit, including the time shift between X- and Z-type stabilizer measurements described by [181]. Qubit labels correspond to the position in Fig. 6. (b) Full quantum circuit of the QEC cycle. The C-Z gates within each group are slightly offset horizontally for visibility (in reality they are performed simultaneously).

### 4.A. Full circuit diagram for Surface-17 implementation

The quantum circuit [181] (Fig. 4.7) consists of  $R_y(\pi/2)$  (“+”) and  $R_y(-\pi/2)$  (“−”) rotations, C-Z gates, and ancilla measurements. The coherent steps of the X and Z ancillas are pipelined (shifted in time with respect to each other) to prevent transmon-transmon avoided crossings. As long as  $\tau_m + \tau_d \geq \tau_c$ , no time is lost due to this separation.

In a simulation of the given circuit, gates on different qubits commute and may be applied to the density matrix in any order, regardless of the times at which they are performed in an experiment. As described in Sec. IV A 3, by simulating gates in a specific order (Fig. 4.8), one can ensure that only one ancilla is ancilla is entangled with the data qubits at any point in the simulation. This allows a reduction in the maximum size of the density matrix from  $2^{17} \times 2^{17}$  to  $2^{10} \times 2^{10}$ .

## 4.B. Parameters of error models

Parameter	Symbol	Value	Reference
In-axis rotation error	$p_{\text{axis}}$	$10^{-4}$	[183]
In-plane rotation error	$p_{\text{plane}}$	$5 \times 10^{-4}$	[183]
$1/f$ flux noise.	$A$	$(1\mu\Phi_0)^2$	[200, 201]
Readout infidelity	$\epsilon_{\text{RO}}$	$5 \times 10^{-3}$	[182]
Photon relaxation time	$1/\kappa$	250 ns	[182]
Dispersive shift	$\chi/\pi$	−2.6 MHz	[182]
photon # post-measurement	$n_0$	0.8 photons	[182]

**Table 4.2.:** Standard parameters of error models used in quantumsim, unless indicated otherwise.

This appendix provides mathematical details of the sources of error described in the main text. Standard values for the parameters used throughout the text are given in Table 4.2.

In the quantumsim module, all gates are applied in the Pauli transfer matrix representation [324]. These are given in the form

$$(R_\Lambda)_{ij} = \frac{1}{2} \text{Tr}(\sigma_i \Lambda \sigma_j), \quad (4.7)$$

where matrices  $\sigma_i$  are the Pauli operators:  $\sigma_0 = I$ ,  $\sigma_1 = X$ ,  $\sigma_2 = Y$  and  $\sigma_3 = Z$ .

### 4.B.1. Qubit idling

Idling qubits are described by the amplitude-phase damping model [23], corresponding to the transfer matrices

$$R_{\Lambda_{T_1}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-p_1} & 0 & 0 \\ 0 & 0 & \sqrt{1-p_1} & 0 \\ p_1 & 0 & 0 & 1-p_1 \end{pmatrix} \quad (4.8)$$

$$R_{\Lambda_{T_\phi}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-p_\phi} & 0 & 0 \\ 0 & 0 & \sqrt{1-p_\phi} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (4.9)$$

Idling for a duration  $t$  is thus described by

$$R_{AP(t)} = R_{\Lambda_{T_1}} R_{\Lambda_{T_\phi}} \quad (4.10)$$

with  $p_1 = 1 - e^{-t/T_1}$  and  $p_\phi = 1 - e^{-t/T_\phi}$ .

### 4.B.2. Photon decay

In the presence of photons in a readout resonator, the coupled qubit is affected according to the effective stochastic master equation [186]:

$$\frac{d\rho}{dt} = -i\frac{B}{2}[\sigma_z, \rho] + \frac{\Gamma_d}{2}\mathcal{D}[\sigma_z]\rho.$$

Here,  $\rho$  is the qubit density matrix,  $\mathcal{D}[X]$  is the Lindblad operator  $\mathcal{D}[X]\rho = X\rho X^\dagger - \frac{1}{2}X^\dagger X\rho - \frac{1}{2}\rho X^\dagger X$ ,  $B = 2\chi\text{Re}(\alpha_g\alpha_e^*)$  is the measurement-induced detuning (Stark shift), and  $\Gamma_d = 2\chi\text{Im}(\alpha_g\alpha_e^*)$  is the measurement-induced dephasing, with  $\alpha_i$  the qubit-state-dependent photon field in the resonator and  $2\chi$  the qubit frequency shift per photon. At time  $t - t_g$  after the qubit superposition is created,

$$\alpha_g\alpha_e^* = \alpha(t_m)\exp(-\kappa(t - t_m))\exp(2i\chi(t - t_g)),$$



with  $t - t_m$  the time since the end of measurement excitation pulse. Integrating over the interval  $[t_1, t_2]$  gives a dephasing term with coefficient

$$\begin{aligned} p_{\phi, \text{photon}} &= \exp\left(-\int_{t_1}^{t_2} \Gamma_d(t) dt\right) \\ &= \exp\left(2\chi\alpha(0) \exp(\kappa(t_m - t_g))\right. \\ &\quad \times \left[\frac{e^{-\kappa t}}{4\chi^2 + \kappa^2} [-\kappa \sin(2\chi t) - 2\chi \cos(2\chi t)]\right]_{t_1 - t_g}^{t_2 - t_g}\bigg). \end{aligned}$$

This dephasing is then implemented via the same Pauli transfer matrix as (4.9).

### 4.B.3. Single-qubit $R_y(\pi/2)$ rotations

Single-qubit rotations are modeled by sandwiching an instantaneous Pauli transfer matrix, representing the rotation, with periods of duration  $\tau_{g,1Q}/2$  of amplitude and phase damping. This allows to model the gate for different  $T_1$  and  $T_\phi$ . However, comparison of this model with Pauli transfer matrices obtained from gate-set tomography experiments shows that actual gates are more accurately described when adding a phenomenological depolarizing noise to the instantaneous part. In the Bloch sphere, this decay corresponds to shrinking toward the origin, with factor  $1 - p_{\text{axis}}$  along the  $y$  axis and  $1 - p_{\text{plane}}$  along the  $x$ - and  $z$ -axes. We thus model

$$R_{R_y(\pi/2)} = R_{AP(\tau_{g,1Q}/2)} R'_{R_y(\pi/2)} R_{\text{dep}} R_{AP(\tau_{g,1Q}/2)}, \quad (4.11)$$

where

$$R_{\text{dep}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - p_{\text{plane}} & 0 & 0 \\ 0 & 0 & 1 - p_{\text{axis}} & 0 \\ 0 & 0 & 0 & 1 - p_{\text{plane}} \end{pmatrix},$$

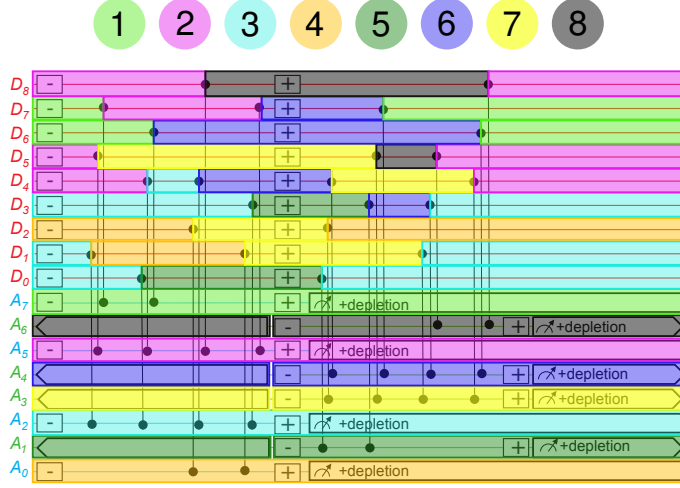
and  $R'_{R_y(\pi/2)}$  is the Pauli transfer matrix describing a perfect  $\pi/2$  rotation around the  $y$  axis.

### 4.B.4. Flux noise

Shifting the transmon from its sweetspot  $f_{q,\text{max}}$  to a lower frequency

$$f_q(t) = (f_{q,\text{max}} + E_C) \sqrt{|\cos(\pi\Phi(t)/\Phi_0)|} - E_C$$

#### 4. Density-matrix simulation of small surface codes



**Figure 4.8.:** Isolation of ancilla interactions in the Surface-17 circuit given in Fig. 4.7. Throughout a simulation, quantumsim stores the density matrix of all data qubits. Each error correction cycle is split up into 8 steps as labeled. In each step, a single ancilla qubit is added to the density matrix, the correspondingly colored pieces of the circuit are executed, and the ancilla is read out and removed from the density matrix. This scheme is only possible because on each data qubit all gates are executed in order. Note that steps after the final C-Z gate on a data qubit are executed during the next cycle.

makes it first-order sensitive to flux noise, with sensitivity

$$\frac{\partial f_q}{\partial \Phi} = \frac{-\pi}{2\Phi_0}(f_q + E_C) \tan\left(\frac{\pi\Phi}{\Phi_0}\right).$$

Here,  $\Phi$  is the flux bias and  $\Phi_0 = h/2e$  is the flux quantum. For a deviation of  $\delta\Phi$ , the pulsed transmon incurs a phase error

$$\delta\phi = -2\pi\tau_{g,2Q} \frac{\partial f_q}{\partial \Phi} \delta\Phi.$$

Flux noise has a characteristic (single-sided) spectral density

$$S_\Phi(f) \approx A/f,$$

where  $A \approx (1 \mu\Phi_0)^2$  with  $f$  in Hz. We model this noise as quasi-static over the duration ( $1/f_{\min} \sim 20 \mu\text{s}$ , or 20 QEC cycles) of individual runs,

but fluctuating between subsequent runs ( $1/f_{\max} \sim 20$  sec, or  $10^5$  runs at  $200 \mu\text{s}$  intervals). The root-mean-square (rms) fluctuations of flux are therefore

$$\begin{aligned}\delta\Phi_{\text{rms}} &= \left( \int_{f_{\min}}^{f_{\max}} S_{\Phi}(f) df \right)^{1/2} \\ &= A(\ln(f_{\max}/f_{\min}))^{1/2} \\ &\approx 4 \mu\Phi_0.\end{aligned}$$

For our quantum circuit based on [181], we estimate the corresponding rms phase error induced in a pulsed transmon to be

$$\delta\phi_{\text{rms}} \approx 0.01 \text{ rad.}$$

#### 4.B.5. C-Z gates

We now focus on the two-qubit phase error. For an adiabatic gate,

$$\phi_{2Q} = \phi_{11} - \phi_{01} = -2\pi \int_{t_1}^{t_2} \zeta(t) dt,$$

with  $t_1$  and  $t_2 = t_1 + \tau_{g,2Q}$  the start and end of the gate and  $\zeta$  the time-dependent frequency deviation of the lower branch of the  $|11\rangle \leftrightarrow |02\rangle$  avoided crossing from the sum of frequencies for  $|01\rangle$  and  $|10\rangle$ . Near the flux center  $\Phi_c$  of the  $|11\rangle - |02\rangle$  avoided crossing,

$$\zeta \approx \beta(\Phi - \Phi_c) - \sqrt{\beta^2(\Phi - \Phi_c)^2 + (2J/2\pi)^2},$$

where  $2J/2\pi \sim 50$  MHz is the minimum splitting between  $|11\rangle$  and  $|02\rangle$ , and

$$\beta = \frac{1}{2} \frac{\partial f_q}{\partial \Phi} \big|_{\Phi=\Phi_c}.$$

Differentiating with respect to  $\Phi$  at  $\Phi_c$  gives

$$\frac{\partial \zeta}{\partial \Phi} \big|_{\Phi=\Phi_c} = \beta.$$

To estimate the  $\delta\phi_{2Q}$  error, we make the following simplification: we replace the exact trajectory created by the flux pulse by a shift to  $\Phi = \Phi_c + \delta\Phi$  with duration  $\tau_{g,2Q}$ . For a deviation of  $\delta\Phi$ ,

$$\delta\phi_{2Q} \approx -2\pi\tau_{g,2Q} \frac{\partial \zeta}{\partial \Phi} \big|_{\Phi=\Phi_c} \delta\Phi.$$

#### 4. Density-matrix simulation of small surface codes

Probability	Value	Probability	Value
$\epsilon_0^{+1,0}$	0.9985	$\epsilon_1^{+1,0}$	0.0050
$\epsilon_0^{+1,1}$	0.0000	$\epsilon_1^{+1,1}$	0.0015
$\epsilon_0^{-1,0}$	0.0015	$\epsilon_1^{-1,0}$	0.0149
$\epsilon_0^{-1,1}$	0.000	$\epsilon_1^{-1,1}$	0.9786

**Table 4.3.:** Measurement butterfly matching a recent characteristic experiment [182] using a Josephson parametric amplifier [203] in phase-preserving mode as the front end of the readout amplification chain.

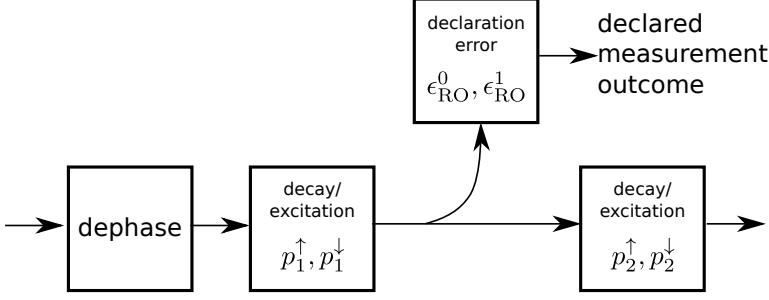
Note that this two-qubit phase error is correlated with the single-qubit phase error on the fluxed transmon. The former is smaller by a factor  $\approx 2$ .

#### 4.B.6. Measurement

The probabilities  $\epsilon_i^{m,o}$  are calibrated using the statistics of outcomes in back-to-back measurements ( $a$  followed by  $b$ ) with the qubit initialized in  $|i\rangle$ .

$$\begin{aligned}
P(m_a = +1)_i &= \epsilon_i^{+1,0} + \epsilon_i^{+1,1}, \\
P(m_a = +1)_i &= \epsilon_i^{-1,0} + \epsilon_i^{-1,1}, \\
P(m_b = m_a = +1)_i &= \left( \epsilon_0^{+1,0} + \epsilon_0^{+1,1} \right) \epsilon_i^{+1,0} \\
&\quad + \left( \epsilon_1^{+1,0} + \epsilon_1^{+1,1} \right) \epsilon_i^{+1,1}, \\
P(m_b = -m_a = +1)_i &= \left( \epsilon_0^{+1,0} + \epsilon_0^{+1,1} \right) \epsilon_i^{-1,0} \\
&\quad + \left( \epsilon_1^{+1,0} + \epsilon_1^{+1,1} \right) \epsilon_i^{-1,1}, \\
P(-m_b = m_a = +1)_i &= \left( \epsilon_0^{-1,0} + \epsilon_0^{-1,1} \right) \epsilon_i^{+1,0} \\
&\quad + \left( \epsilon_1^{-1,0} + \epsilon_1^{-1,1} \right) \epsilon_i^{+1,1}, \\
P(-m_b = -m_a = +1)_i &= \left( \epsilon_0^{-1,0} + \epsilon_0^{-1,1} \right) \epsilon_i^{-1,0} \\
&\quad + \left( \epsilon_1^{-1,0} + \epsilon_1^{-1,1} \right) \epsilon_i^{-1,1}.
\end{aligned}$$

We obtain the six free parameters of the black-box description from these 12 equations, using experimental values on the left-hand side [202]. Table 4.3 shows the values used, achieved in a recent experiment [182]. For the simulation, we reproduce this behaviour of the measurement process by a model with several steps. The qubit undergoes dephasing, followed by periods of decay or excitation between which the measurement result is



**Figure 4.9.:** The model for measurements consists of a dephasing of the qubit followed by a period of decay and excitation with probability  $p_{\downarrow/\uparrow}^{(1)}$ . At this point, the qubit state is sampled. The sampling result is subject to a declaration error  $\epsilon_{\text{RO}}$ , and the qubit state is subject to further decay or excitation with probabilities  $p_{\downarrow/\uparrow}^{(2)}$  before the end of the measurement block.

sampled. This measurement result is further subject to a state-dependent declaration error  $\epsilon_{\text{RO}}$  before reported to the decoder (see Fig.4.9). The six parameters of this model are in a one-to-one correspondence with the butterfly parameters described above, and can be mapped by solving the corresponding system of equations. The experimental results in Tab.4.3 are very well explained by assuming unmodified amplitude-phase damping (with the zero excitation probabilities) during the measurement period, and an outcome-independent declaration error of  $\epsilon_{\text{RO}} = \epsilon_{\text{RO}}^0 = \epsilon_{\text{RO}}^1 = 0.15\%$ . We use this result to extrapolate measurement performance to different values of  $T_1$ .

Reduction of measurement time is expected to reduce assignment fidelity. For the results presented in Fig. 2, we do not rely on experimental results, but assume a simplified model for measurement, following Ref. 186. A constant drive pulse of amplitude  $\epsilon$  and tuned to the bare resonator frequency,  $\Delta_r = 0$ , excites the readout resonator for time  $\tau_m$ . The dynamics of the resonator is dependent on the transmon state (we approximate linear behavior), and the transmitted signal is amplified and detected in a homodyne measurement as a noisy transient. This transient is processed by a linear classifier, which declares the measurement outcome. For resonator depletion, we use a two-step clearing pulse with amplitude  $\epsilon_{c1}$  and  $\epsilon_{c2}$ , each active for  $\tau_d/2$  and chosen (by numerical minimization) so that, at the end of the depletion pulse, the transients for both transmon states return to zero. While the resonator dynamics is easily found if the transmon is in the ground state, amplitude damping of the transmon in the excited

#### 4. Density-matrix simulation of small surface codes

state leads to non-deterministic behavior. We thus numerically obtain an ensemble of noisy transients for each input qubit state, and optimize the decision boundary of the linear classifier for this ensemble. Generating a second verification ensemble, the “butterfly” of the measurement setup is estimated.

The dynamics of the resonator is determined by the resonator linewidth  $\kappa$  as well as the dispersive shift  $\chi$ . We chose the parameters of the setup used in [182],  $1/\kappa = 250$  ns and  $\chi/\pi = -2.6$  MHz. The signal-to-noise ratio of the detected transient is reduced by the quantum efficiency  $\eta = 12.5\%$ . The driving strength  $\epsilon$  is chosen to approximate the “butterfly” used in most of the main text, and corresponds to a steady-state average photon population of about  $\bar{n} = 15$ . We then keep  $\epsilon$  constant while changing the measurement time, keeping  $\tau_m = \tau_d$ , to obtain the butterflies used in the density matrix simulation. We ignore effects leading to measurement-induced mixing and non-linearity of the readout resonator. Finally, since these simulations do not allow to make a realistic prediction about residual photon numbers achievable in experiments, we ignore this effect when using these results.

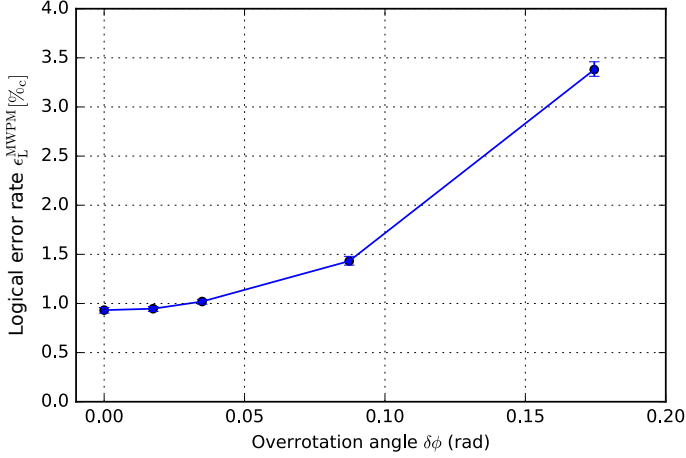
### 4.C. Effect of over-rotations and two-qubit phase noise on logical error rate

In this section we provide additional numerical data showing the effect of some common noise sources on the logical error rate. In Fig. 4.10 we show the effect of a coherent over-rotation, whereby the  $R'_Y(\pi/2)$  operator in Eq. 4.11 is replaced by  $R'_Y(\pi/2 + \delta\phi)$ . This can be caused by inaccurate calibration of the flux pulse used to perform the gate. In Fig. 4.11 we show the effect of an increase in the two-qubit flux noise  $\delta\phi_{\text{rms}}$  as described in Sec. 4.B.4.

### 4.D. Calculation of decoder upper bound

We provide a detailed description how the decoder upper bound is obtained from the simulation results. As described in the main text, after each cycle of simulation, the diagonal of the reduced density matrix of the data qubits in the Z basis is stored. It contains the probability distribution for the  $2^9 = 512$  different possible measurement outcomes of the data qubits. In the quantum memory experiment described in the main text, each of

#### 4.D. Calculation of decoder upper bound



**Figure 4.10.:** Logical error rate for Surface-17 as a function of single-qubit over-rotation, using the MWPM decoder. Other parameters are as given in the main text.

these outcomes are passed to the decoder, which then declares a logical measurement outcome.

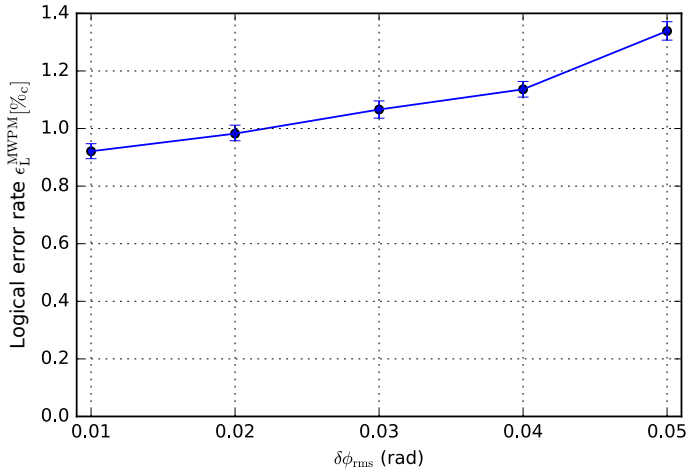
It is evident that any decoder must declare opposite logical outcomes if two of the 512 possible measurements  $m$  and  $m'$  are related by the application of a logical X operator. Thus, any decoder can give the correct result only for half of the measurement outcomes. Subject to this constraint, we can find the set of 256 declarations which maximize the probability that the declaration is correct. It immediately follows that no decoder can achieve a declaration fidelity larger than this maximal probability. We thus refer to it as the decoder upper bound.

In practice, the upper bound is found according to the following approach. Since declarations are opposite if two outcomes differ by a logical X operator, they must be equal if they differ by the application of one or more X stabilizers (applying two different logical X operators amounts to the application of a product of X stabilizers). We thus group the outcomes in 32 cosets which are related by the application of X-stabilizers. (There are 4 X-stabilizers in Surface-17, so there are  $512/2^4 = 32$  cosets). For outcomes from the same coset, the declaration from a decoder must be the same. We obtain the probability of a final measurement falling within each coset by summing the probabilities from the density matrix diagonal. We further group the 32 cosets to 16 pairs, which differ by the application

#### 4. Density-matrix simulation of small surface codes

of a logical operator. The upper bound is then obtained by selecting the more probable coset from each pair and summing the corresponding probabilities. This upper bound can also be interpreted as the internal decoherence of the logical qubit: it represents the maximal overlap of the final state with the initial state, under any possible correction of errors.

We finally emphasize that this upper bound can be found only because we have access to the complete probability distribution of outcomes (for a given result of syndrome measurements), a major advantage of the density matrix simulation. However, we do not expect that any decoder can actually achieve this upper bound: This is because we add syndrome measurement events independently after the situation, which will decrease the logical error rate further.



**Figure 4.11.:** Logical error rate for Surface-17 as a function of two-qubit phase error, using the MWPM decoder. Other parameters are as given in the main text.

### 4.E. Hardware requirements of simulation

The simulations are performed using the `quantumsim` package \*, which were developed by the authors for this chapter. The package is accelerated by performing the density matrix manipulations on a GPU (graphics

---

\*The `quantumsim` package can be found at <http://github.com/brianzi/quantumsim>



card). The simulations for this chapter were performed on a NVidia Tesla K40 GPU, on which we observed runtimes of about 0.5 seconds for the simulation of a run of  $k=20$  cycles (25 ms per QEC cycle). We also had the opportunity to test the software on a more modern GPU (NVidia Tesla P100), observing about 15 ms per cycle, and on a consumer-grade GPU (NVidia Quadro M2000), observing about 40 ms per cycle. By comparison, the CPU is mostly idle during the simulation, except for handling of input and output. The memory requirements are modest for both CPU and GPU RAM. They are dominated by the storage of the density matrices and amount to a few ten megabytes.

## 4.F. Homemade MWPM decoder with asymmetric weight calculation

Every QEC code requires a decoder to track the most likely errors consistent with a given set of stabilizer measurements. The MWPM decoder has gained popularity since it was shown to have threshold values above 1% [177]. The motivation behind MWPM is that single  $X$  or  $Z$  errors on data qubits in the bulk of a surface-code fabric cause changes of two stabilizers in the code. These signals can then be considered vertices on a graph, with the error the edge connecting them. Errors in measurement, or errors on a single ancilla qubit, behave as changes in the stabilizer that are separated in time. Multiple errors that would join the same vertices create longer paths in the graph, of which an experiment only records the endpoints. Thus, the problem becomes that of finding the most likely set of generating errors given the error signals that mark their ends. This is made slightly simpler, as in the surface code any chain of errors that forms a closed loop does not change the logical state. This implies that all paths that connect two points are equivalent, and can be considered together. The problem then is to join error signals, either in pairs, or to a ‘boundary’ vertex. The latter corresponds to errors on data qubits at the boundary, which belong to only one  $X$  or  $Z$  stabilizer. This pairing  $P$  should be chosen as the most likely combination of single-qubit errors that could generate the measured error signals. This has then been reduced to the problem of minimum-weight perfect matching on a graph, which can be solved in polynomial time by the blossom algorithm [173, 209].

The MWPM decoder we use differs from previous methods by its weight calculation. As part of the decoding process, it is required to calculate to some degree of accuracy [204] the probability  $p_{e_1, e_2}$  of two measured

#### 4. Density-matrix simulation of small surface codes

error signals  $e_1$  and  $e_2$  being connected by a chain of individual logical errors. This is then converted to a weight  $w_{e_1, e_2} = -\log(p_{e_1, e_2})$ , which form the input to the blossom algorithm of Edmonds to find the most likely matching of error signals [173, 209]. An exact calculation of  $p_{e_1, e_2}$  requires a sum over all such chains between  $e_1$  and  $e_2$  that do not cross the boundary (these are equivalent modulo stabilizer operators that do not change the logical state). In this appendix we detail a method of computing this sum, and approximations to make it viable within the runtime of the experiment.

Let us define the ancilla graph  $\mathcal{G}_A = (V_A, E_A)$  containing a vertex  $v \in V_A$  for every ancilla measurement, and an edge  $e \in E_A$  connecting  $v, u \in V_A$  if a single component (gate, single-qubit rest period, or faulty measurement) in the simulation can cause the  $u$  and  $v$  measurements to return an error. We include a special ‘boundary’ vertex  $v_B$ , to which we connect another vertex  $v$  if single components can cause errors on  $v$  alone. Then, to each edge  $e$  we associate a probability  $p_e$ , being the sum of the probabilities of each component causing this error signal. These error rates can be obtained directly from quantumsim, by cutting the circuit at each C-Z gate and measuring the decay of single qubits between. Then, for a given experiment with given syndrome measurements, let us define the syndrome graph  $\mathcal{G}_S = (V_S, E_S)$  containing a vertex  $v \in V_S$  for each syndrome measurement that records an error, and an edge  $\lambda_{u, v} \in E_S$  connecting  $u, v \in V_S$  if  $u$  and  $v$  are either both  $X$  ancilla qubits or both  $Z$  ancilla qubits. To each edge  $\lambda_{u, v}$  we associate a probability  $p_{u, v}$  given by the sum of the probabilities of a chain of errors causing error signals solely on  $u$  and  $v$ .

If we assume that single-qubit errors are uncorrelated, we have to lowest order

$$p_{u, v} \approx \sum_{\text{paths } (e_1, e_2, \dots, e_n) \text{ between } u \text{ and } v} \prod_{j=1}^n p_{e_j}, \quad (4.12)$$

Let  $A_A$  be the adjacency matrix on  $\mathcal{G}_A$  weighted by the probabilities  $p_e$  (i.e.,  $(A_A)_{u, v} = p_e$  with  $e$  connecting  $u$  and  $v$ ), and  $A_S$  the same for  $\mathcal{G}_S$ . Then, the above becomes

$$A_S = A_A + A_A^2 + A_A^3 + \dots = \frac{\mathbf{1}}{\mathbf{1} - A_A} - \mathbf{1}, \quad (4.13)$$

noting that  $A_S$  contains a subset of the indices that are used to construct  $A_A$ .

The boundary must be treated specially in the above calculation. For

#### 4.F. Homemade MWPM decoder with asymmetric weight calculation

the purposes of the surface code, the boundary can be described as a single vertex which has no limit on the number of other vertices it may pair to [173]. For the purposes of weight calculation, any path that passes through the boundary is already counted by pairing both end vertices to the boundary. This can be treated by making  $\mathcal{G}_A$  directed, and breaking the symmetry  $A_A^T = A_A$ . In particular, either  $(A_A)_{v_B, u} = 0$  for all  $u$  or  $(A_A)_{u, v_B} = 0$  for all  $u$ .

The above calculation requires inversion of a  $N_{\text{mat}} \times N_{\text{mat}}$  matrix, with  $N_{\text{mat}}$  the total number of ancilla measurements per experiment. Furthermore, as ancilla error rates depend upon the previous ancilla state, elements in  $A_A$  are not completely known until the previous cycle. This implies that in an actual computation with runtime decoding, this inversion would need to be completed within a few microseconds (with a transmon-cQED architecture), which is practically unfeasible. We suggest two approximations that can be made to shorten the decoding time. The first is to average all errors over the ancilla population, ignoring any asymmetry in the system. The adjacency matrix is now the same for any experiment, and can be precalculated and stored as a look-up table for the run-time decoder. We call this the decoder with symmetrized weights. The size of such a look-up table scales poorly with the number of qubits and the number of cycles. However,  $(A_S)_{u, v}$  is approximately invariant under simultaneous translation of  $u$  and  $v$  (excluding boundary effects). This implies that a precalculated  $A_S$  can be vastly compressed, making this method feasible.

The second approximation to the full  $A_S$  calculation is to perform it iteratively. We divide our graph  $\mathcal{G}_A$  ( $\mathcal{G}_S$ ) by time steps; let  $\mathcal{G}_A^t$  ( $\mathcal{G}_S^t$ ) be the subgraph of  $\mathcal{G}_A$  ( $\mathcal{G}_S$ ) containing only ancillas measured before time step  $t$ , and let  $\partial\mathcal{G}_A^t$  ( $\partial\mathcal{G}_S^t$ ) be the subgraph of  $\mathcal{G}_A$  ( $\mathcal{G}_S$ ) containing only ancillas measured during time step  $t$ . Then, if we assume we have an approximation to the matrix  $A_S^t$  (being the adjacency matrix of  $\mathcal{G}_S^t$ ), we can approximate

$$A_S^{t+1} \approx \begin{pmatrix} A_S^t & C_S^{t+1} \\ (C_S^{t+1})^T & (\mathbf{1} - \partial A_A^{t+1})^{-1} \end{pmatrix} \quad (4.14)$$

to lowest order in physical errors. Here,  $\partial A_A^{t+1}$  is the weighted adjacency matrix on  $\partial\mathcal{G}_A^{t+1}$ , and the coupling matrix  $C_S^{t+1}$  is approximated by

$$C_S^{t+1} = A_S^t C_A^{t+1} (\mathbf{1} - \partial A_A^{t+1})^{-1}, \quad (4.15)$$

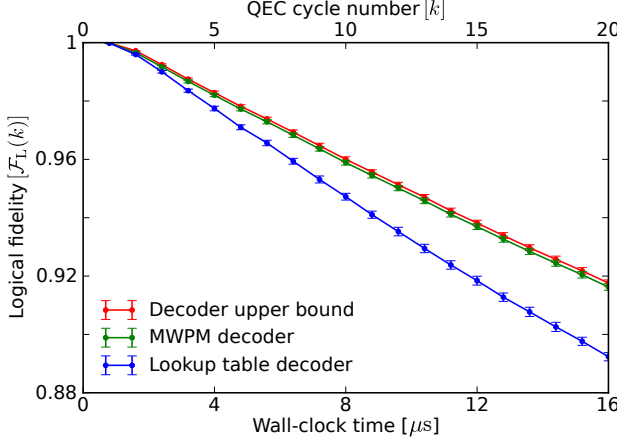
#### 4. Density-matrix simulation of small surface codes

with  $C_A^{t+1}$  the adjacency matrix containing only edges between  $\partial\mathcal{G}_A^{t+1}$  and  $\mathcal{G}_A^t$ . This procedure corresponds to a sum over all paths that are made by moving within  $\partial\mathcal{G}_A^{t+1}$ , shifting back in time to  $\mathcal{G}_A^t$ , and then taking any precalculated path in  $\mathcal{G}_A^t$ .  $C_A^{t+1}$  and  $(\mathbf{1} - \partial A_A^{t+1})^{-1}$  can be precalculated, and so the runtime computation requirement is reduced to the product in Eq. (4.15). This in turn can be sparsified, as  $C_A^{t+1}$  only contains connections to vertices in  $\mathcal{G}_A^t$  close to the time boundary, and we can delete all terms in  $A_S^t$  that do not connect from these vertices to errors.

We have used the second method for our MWPM decoder, as we expect the error from neglecting higher-order combinations of errors to be small. In order to check this assumption, in Fig. 4.12 we repeat our simulation protocol with a modified physical error model that excludes all  $Y$  and measurement errors. We see that in the absence of these errors, the MWPM decoder performs within the error margin of the decoder upper bound. Note that a small deviation is expected from the discrepancy between a MWPM decoder and a maximum-likelihood decoder [178]. With the parameters used in this chapter, we do not observe any loss of fidelity when we stop accounting for the difference in error rates between ancilla states. We account this to the large error contribution from photon noise and gate infidelity on the ancilla qubits, which do not have this asymmetry. We further note that we operate in a regime of large ancilla error; as described in the text this makes the system counter-intuitively less sensitive to ancilla noise. In systems where this is not the case, it could be that accounting for ancilla asymmetry provides a useful computational method to improve  $\epsilon_L$ .

### 4.G. Implementation of a look-up table decoder

In [176], the authors describe a decoding scheme specific to Surface-17, which is optimized to be implementable with limited computational resources in a short cycle time. This decoding scheme works by using a short decision tree to connect errors to each other in a style similar to blossom. Indeed, this scheme is equivalent to a blossom decoder with all horizontal, vertical and diagonal weights equal [176]. As such, we have implemented the new weights in the blossom decoder rather than utilizing the exact method given.



**Figure 4.12.:** Simulation of the experimental protocol used throughout the chapter, but using an error model that has  $Y$  errors and readout infidelity removed. With these errors absent, the MWPM decoder achieves the decoder upper bound within simulation error. The look-up table approach (blue) retains some inaccuracy beyond this.

## 4.H. Details of lowest-order approximation

We detail the approximation made to study Surface-49 in Sec. II C. Note that this calculation is only for  $X$  errors, which are measured by the  $Z$  ancillas. This implies that our approximation should attempt to realize the result of blossom, rather than the decoder upper bound.

We begin with the  $\mathcal{G}_A$  graph defined in App. 4.F. In the absence of correlated errors that cause more than two error signals, any experiment can be approximately described by choosing a set  $S \subset E_A$  of edges on the graph and assuming the errors that correspond to these edges have occurred. Each ancilla measurement corresponds to a vertex in  $\mathcal{G}_A$ , which records an error if an odd number of edges in  $S$  point to the vertex. Each combination  $M_a$  of ancilla measurements can be generated by multiple error sets  $S$ .

Formally, let us write  $\mathcal{M}$  for the set of all combinations of ancilla measurements and  $\mathcal{S}$  for the set of all combinations of errors (so  $\mathcal{S} = 2^{E_A}$ ). We then define a function  $\phi : \mathcal{S} \rightarrow \mathcal{M}$  that takes a combination of errors to the resultant measurement outcomes. Let us fix a logical  $Z$  operator  $Z_L$  on the surface-code fabric. Then to each  $S \in \mathcal{S}$  we can assign a parity

#### 4. Density-matrix simulation of small surface codes

$p(S) = \pm 1$  depending on whether the product of all errors in  $S$  commute with  $Z$  or not. A decoding then consists of a choice of parity  $p_d(M)$  for each  $M \in \mathcal{M}$ . Such a decoding correctly decodes  $S \in \mathcal{S}$  if  $p_d(\phi(S)) = p(S)$ , and creates a logical error otherwise. The source of logical errors in a perfect decoder is then precisely the fact that we can have two error combinations  $S_1, S_2 \in \mathcal{S}$  such that  $\phi(S_1) = \phi(S_2)$  but  $p(S_1) \neq p(S_2)$ .

The above suggests a method by which a perfect decoder can be constructed. As defined,  $\phi^{-1}(M) \subset \mathcal{S}$  is the set of error combinations  $S$  that return a measurement  $M \in \mathcal{M}$ . For each error combination  $S$ , we can calculate the probability of this occurring:

$$r(S) = \prod_{e \in S} p_e \prod_{e \notin S} (1 - p_e). \quad (4.16)$$

The optimal choice of  $p_d(M_a)$  is the one maximizing

$$\sum_{S \in \phi^{-1}(M), p(S)=p_d(M)} r(S), \quad (4.17)$$

and the fidelity of such a decoder (over the entire experiment) can be calculated as

$$\mathcal{F}_L = 1 - \sum_{M \in \mathcal{M}} \min \left( \sum_{S \in \phi^{-1}(M)} \delta_{p(S), +1} r(S), \sum_{S \in \phi^{-1}(M)} \delta_{p(S), -1} r(S) \right). \quad (4.18)$$

At this point the only approximation that has been made is to neglect the  $T_1$  asymmetry in the system, which we have shown previously in this chapter to be negligible. Unfortunately, the above function cannot be evaluated exactly; the number of error combinations  $S$  is approximately  $2^{200}$  for 4 cycles of Surface-49. Our goal instead is to approximate this to the lowest order in the physical qubit error rate.

Let us make the approximation that our error combinations  $S$  can be split into small, well-separated pieces of errors containing separate correctable and non-correctable parts,  $S = \cup_i S^i$ . To each  $S^i$  we can assign a time step  $t(S^i)$ , being the earliest time of the first error measurement observed (in  $\phi(S^i)$ ). The error rate per round,  $\epsilon_L$ , can be determined by summing Eq. 4.18 over all pieces  $S^i$  of all combinations  $S$  such that  $t(S^i) = T$  (with

#### 4.H. Details of lowest-order approximation

arbitrary  $T$ ), as the effect of repeated errors from  $S^i, S^j \subset S$  is taken into account during the derivation of the logical fidelity equation (Eq. 2 in the main text).

$$\epsilon_L = \sum_{M \in \mathcal{M}} \min \left( \sum_{S \in \phi^{-1}(M)} \sum_{t(S^i)=T} \delta_{p(S^i),+1} r(S), \right. \\ \left. \sum_{S \in \phi^{-1}(M)} \sum_{t(S^i)=T} \delta_{p(S^i),-1} r(S) \right). \quad (4.19)$$

Let us also extend the above division of  $S$  to a division of  $M$  into separate pieces  $M^a$ , and rewrite our sum slightly,

$$\epsilon_L = \sum_{M^a} \min \left( \sum_{S^i \in \phi^{-1}(M^a), t(S^i)=T} \delta_{p(S^i),+1} \bar{r}(S^i), \right. \\ \left. \sum_{S^i \in \phi^{-1}(M^a), t(S^i)=T} \delta_{p(S^i),-1} \bar{r}(S^i) \right), \quad (4.20)$$

Where here we have brought the sum over the global combinations of syndromes and measurements inside a new function  $\bar{r}$

$$\bar{r}(S^i) = \prod_{e \in S^i} p_e \sum_{M \supset M^a} \sum_{(S \supset S^i, S \in \phi^{-1}(M))} \\ \prod_{f \in S/S^i} p_f \prod_{g \notin S^i} (1 - p_g) \\ = \prod_{e \in S^i} p_e \sum_{S \supset S^i} \prod_{f \in S/S^i} p_f \prod_{g \notin S^i} (1 - p_g) \quad (4.21)$$

If we took this approximation literally and considered the sum over every possible combination  $S$  containing  $S^i$ , the final sum in Eq. 4.21 would reduce to

$$\bar{r}^{(u)}(S^i) = \prod_{e \in S^i} p_e. \quad (4.22)$$

However, this includes error combinations  $S$  that cannot be easily separated into  $S^i$  and ‘something else’, i.e. they contain other errors  $e$  that cannot be separated from  $S^i$ . Eq. 4.22 is then equivalent to assuming that if  $S^i$  is an uncorrectable logical error, no nearby combination of physical errors

#### 4. Density-matrix simulation of small surface codes

$S'$  can be combined such that  $S^i \cup S'$  is correctable unless  $S'$  itself is an uncorrectable logical error. Such combinations would serve to reduce the calculated  $\epsilon_L$ , and so  $\bar{r}^{(u)}$  gives an upper bound for  $\epsilon_L$  in Eq. 4.20. For a lower bound, we approximate that for any uncorrectable error combination  $S^i$ , approximately one rounds-worth of single errors would undo the logical error, leading to the approximation

$$\bar{r}^{(l)}(S^i) = \prod_{e \in S^i} p_e \prod_{t(\{e\})=T} (1 - p_e). \quad (4.23)$$

We now make one further approximation, and sum Eq. 4.20 only over the shortest  $S^i$  that can be expected to contribute to the final error rate. That is, we sum over those  $S^i$  with  $|S^i| \leq (d+1)/2$ , and that spread directly across the chain. The error incurred from this approximation is roughly proportional to the largest single error, which is no more than 5% throughout our study. We use  $\bar{r}^{(u)}$  and  $\bar{r}^{(l)}$  to give the error bars shown in Fig. 5. Points in the plot are taken as a log average of the upper and lower bounds, and thus have no particular relevance themselves. We see that the numerical calculation falls within the corresponding error bars for almost the entire dataset, giving verification for our method, save a slight deviation at one point where it falls below. Moreover, as the simulated Surface-17 error rate lies above the upper bound found for the Surface-49 error rate (with the standard set of parameters from the main text), our claim that Surface-17 will operate below the fault-tolerant threshold is quite strong.



# 5. Adaptive weight estimator for quantum error correction in a time-dependent environment

## 5.1. Introduction

To execute algorithms on a quantum computer, one must prevent the accumulation of errors by monitoring and correcting them in control hardware. The monitoring is made possible by a nonlocal encoding of the quantum information in a redundant set of qubits, allowing for repeated measurements via auxiliary (ancilla) qubits without collapsing and destroying the quantum superposition of the logical degrees of freedom [66, 205]. Parity-check measurements produce strings of bits, the so-called error syndrome, that must be decoded to infer the correction which should be applied to the logical qubits.

For an important class of error correcting codes, the syndrome identifies the end points of an error chain in a space-time graph of ancilla measurements. (See Fig. 5.1.) The dimensionality of space can differ; it equals 1 in the repetition code [67, 206], 2 in the surface code [167, 173, 207], and 3 for topological cluster states [208]. The identification is not unique: there is in general no unique way to construct a chain of error events consistent with a given syndrome (the decoding problem). One approach to decoding refers to the optimization problem of minimum-weight perfect matching on a graph, which may be solved by the “blossom” algorithm [209, 210] in

---

The contents of this chapter have been published in S. Spitz, B. Tarasinski, C. W. J. Beenakker, and T. E. O’Brien, *Adv. Quant. Tech.* **1** (1), 1870015 (2018)

## 5. Quantum error correction in a time-dependent environment

polynomial time. The blossom decoder is sub-optimal [178, 190, 191, 211], but it performs sufficiently well for current quantum hardware to achieve the fault-tolerance threshold [212].

The weights that govern the optimization problem can be readily obtained if one has a calibrated model of the sources of error in the system [213]. Such an error model may not be available, and moreover the error rates may vary in time during the quantum computation. This complication has motivated the search for an adaptive decoder, that would infer the weights from the syndrome without requiring updates of the error model [189, 214–217]. Since the syndrome depends nonlinearly on the weights, this inversion problem is nontrivial — a recent approach [217] employs a machine learning algorithm to learn the weights from the measured data.

Here we show that the inversion can be actually carried out by purely algebraic means: The covariance of measurements on pairs of ancillas exactly determines the weight of their matching. We demonstrate the method on the repetition code with time-dependent error rates.

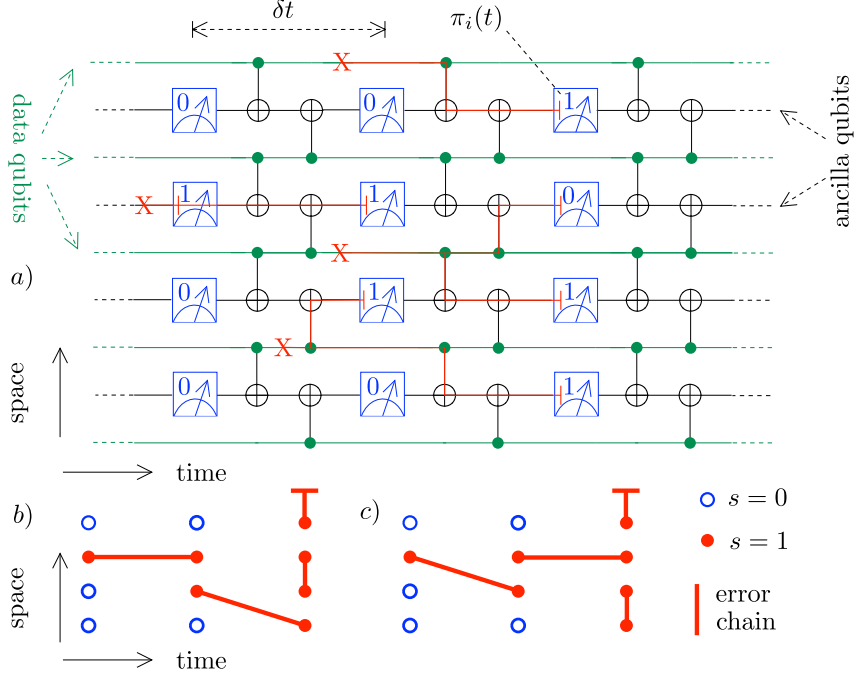
### 5.2. Quantum error correction and the repetition code

To set the stage, we summarize the elements of quantum error correction [66, 218] that we need in what follows. The expert reader may skip this section.

A quantum error correcting code stores quantum information nonlocally in an array of physical qubits, such that it is protected from local errors (bit flips or sign flips). The encoded state  $|\psi\rangle$  evolves for a cycle time  $\delta t$ , after which a set of ‘stabilizer’ measurements is carried out. The stabilizers project  $|\psi\rangle$  onto a state  $|\psi'\rangle$  that may differ from  $|\psi\rangle$  if an error occurred during the cycle. The outcome of the stabilizer measurements, called the syndrome, identifies the error and allows for a correction. It is crucial that the stabilizer measurements do not measure the degrees of freedom of  $|\psi\rangle$  in which the relevant quantum information is stored, otherwise this information will be lost upon projection.

The simplest example of error correction via stabilizer measurement is a one-dimensional array, in which a logical qubit is encoded into  $d$  data qubits via  $|0\rangle_L = |00\cdots 0\rangle$ ,  $|1\rangle_L = |11\cdots 1\rangle$ . In a classical setting, this would correspond to a distance- $d$  repetition code, for which one would compare the value of adjacent bits to identify up to  $(d-1)/2$  bit flips. A quantum parity check achieves this goal without collapsing the

## 5.2. Quantum error correction and the repetition code



**Figure 5.1.:** *Panel a:* Space-time circuit of a  $d = 5$  quantum repetition code. Five data qubits (green) are entangled with four ancilla qubits (black) by means of a CNOT gate ( $\oplus$ ). The ancillas are measured at the end of each cycle (blue boxes, spaced by  $\delta t$ , with measurement outcomes  $\pi_i(t)$ ,  $i = 1, 2, 3, 4$ ). A bit flip (red X) produces an error chain (red line) with end points on an ancilla measurement or on the boundary of the array. *Panel b:* Syndrome  $s_i(t) = \pi_i(t) \oplus \pi_i(t - 2\delta t)$  corresponding to the error events in panel a). The measurements that are connected by an error chain can be separated in space, in time, or in both space and time. *Panel c:* Alternative matching consistent with the same syndrome. The minimum weight decoder associates a weight to each error chain and finds the matching with the smallest total weight.

## 5. Quantum error correction in a time-dependent environment

superposition  $|\psi\rangle = a_0|0\rangle_L + a_1|1\rangle_L$  onto the state  $|0\rangle_L$  or  $|1\rangle_L$ . The parity-check measurements are performed on  $d - 1$  ancilla qubits, which are entangled with pairs of data qubits (see Fig. 5.1). Each ancilla measures the stabilizer operator  $Z_i Z_{i+1}$  ( $i = 1, 2, \dots, d - 1$ , with  $Z \equiv \sigma_z$  a Pauli matrix). The stabilizer does not distinguish between the states  $|0\rangle_L$  and  $|1\rangle_L$  and thus preserves their quantum superposition. A bit flip error of qubit  $j$  ( $X_j$  error) is detected by the stabilizer measurements  $Z_j Z_{j+1}$  and  $Z_{j-1} Z_j$ , which change their value from  $+1$  to  $-1$ . A decoder may infer the underlying error from this signature and correct it without needing to measure qubit  $j$  itself (which would collapse the state).

The stabilizer measurements in cycle  $t$  form a binary parity-check vector  $\vec{\pi}(t)$ . The error syndrome

$$\vec{s}(t) = \vec{\pi}(t) + \vec{\pi}(t - 2\delta t) \quad (5.1)$$

is defined such that an error event is signaled by a nonzero element  $s_i(t) = 1$ . A single error event is not sufficient to diagnose an error, as  $Z_i Z_{i+1}$  would trigger an error event for either  $X_i$  or  $X_{i+1}$ . To identify which qubit flipped we match pairs of error events. As indicated in Fig. 5.1, the match can be between error events that are separated in both space and time, at the end points of an error chain from ancilla  $i_0$  at time  $t$  to ancilla  $j_0$  at time  $t + n\delta t$ . The error chain may terminate at the boundary of the lattice (corresponding to errors on the boundary data qubits), so some error events may remain unmatched.

This simple description to detect bit flips in a repetition code can be extended to the detection of both bit flips and phase flips ( $X_i$  and  $Z_i$  errors) and by encoding in 2D and 3D (surface codes and topological cluster states). The generic feature of this class of stabilizer codes is that the decoding entails the pairwise matching of error events in a space-time graph. The method of adaptive quantum error correction presented in the next section applies to this general setting, while for a demonstration we will return to the repetition code.

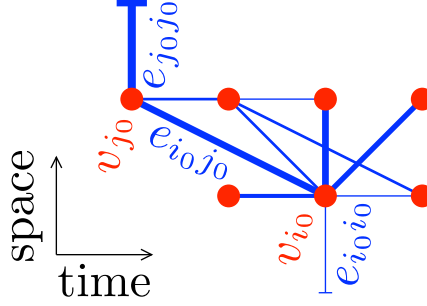
## 5.3. Weight inference from error syndromes

### 5.3.1. Formulation of the inversion problem

We collect the binary output of the stabilizer measurements in the error syndrome  $\vec{s}(t)$ . The discrete time variable  $t$  counts the error correction cycle and the elements of the vector  $\vec{s}$  identify the ancilla qubits. For  $N$  ancillas and  $T$  cycles there are a total of  $NT$  variables  $v_i \in \{0, 1\}$ ,

### 5.3. Weight inference from error syndromes

arranged as vertices in a space-time graph. (See Fig. 5.2.) An error event corresponds to  $v_i = 1$ , while  $v_i = 0$  if the ancilla has not detected an error.



**Figure 5.2.:** Space-time graph showing a pair of vertices  $v_{i_0}$ ,  $v_{j_0}$  connected by an edge  $e_{i_0 j_0}$ . A few other vertices and connecting edges are also shown, as well as edges that connect a vertex to a boundary ( $e_{i_0 i_0}$  and  $e_{j_0 j_0}$ ). The line thickness of an edge  $e_{ij}$  is proportional to the probability  $p_{ij}$  that a single error affects the ancilla qubit measurements on vertices  $i$  and  $j$ . We seek to determine these probabilities from measurements of the error syndrome.

The vertices are pairwise connected by undirected edges  $e_{ij} \equiv e_{ji} \in \{0, 1\}$  such that  $e_{ij} = 1$  with probability  $p_{ij}$ . We allow for  $i = j$ , the edge  $e_{ii}$  connects a vertex to the boundary of the graph. We say that the edge is *on* or *off* depending on whether  $e_{ij} = 1$  or 0. The state of a vertex depends on the edges according to

$$v_i = \frac{1}{2} [1 - (-1)^{\sum_j e_{ij}}]. \quad (5.2)$$

Each edge that is *on* toggles its vertex between the states 0 and 1, so that  $v_i = 1$  if an odd number of connecting edges is *on*.

The  $p_{ij}$ 's are probabilities of a single-qubit error that correlates ancilla measurements  $i$  and  $j$ . Correlations of ancilla measurements due to uncorrelated multiple-qubit errors are described by weights  $w_{ij}$ . The weight  $w_{ij}$  for  $i \neq j$  is determined from the  $p$ 's by following all paths through the graph from  $i$  to  $j$ :

$$w_{ij} = -\ln \left( p_{ij} + \sum_n \sum'_{k_1, k_2, \dots, k_n} p_{ik_1} p_{k_1 k_2} \cdots p_{k_n j} \right). \quad (5.3)$$

The prime in the sum indicates that the path should not pass through the boundary ( $i \neq k_1 \neq k_2 \cdots \neq k_n \neq j$ ). For a boundary weight  $w_{ii}$  the path

## 5. Quantum error correction in a time-dependent environment

terminates on the boundary,

$$w_{ii} = -\ln \left( p_{ii} + \sum_n \sum'_{k_1, k_2, \dots, k_n} p_{ik_1} p_{k_1 k_2} \cdots p_{k_n k_n} \right). \quad (5.4)$$

These sums over error chains can be carried out in terms of the matrix  $A_{ij} = (1 - \delta_{ij})p_{ij}$  by matrix inversion [212],

$$e^{-w_{ij}} = \begin{cases} [(1 - A)^{-1}]_{ij} & \text{if } i \neq j, \\ \sum_k [(1 - A)^{-1}]_{ik} p_{kk} & \text{if } i = j. \end{cases} \quad (5.5)$$

Given a set of error events  $\mathcal{V} = \{v_i | v_i = 1\}$  the minimum-weight perfect matching decoder searches for a subset  $\mathcal{M} = \{e_{ij}, w_{ij}\}$  of weighted edges such that each vertex in  $\mathcal{V}$  is connected either to one other vertex in  $\mathcal{V}$  or to the boundary, at minimal total weight  $\sum_{\mathcal{M}} w_{ij}$ .

Modern implementations [210] of the blossom algorithm [209] solve this optimization problem efficiently given an error model: A physical model for qubit errors from which the error probabilities  $p$  and hence the weights  $w$  can be calculated. Here we consider the opposite approach: can we infer the weights from the measured error syndromes, without an underlying error model? This is an inversion problem for Eq. (5.2), where we seek to reconstruct the statistics of the edges  $e_{ij}$  from the measured statistics of the vertices  $v_i$ . The inversion is possible, in spite of the nonlinearity of Eq. (5.2), as we now show.

### 5.3.2. Solution for edges connecting pairs of vertices

We first consider a pair of distinct vertices  $i_0 \neq j_0$ , connected by an edge  $e_{i_0 j_0}$ . (The case of a single vertex connected to the boundary will be dealt with later.) We define

$$E_{i_0 \setminus j_0} = \frac{1}{2} [1 + (-1)^{\sum_{j \neq j_0} e_{i_0 j}}]. \quad (5.6)$$

In words,  $E_{i_0 \setminus j_0}$  equals 1 or 0 depending on whether the vertex  $i_0$  has an even or an odd number of connecting edges that are *on* — excluding the connection to vertex  $j_0$ . Note that the sum over  $j$  includes  $j = i_0$ , it only excludes  $j = j_0$ .

We then rewrite Eq. (5.2) for vertex  $i_0$  as

$$v_{i_0} = e_{i_0 j_0} E_{i_0 \setminus j_0} + (1 - e_{i_0 j_0})(1 - E_{i_0 \setminus j_0}). \quad (5.7)$$

### 5.3. Weight inference from error syndromes

Similarly, for vertex  $j_0$  one has

$$v_{j_0} = e_{j_0 i_0} E_{j_0 \setminus i_0} + (1 - e_{j_0 i_0})(1 - E_{j_0 \setminus i_0}). \quad (5.8)$$

Since  $e_{i_0 j_0} = e_{j_0 i_0} = e_{i_0 j_0}^2$ , the product (AND) of  $v_{i_0}$  and  $v_{j_0}$  equals

$$\begin{aligned} v_{i_0} v_{j_0} &= (1 - e_{i_0 j_0})(1 - E_{i_0 \setminus j_0} - E_{j_0 \setminus i_0}) \\ &\quad + E_{i_0 \setminus j_0} E_{j_0 \setminus i_0}, \end{aligned} \quad (5.9)$$

while the binary sum (XOR) equals

$$\begin{aligned} v_{i_0} \oplus v_{j_0} &\equiv v_{i_0} + v_{j_0} \pmod{2} \\ &= E_{i_0 \setminus j_0} + E_{j_0 \setminus i_0} - 2E_{i_0 \setminus j_0} E_{j_0 \setminus i_0}. \end{aligned} \quad (5.10)$$

By construction, all three variables  $e_{i_0 j_0}$ ,  $E_{i_0 \setminus j_0}$ , and  $E_{j_0 \setminus i_0}$  are statistically independent. We denote the average by  $\langle \dots \rangle$ , with

$$\langle e_{i_0 j_0} \rangle = p_{i_0 j_0} \quad (5.11)$$

by definition. The averages of the  $E$ 's are unknown, but they can be eliminated by combining the four equations (5.7)–(5.10). We thus arrive at

$$p_{i_0 j_0} (1 - p_{i_0 j_0}) = \frac{\langle v_{i_0} v_{j_0} \rangle - \langle v_{i_0} \rangle \langle v_{j_0} \rangle}{1 - 2\langle v_{i_0} \oplus v_{j_0} \rangle}. \quad (5.12)$$

The left-hand-side is symmetric under the exchange  $p_{i_0 j_0} \leftrightarrow 1 - p_{i_0 j_0}$ . We may safely assume that the error probabilities are  $< 1/2$ , resulting in the probability

$$p_{i_0 j_0} = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{\langle v_{i_0} v_{j_0} \rangle - \langle v_{i_0} \rangle \langle v_{j_0} \rangle}{1 - 2\langle v_{i_0} \oplus v_{j_0} \rangle}}. \quad (5.13)$$

This is an exact relation between the probability of an edge and correlators of the pair of connected vertices. These correlators are measurable from the error syndrome, without any prior knowledge of the error model.

#### 5.3.3. Solution for boundary edges

The probability  $p_{i_0 i_0}$  of an edge  $e_{i_0 i_0}$  connecting the vertex  $v_{i_0}$  to the boundary cannot be determined by a correlator, since there is nothing to

## 5. Quantum error correction in a time-dependent environment

correlate with. We do have access to the average

$$\langle v_{i_0} \rangle = 1 - p_{i_0 i_0} - (1 - 2p_{i_0 i_0}) \langle E_{i_0 \setminus i_0} \rangle, \quad (5.14)$$

$$\begin{aligned} E_{i_0 \setminus i_0} &= \frac{1}{2} [1 + (-1)^{\sum_{j \neq i_0} e_{i_0 j}}] \\ &= \frac{1}{2} + \frac{1}{2} \prod_{j \neq i_0} (1 - 2e_{i_0 j}). \end{aligned} \quad (5.15)$$

Using again the independence of the variables, we find

$$p_{i_0 i_0} = \frac{1}{2} + \frac{\langle v_{i_0} \rangle - 1/2}{\prod_{j \neq i_0} (1 - 2p_{i_0 j})}. \quad (5.16)$$

So once the probabilities  $p_{i_0 j}$  for non-boundary edges are determined from Eq. (5.13), we can use Eq. (5.16) to obtain the probability of a boundary edge.

## 5.4. Implementation of the adaptive decoder

### 5.4.1. Convergence in the large-time limit

We test the adaptive decoder on the repetition code of Fig. 5.1, for a bit-flip error model: at the end of a cycle of duration  $\delta t$  each qubit  $i$  is flipped independently with probability  $\gamma_i$ . The time-dependent density matrix of the quantum circuit is calculated using the `quantumsim` simulator of Ref. 212.

We implement the blossom decoder without any prior knowledge of the error probabilities, using Eqs. (5.13) and (5.16) to determine them from the measured syndrome data. We assume local sources of error and set  $p_{ij} \equiv 0$  for ancilla measurements  $i$  and  $j$  that are not connected by any local error. In a nonlocal situation, *e.g.* because of non-negligible crosstalk, a proliferation of negligibly small error probabilities can be avoided by setting  $p_{ij} \equiv 0$  when the deviation from zero is statistically insignificant.

The adaptive decoder needs sufficient syndrome data in the training stage to estimate the probabilities. Since  $p_{ij}$  is the mean of a Bernoulli random variable with variance  $\sigma_{ij}^2 = p_{ij}(1 - p_{ij})$ , the statistical uncertainty  $\delta p_{ij}$  in the estimation after  $N = t/\delta t$  error cycles is of order

$$\delta p_{ij} = N^{-1/2} \sqrt{p_{ij}(1 - p_{ij})}. \quad (5.17)$$



## 5.4. Implementation of the adaptive decoder

The requirement that  $\delta p_{ij} \ll p_{ij} \ll 1$  implies that a minimum of

$$N_{\min} \simeq 1/\bar{p} \quad (5.18)$$

measurements are needed for a reliable estimation of error probabilities of typical magnitude  $\bar{p}$ . After the training stage the probabilities are inserted into Eq. (5.5) to determine the weights which are passed to the minimum-weight perfect matching (blossom) decoder for error correction.

As a figure of merit we introduce a testing stage after the training stage in which we calculate the probability  $\epsilon_{\text{adaptive}}(N)$  of a logical error per cycle using the adaptive decoder trained on  $N$  rounds of data. The error rates are calculated following the method of Ref. 212, measuring the average logical qubit fidelity over 100 cycles. The combination of training and testing is repeated a few hundred times to obtain an accurate value of  $\epsilon_{\text{adaptive}}(N)$ . We compare this with the probability  $\epsilon_0$  that would follow from a blossom decoder with pre-determined weights calculated from the error model. The relative error

$$\Delta = \epsilon_{\text{adaptive}}/\epsilon_0 - 1 \quad (5.19)$$

measures how well the adaptive decoder has converged to the ideal blossom decoder.

Results are shown in Fig. 5.3, for a depth  $d = 3$  repetition code with uniform single-qubit error rate  $\gamma_i = 5 \cdot 10^{-3}$ . We observe a power law convergence  $\Delta \propto N^{-\alpha}$  with  $\alpha \approx 1.2$ . (We do not have an analytical result for this exponent.)

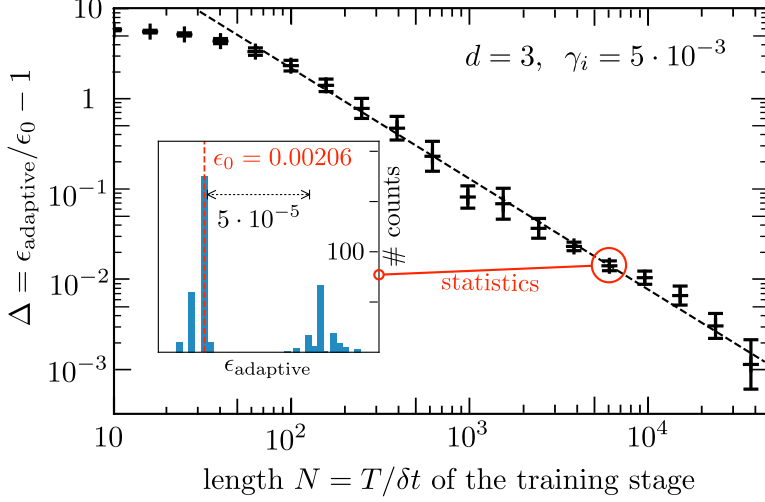
### 5.4.2. Performance in a time-dependent environment

The adaptive decoder can be readily applied to sources of noise that vary in time, by recalibration of the weights as time proceeds. We implement this by estimating the error probabilities at time  $t$  from the syndrome data in the time interval  $(t - T, t)$ . The optimal time window  $T = N\delta t$  should not be too short in view of the statistical error (5.17), and it should not be too large in view of the variation  $\omega T p_{ij}$  of the probabilities in the time-dependent environment (with characteristic frequency  $\omega$ ). The sum of these sources of error is minimized for

$$N_{\text{opt}} \simeq (p_{ij} \omega^2 \delta t^2)^{-1/3} \Rightarrow \delta p_{ij}^{\text{opt}} \simeq p_{ij}^{2/3} (\omega \delta t)^{1/3}. \quad (5.20)$$

The adaptive decoder fails if the noise fluctuates too rapidly to acquire sufficient data for the probability estimation. The condition  $\delta p_{ij}^{\text{opt}} \ll p_{ij}$

## 5. Quantum error correction in a time-dependent environment



**Figure 5.3.:** Convergence of the adaptive decoder towards the ideal blossom decoder, as determined by the relative decoder error  $\Delta$  as a function of the number of cycles  $N$  used to estimate the error probabilities in the training stage. Each data point with error bars results from the repetition of 400 training stages, the inset shows the statistics for one particular data point. The dashed line through the data points is a guide to the eye.

implies an upper bound

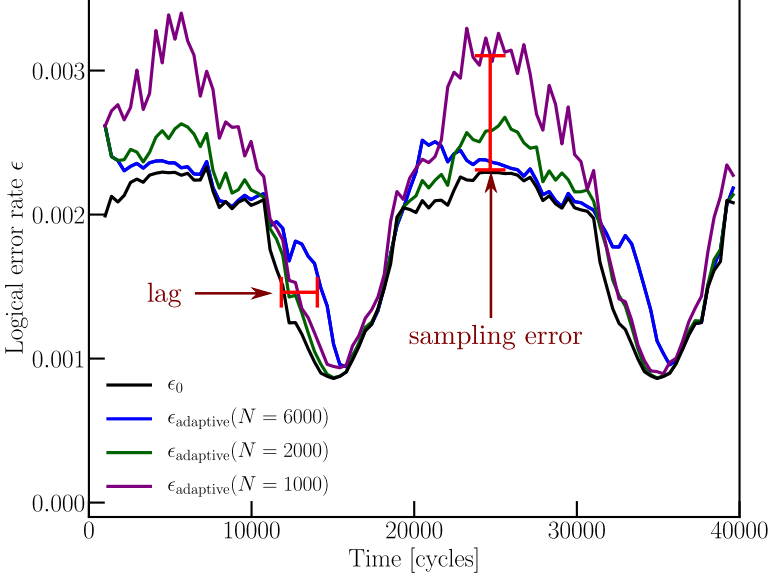
$$\omega_c \simeq \bar{p}/\delta t \quad (5.21)$$

on the frequency of the noise variations that is adaptable for a typical error probability  $\bar{p}$ .

We test the adaptive decoder in the presence of time dependent errors by taking  $\gamma_i = \gamma_0$  for the data qubits and  $\gamma_i = \gamma_0(1 + \sin \omega t)$  for the ancilla qubits (with  $\gamma_0 = 5 \cdot 10^{-3}$  and  $2\pi/\omega = 2 \cdot 10^4 \delta t$ ). The predicted optimal time window at this frequency, for  $\bar{p} = 5 \cdot 10^{-3}$ , is  $N_{\text{opt}} \approx 1265$ . As shown in Fig. 5.4, when a larger window  $N \gg N_{\text{opt}}$  is used, the decoder experiences a time lag in determining optimal weights; for a smaller window  $N < N_{\text{opt}}$  the weight estimation is degraded by sampling errors.

## 5.5. Conclusion

We have demonstrated that it is possible to analytically calculate the underlying error probabilities from measured error syndromes in a broad



**Figure 5.4.:** Performance of the adaptive decoder in the presence of a fluctuating noise ( $d = 3$ ,  $\gamma_i = 0.005$  for data qubits,  $\gamma_i = 0.005 + 0.005 \sin(\pi t/10^4 \delta t)$  for ancilla qubits) using three different time windows  $T = N\delta t$  for the error estimation. The average over 200 training stages is compared to a blossom decoder (black) with optimally chosen weights at every point in time. Small time windows suffer from sampling error, but adapt quickly to changing error rates, while a decoder with a larger time window lags behind. The optimal time window that balances the two effects is around  $T = 2000 \delta t$  in this case.

class of stabilizer codes. As this requires inverting a set of non-linear equations, it is surprising that it should be possible at all, let alone with such small overhead. Because the inversion is exact, the convergence of our adaptive decoder to the ideal blossom decoder should be optimal in the absence of additional information about the error rates. This implies that fluctuations faster than a critical frequency  $\omega_c$  are uncorrectable; we have estimated  $\omega_c \simeq \bar{p}/\delta t$ , with  $\bar{p}$  the single-qubit error probability and  $\delta t$  the duration of one error-correction cycle. Such rapid fluctuations will contribute relatively more to the logical error rate of a quantum error correcting code than slow fluctuations to which the decoder can adapt.

It would be interesting for future work to test the adaptive decoder on more complex noise models, where the optimal window must be chosen for an entire noise frequency spectrum, instead of for a single frequency. We

## 5. *Quantum error correction in a time-dependent environment*

expect white noise to be significantly worse for quantum error correction than  $1/f$  noise, due to the much larger contributions from high frequencies. Future work could also extend our results to simulations of the surface code or topological cluster states.

## 6. Neural network decoder for topological color codes with circuit level noise

### 6.1. Introduction

In fault-tolerant quantum information processing, a topological code stores the logical qubits nonlocally on a lattice of physical qubits, thereby protecting the data from local sources of noise [66, 205]. To ensure that this protection is not spoiled by logical gate operations, they should act locally. A gate where the  $j$ -th qubit in a code block interacts only with the  $j$ -th qubit of another block is called “transversal” [219]. Transversal gates are desirable both because they do not propagate errors within a code block, and because they can be implemented efficiently by parallel operations.

Two families of two-dimensional (2D) topological codes have been extensively investigated, surface codes [79, 173, 207, 220] and color codes [75, 222]. The two families are related: a color code is equivalent to multiple surface codes, entangled using a local unitary operation [223, 224] that amounts to a code concatenation [225]. There are significant differences between these two code families in terms of their practical implementation. On the one hand, the surface code has a favorably high threshold error rate for fault tolerance, but only CNOT,  $X$ , and  $Z$  gates can be performed transversally [221]. On the other hand, while the color code has a smaller threshold error rate than the surface code [174, 226], it allows for the transversal implementation of the full Clifford group of quantum gates (with Hadamard,  $\pi/4$  phase gate, and CNOT gate as generators) [227, 228]. While this is not yet computationally universal, it can be rendered universal using gate teleportation [229] and magic state distillation [48]. Moreover, color codes are particularly suitable for topological quantum computation

---

The contents of this chapter have been published in P. Baireuther, M.D. Caio, B. Criger, C.W.J. Beenakker and T.E. O’Brien, New J. Phys. 21, 013003 (2019)

with Majorana qubits, since high-fidelity Clifford gates are accessible by braiding [230, 231].

A drawback of color codes is that quantum error correction is more complicated than for surface codes. The identification of errors in a surface code (the “decoding” problem) can be mapped onto a matching problem in a graph [232], for which there exists an efficient solution called the “blossom” algorithm [209]. This graph-theoretic approach does not carry over to color codes, motivating the search for decoders with performance comparable to the blossom decoder, some of which use alternate graph-theoretic constructions [233–237].

An additional complication of color codes is that the parity checks are prone to “hook” errors, where single-qubit errors on the ancilla qubits propagate to higher weight errors on data qubits, reducing the effective distance of the code. There exist methods due to Shor [238], Steane [239], and Knill [240] to mitigate this, but these error correction methods come with much overhead because of the need for additional circuitry. An alternative scheme with reduced overhead uses dedicated ancillas (“flag qubits”) to signal the hook errors [241–245].

Here we show that a neural network can be trained to fault-tolerantly decode a color code with high efficiency, using only measurable data as input. No *a priori* knowledge of the error model is required. Machine learning approaches have been previously shown to be successful for the family of surface codes [211, 246–249], and applications to color codes are now being investigated [250, 251, 253]. We adapt the recurrent neural network of Ref. 211 to decode color codes with distances up to 7, fully incorporating the information from flag qubits. A test on a density matrix-based simulator of a superconducting quantum computer [212] shows that the performance of the decoder is close to optimal, and would surpass the quantum memory threshold under realistic experimental conditions.

## 6.2. Description of the problem

### 6.2.1. Color code

The color code belongs to the class of stabilizer codes [68], which operate by the following general scheme. We denote by  $I, X, Y, Z$  the Pauli matrices on a single qubit and by  $\Pi^n = \{I, X, Y, Z\}^{\otimes n}$  the Pauli group on  $n$  qubits. A set of  $k$  logical qubits is encoded as a  $2^k$ -dimensional Hilbert space  $\mathcal{H}_L$  across  $n$  noisy physical qubits (with  $2^n$ -dimensional Hilbert space  $\mathcal{H}_P$ ). The logical Hilbert space is stabilized by the repeated measurement of

$n - k$  parity checks  $S_i \in \Pi^n$  that generate the stabilizer  $\mathcal{S}(\mathcal{H}_L)$ , defined as

$$\mathcal{S}(\mathcal{H}_L) = \{S \in \mathcal{B}(\mathcal{H}_P), S|\psi_L\rangle = |\psi_L\rangle \forall |\psi_L\rangle \in \mathcal{H}_L\}, \quad (6.1)$$

where  $\mathcal{B}(\mathcal{H}_P)$  is the algebra of bounded operators on the physical Hilbert space.

As errors accumulate in the physical hardware, an initial state  $|\psi_L(t=0)\rangle$  may rotate out of  $\mathcal{H}_L$ . Measurement of the stabilizers discretizes this rotation, either projecting  $|\psi_L(t)\rangle$  back into  $\mathcal{H}_L$ , or into an error-detected subspace  $\mathcal{H}_{\vec{s}(t)}$ . The syndrome  $\vec{s}(t) \in \mathbb{Z}_2^{n-k}$  is determined by the measurement of the parity checks:  $S_i \mathcal{H}_{\vec{s}(t)} = (-1)^{s_i(t)} \mathcal{H}_{\vec{s}(t)}$ . It is the job of a classical decoder to interpret the multiple syndrome cycles and determine a correction that maps  $\mathcal{H}_{\vec{s}(t)} \mapsto \mathcal{H}_L$ ; such decoding is successful when the combined action of error accumulation and correction leaves the system unperturbed.

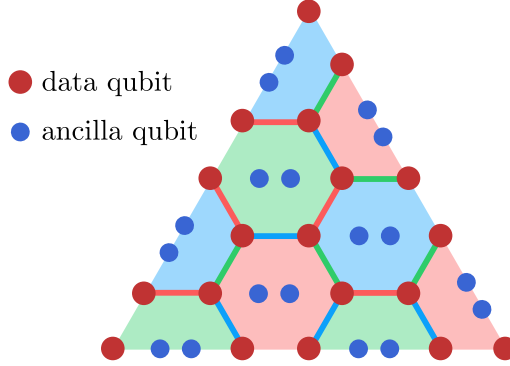
This job can be split into a computationally easy task of determining a unitary that maps  $\mathcal{H}_{\vec{s}(t)} \mapsto \mathcal{H}_L$  (a so-called ‘pure error’ [252]), and a computationally difficult task of determining a logical operation within  $\mathcal{H}_L$  to undo any unwanted logical errors. The former task (known as ‘excitation removal’ [253]) can be performed by a ‘simple decoder’ [247]. The latter task is reduced, within the stabilizer formalism, to determining at most two parity bits per logical qubit, which is equivalent to determining the logical parity of the qubit upon measurement at time  $t$  [211].

We implement the color code [75, 222] on an hexagonal lattice inside a triangle, see Fig. 6.1. (This is the 6,6,6 color code of Ref. 174.) One logical qubit is encoded by mapping vertices  $v$  to data qubits  $q_v$ , and tiles  $T$  to the stabilizers  $X_T = \prod_{v \in T} X_v$ ,  $Z_T = \prod_{v \in T} Z_v$ . The simultaneous  $+1$  eigenstate of all the stabilizers (the “code space”) is twofold degenerate [228], so it can be used to define a logical qubit. The number of data qubits that encodes one logical qubit is  $n_{\text{data}} = 7, 19$ , or  $37$  for a code with distance  $d = 3, 5$ , or  $7$ , respectively. (For any odd integer  $d$ , a distance- $d$  code can correct  $(d-1)/2$  errors.) Note that  $n_{\text{data}}$  is less than  $d^2$ , being the number of data qubits used in a surface code with the same  $d$  [79].

An  $X$  error on a data qubit switches the parity of the surrounding  $Z_T$  stabilizers, and similarly a  $Z$  error switches the parity of the surrounding  $X_T$  stabilizers. These parity switches are collected in the binary vector of syndrome increments  $\delta \vec{s}(t)^*$ , such that  $\delta s_i = 1$  signals some errors on the qubits surrounding ancilla  $i$ . The syndrome increments themselves

---

\*The syndrome increment is usually  $\delta \vec{s}(t) \equiv \vec{s}(t) - \vec{s}(t-1) \pmod{2}$ . When ancilla qubits are not reset between QEC cycles, we use a somewhat different definition, see App. 6.A.2 for details.



**Figure 6.1.:** Schematic layout of the distance-5 triangular color code. A hexagonal lattice inside an equilateral triangle encodes one logical qubit in 19 data qubits (one at each vertex). The code is stabilized by 6-fold  $X$  and  $Z$  parity checks on the corners of each hexagon in the interior of the triangle, and 4-fold parity checks on the boundary. For the parity checks, the data qubits are entangled with a pair of ancilla qubits inside each tile, resulting in a total of  $\frac{3d^2-1}{2}$  qubits used to realize a distance- $d$  code. Pauli operators on the logical qubit can be performed along any side of the triangle, single-qubit Clifford operations can be applied transversally, and two-qubit joint Pauli measurements can be performed through lattice surgery to logical qubits on adjacent triangles.

are sufficient for a classical decoder to infer the errors on the physical data qubits. Parity checks are performed by entangling ancilla qubits at the center of each tile with the data qubits around the border, and then measuring the ancilla qubits (see App. 6.A for the quantum circuit).

### 6.2.2. Error model

We consider two types of circuit-level noise models, both of which incorporate flag qubits to signal hook errors. Firstly, a simple Pauli error model allows us to develop and test the codes up to distance  $d = 7$ . (For larger  $d$  the training of the neural network becomes computationally too expensive.) Secondly, the  $d = 3$  code is applied to a realistic density-matrix error model derived for superconducting qubits.

In the Pauli error model, one error correction cycle of duration  $t_{\text{cycle}} = N_0 t_{\text{step}}$  consists of a sequence of  $N_0 = 20$  steps of duration  $t_{\text{step}}$ , in which a particular qubit is left idle, measured, or acted upon with a single-qubit rotation gate or a two-qubit conditional-phase gate. Before the first cycle we prepare all the qubits in an initial state, and we reset the ancilla qubits



after each measurement. Similarly to Ref. 173, we allow for an error to appear at each step of the circuit and during the preparation, including the reset of the ancilla qubits, with probability  $p_{\text{error}}$ . For the preparation errors, idle errors, or rotation errors we introduce the possibility of an  $X$ ,  $Y$ , or  $Z$  error with probability  $p_{\text{error}}/3$ . Upon measurement, we record the wrong result with probability  $p_{\text{error}}$ . Finally, after the conditional-phase gate we apply with probability  $p_{\text{error}}/15$  one of the following two-qubit errors:  $I \otimes P$ ,  $P \otimes I$ ,  $P \otimes Q$ , with  $P, Q \in \{X, Y, Z\}$ . We assume that  $p_{\text{error}} \ll 1$  and that all errors are independent, so that we can identify  $p_{\text{error}} \equiv \epsilon_{\text{phys}}$  with the physical error rate per step.

The density matrix simulation uses the *quantumsim* simulator of Ref. 212. We adopt the experimental parameters from that work, which match the state-of-the-art performance of superconducting transmon qubits. In the density-matrix error model the qubits are not reset between cycles of error correction. Because of this, parity checks are determined by the difference between subsequent cycles of ancilla measurement. This error model cannot be parametrized by a single error rate, and instead we compare to the decay rate of a resting, unencoded superconducting qubit.

### 6.2.3. Fault-tolerance

The objective of quantum error correction is to arrive at a error rate  $\epsilon_L$  of the encoded logical qubit that is much smaller than the error rate  $\epsilon_{\text{phys}}$  of the constituting physical qubits. If error propagation through the syndrome measurement circuit is limited, and a “good” decoder is used, the logical error rate should exhibit the power law scaling [173]

$$\epsilon_L = C_d \epsilon_{\text{phys}}^{(d+1)/2}, \quad (6.2)$$

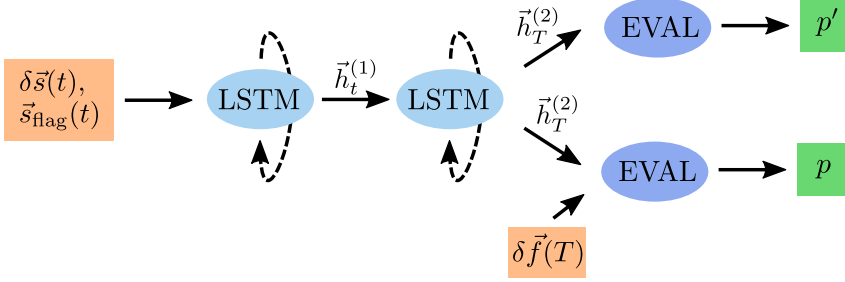
with  $C_d$  a prefactor that depends on the distance  $d$  of the code but not on the physical error rate. The so-called “pseudothreshold”\* [254],

$$\epsilon_{\text{pseudo}} = \frac{1}{C_d^{2/(d-1)}} \quad (6.3)$$

is the physical error rate below which the logical qubit can store information for a longer time than a single physical qubit.

---

\*The quantity  $\epsilon_{\text{pseudo}}$  defined in Eq. (6.3) is called a *pseudo*-threshold because it is  $d$ -dependent. In the limit  $d \rightarrow \infty$  it converges to the true threshold.



**Figure 6.2.:** Architecture of the recurrent neural network decoder. After a body of recurrent layers the network branches into two heads, each of which estimates the probability  $p$  or  $p'$  that the parity of bit flips at time  $T$  is odd. The upper head does this solely based on syndrome increments  $\delta \vec{s}$  and flag measurements  $\vec{s}_{\text{flag}}$  from the ancilla qubits, while the lower head additionally gets the syndrome increment  $\delta \vec{f}$  from the final measurement of the data qubits. During training both heads are active, during validation and testing only the lower head is used. Ovals denote the two long short-term memory (LSTM) layers and the fully connected evaluation layers, while boxes denote input and output data. Solid arrows indicate data flow in the system (with  $\vec{h}_t^{(1)}$  and  $\vec{h}_T^{(2)}$  the output of the first and second LSTM layer), and dashed arrows indicate the internal memory flow of the LSTM layers.

#### 6.2.4. Flag qubits

During the measurement of a weight- $w$  parity check with a single ancilla qubit, an error on the ancilla qubit may propagate to as many as  $w/2$  errors on data qubits. This reduces the effective distance of the code in Eq. (6.2). The surface code can be made resilient to such hook errors, but the color code cannot: Hook errors reduce the effective distance of the color code by a factor of two.

To avoid this degradation of the code distance, we take a similar approach to Refs. 241–245 by adding a small number of additional ancilla qubits, so-called “flag qubits”, to detect hook errors. For our chosen color code with weight-6 parity checks, we opt to use one flag qubit for each ancilla qubit used to make a stabilizer measurement. (This is a much reduced overhead in comparison to alternative approaches [238–240].) Flag and ancilla qubits are entangled during measurement and read out simultaneously (circuits given in App. 6.A). Our scheme is not *a priori* fault-tolerant, as previous work has required at least  $(d-1)/2$  flag qubits per stabilizer. Instead, we rely on fitting our numeric results to Eq. (6.2) with  $d$  fixed to the code distance to demonstrate that our scheme is in fact fault tolerant.

## 6.3. Neural network decoder

### 6.3.1. Learning mechanism

Artificial neural networks are function approximators. They span a function space that is parametrized by variables called weights and biases. The task of learning corresponds to finding a function in this function space that is close to the unknown function represented by the training data. To do this, one first defines a measure for the distance between functions and then uses an optimization algorithm to search the function space for a local minimum with respect to this measure. Finding the global minimum is in general not guaranteed, but empirically it turns out that often local minima are good approximations. For a comprehensive review see for example Refs. 255, 256.

We use a specific class of neural networks known as *recurrent* neural networks, where the “function” can represent an algorithm [257]. During optimization the weights and biases are adjusted such that the resulting algorithm is close to the algorithm represented by the training data.

### 6.3.2. Decoding algorithm

Consider a logical qubit, prepared in an arbitrary logical state  $|\psi_L\rangle$ , kept for a certain time  $T$ , and then measured with outcome  $m \in \{-1, 1\}$  in the logical  $Z$ -basis. Upon measurement, phase information is lost. Hence, the only information needed in addition to  $m$  is the parity of bit flips in the measurement basis. (A separate decoder is invoked for each measurement basis.) If the bit flip parity is odd, we correct the error by negating  $m \mapsto -m$ . The task of decoding amounts to the estimation of the probability  $p$  that the logical qubit has had an odd number of bit flips.

The experimentally accessible data for this estimation consists of measurements of ancilla and flag qubits, contained in the vectors  $\delta\vec{s}(t)$  and  $\vec{s}_{\text{flag}}(t)$  of syndrome increments and flag measurements, and, at the end of the experiment, the readout of the data qubits. From this data qubit readout a final syndrome increment vector  $\delta\vec{f}(T)$  can be calculated. Depending on the measurement basis, it will only contain the  $X$  or the  $Z$  stabilizers.

Additionally, we also need to know the true bit flip parity  $p_{\text{true}}$ . To obtain this we initialize the logical qubit at  $|\psi_L\rangle \equiv |0\rangle$  ( $|\psi_L\rangle \equiv |1\rangle$  would be an equivalent choice) and then compare the final measured logical state to this initial logical state to obtain the true bit flip parity  $p_{\text{true}} \in \{0, 1\}$ .

An efficient decoder must be able to decode an arbitrary and unspecified

number of error correction cycles. As a feedforward neural network requires a fixed input size, it is impractical to train such a neural network to decode the entire syndrome data in a single step, as this would require a new network (and new training data) for every experiment with a different number of cycles. Instead, a neural network for quantum error correction must be cycle-based: It must be able to parse repeated input of small pieces of data (e.g. syndrome data from a single cycle) until called upon by the user to provide output. Importantly, this requires the decoder to be translationally invariant in time: It must decode late rounds of syndrome data just as well as the early rounds. To achieve this, we follow Ref. 211 and use a recurrent neural network of long short-term memory (LSTM) layers [258] — with one significant modification, which we now describe.

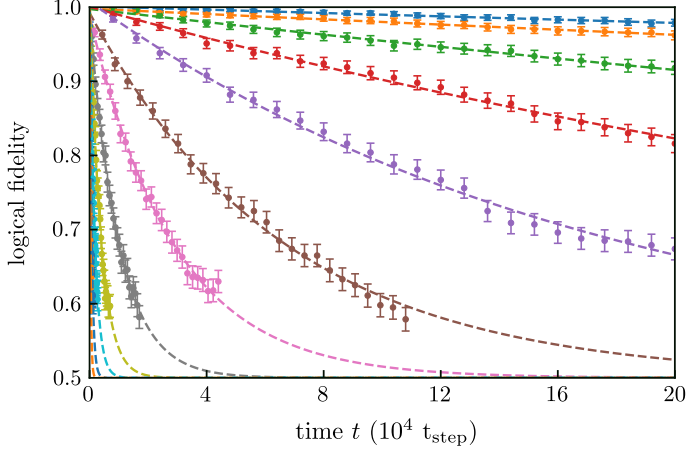
The time-translation invariance of the error propagation holds for the ancilla qubits, but it is broken by the final measurement of the data qubits — since any error in these qubits will not propagate forward in time. To extract the time-translation invariant part of the training data, in Ref. 211 two separate networks were trained in parallel, one with and one without the final measurement input. Here, we instead use a single network with two heads, as illustrated in Fig. 6.2. The upper head sees only the translationally invariant data, while the lower head solves the full decoding problem. In appendix 6.B we describe the details of the implementation.

The switch from two parallel networks to a single network with two heads offers several advantages: (1) The number of LSTM layers and the computational cost is cut in half; (2) The network can be trained on a single large error rate, then used for smaller error rates without retraining; (3) The bit flip probability from the upper head provides a so-called Pauli frame decoder [66].

In the training stage the bit flip probabilities  $p'$  and  $p \in [0, 1]$  from the upper and lower head are compared with the true bit flip parity  $p_{\text{true}} \in \{0, 1\}$ . By adjusting the weights of the network connections a cost function is minimized in order to bring  $p', p$  close to  $p_{\text{true}}$ . We carry out this machine learning procedure using the *TensorFlow* library [259].

After the training of the neural network has been completed we test the decoder on a fresh dataset. Only the lower head is active during the testing stage. If the output probability  $p < 0.5$ , the parity of bit flip errors is predicted to be even and otherwise odd. We then compare this to  $p_{\text{true}}$  and average over the test dataset to obtain the logical fidelity  $\mathcal{F}(t)$ . Using a two-parameter fit to [212]

$$\mathcal{F}(t) = \frac{1}{2} + \frac{1}{2}(1 - 2\epsilon_L)^{(t-t_0)/t_{\text{step}}}, \quad (6.4)$$



**Figure 6.3.:** Decay of the logical fidelity for a distance-3 color code. The curves correspond to different physical error rates  $\epsilon_{\text{phys}}$  per step, from top to bottom:  $1.6 \cdot 10^{-5}$ ,  $2.5 \cdot 10^{-5}$ ,  $4.0 \cdot 10^{-5}$ ,  $6.3 \cdot 10^{-5}$ ,  $1.0 \cdot 10^{-4}$ ,  $1.6 \cdot 10^{-4}$ ,  $2.5 \cdot 10^{-4}$ ,  $4.0 \cdot 10^{-4}$ ,  $6.3 \cdot 10^{-4}$ ,  $1.0 \cdot 10^{-3}$ ,  $1.6 \cdot 10^{-3}$ ,  $2.5 \cdot 10^{-3}$ . Each point is averaged over  $10^3$  samples. Error bars are obtained by bootstrapping. Dashed lines are two-parameter fits to Eq. (6.4).

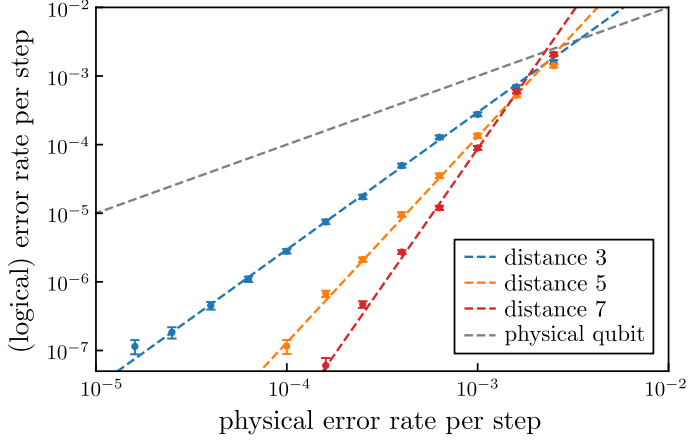
we determine the logical error rate  $\epsilon_L$  per step of the decoder.

## 6.4. Neural network performance

### 6.4.1. Power law scaling of the logical error rate

Results for the distance-3 color code are shown in Fig. 6.3 (with similar plots for distance-5 and distance-7 codes in App. 6.C). These results demonstrate that the neural network decoder is able to decode a large number of consecutive error correction cycles. The dashed lines are fits to Eq. (6.4), which allow us to extract the logical error rate  $\epsilon_L$  per step, for different physical error rates  $\epsilon_{\text{phys}}$  per step.

Figure 6.4 shows that the neural network decoder follows a power law scaling (6.2) with  $d$  fixed to the code distance. This shows that the decoder, once trained using a single error rate, operates equally efficiently when the error rate is varied, and that our flag error correction scheme is indeed fault-tolerant. The corresponding pseudothresholds (6.3) are listed in Table 6.1.



**Figure 6.4.:** In color: Log-log plot of the logical versus physical error rates per step, for distances  $d = 3, 5, 7$  of the color code. The dashed line through the data points has the slope given by Eq. (6.2). Quality of fit indicates that at least  $\lfloor \frac{1}{2}(d+1) \rfloor$  independent physical errors must occur in a round to generate a logical error in that round, so syndrome extraction is fault-tolerant. In gray: Error rate of a single physical (unencoded) qubit. The error rates at which this line intersects with the lines for the encoded qubits are the pseudothresholds.

### 6.4.2. Performance on realistic data

To assess the performance of the decoder in a realistic setting, we have implemented the distance-3 color code using a density matrix based simulator of superconducting transmon qubits [212]. We have then trained and tested the neural network decoder on data from this simulation. In Fig. 6.5 we compare the decay of the fidelity of the logical qubit as it results from the neural network decoder with the fidelity extracted from the simulation [212]. The latter fidelity determines via Eq. (6.4) the logical error rate  $\epsilon_{\text{optimal}}$  of an optimal decoder. For the distance-3 code we find  $\epsilon_L = 0.0148$  and  $\epsilon_{\text{optimal}} = 0.0132$  per microsecond. This can be used to calculate the decoder efficiency [212]  $\epsilon_{\text{optimal}}/\epsilon_L = 0.89$ , which measures the performance of the neural network decoder separate from uncorrectable errors. The dashed gray line is the average fidelity (following Eq. (6.4)) of a single physical qubit at rest, corresponding to an error rate of 0.0164 [212]. This demonstrates that, even with realistic experimental parameters, a logical qubit encoded with the color code has a longer life-time than a physical qubit.

distance $d$	pseudothreshold $\epsilon_{\text{pseudo}}$
3	0.0034
5	0.0028
7	0.0023

**Table 6.1.:** Pseudothresholds calculated from the data of Fig. 6.4, giving the physical error rate below which the logical qubit can store information for a longer time than a single physical qubit.

## 6.5. Conclusion

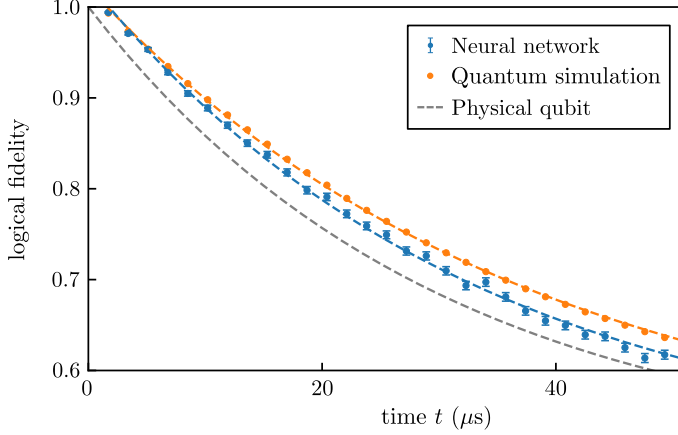
We have presented a machine-learning based approach to quantum error correction for the topological color code. We believe that this approach to fault-tolerant quantum computation can be used efficiently in experiments on near-term quantum devices with relatively high physical error rates (so that the neural network can be trained with relatively small datasets). In support of this, we have presented a density matrix simulation [212] of superconducting transmon qubits (Fig. 6.5), where we obtain a decoder efficiency of  $\eta_d = 0.89$ .

Independently of our investigation, three recent works have shown how a neural network can be applied to color code decoding. Refs. 250 and 253 only consider single rounds of error correction, and cannot be extended to a multi-round experiment or circuit-level noise. Ref. 251 uses the Steane and Knill error correction schemes when considering color codes, which are also fault-tolerant against circuit-level noise, but have larger physical qubit requirements than flag error correction. None of these works includes a test on a simulation of physical hardware.

## 6.A. Quantum circuits

### 6.A.1. Circuits for the Pauli error model

Fig. 6.6 shows the circuits for the measurements of the  $X$  and  $Z$  stabilizers in the Pauli error model. To each stabilizer, measured with the aid of an ancilla qubit, we associate a second “flag” ancilla qubit with the task of spotting faults of the first ancilla [241–245]. This avoids hook errors (errors that propagate from a single ancilla qubit onto two data qubits), which would reduce the distance of the code. After the measurement of the  $X$  stabilizers, all the ancillas are reset to  $|0\rangle$  and reused for the measurement



**Figure 6.5.:** Same as Fig. 6.3, but for a density matrix-based simulation of an array of superconducting transmon qubits. Each point is an average over  $10^4$  samples. The density matrix-based simulation gives the performance of an optimal decoder, with a logical error rate  $\epsilon_{\text{optimal}} = 0.0132$  per microsecond. From this, and the error rate  $\epsilon_L = 0.0148$  per microsecond obtained by the neural network, we calculate the neural network decoder efficiency to be 0.89. The average fidelity of an unencoded transmon qubit at rest with the same physical parameters is plotted in gray.

of the  $Z$  stabilizers. Before finally measuring the data qubits, we allow the circuit to run for  $T$  cycles.

### 6.A.2. Measurement processing for the density-matrix error model

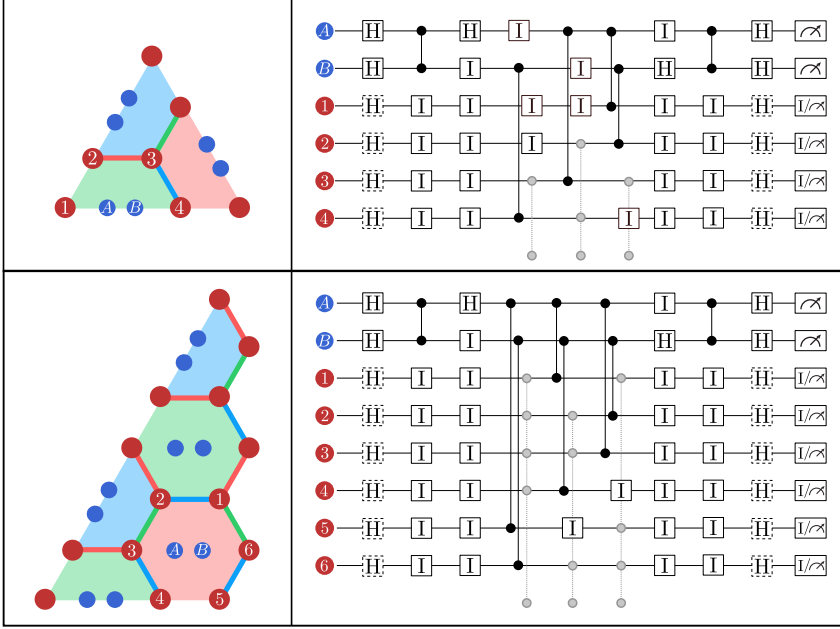
For the density matrix simulation, neither ancilla qubits nor flag qubits are reset between cycles, leading to a more involved extraction process of both  $\delta\vec{s}(t)$  and  $\vec{s}_{\text{flag}}(t)$ , as we now explain.

Let  $\vec{m}(t)$  and  $\vec{m}_{\text{flag}}(t)$  be the actual ancilla and flag qubit measurements taken in cycle  $t$ , and  $\vec{m}^0(t)$ ,  $\vec{m}_{\text{flag}}^0(t)$  be compensation vectors of ancilla and flag measurements that would have been observed had no errors occurred in this cycle. Then,

$$\delta\vec{s}(t) = \vec{m}(t) + \vec{m}^0(t) \pmod{2}, \quad (6.5)$$

$$\vec{s}_{\text{flag}}(t) = \vec{m}_{\text{flag}}(t) + \vec{m}_{\text{flag}}^0(t) \pmod{2}. \quad (6.6)$$





**Figure 6.6.:** Top left: Schematic of a 6-6-6 color code with distance 3. Top right: Stabilizer measurement circuits for a plaquette on the boundary. Bottom left: Partial schematic of a 6-6-6 color code with distance larger than 3. Bottom right: Stabilizer measurement circuits for a plaquette in the bulk. For the circuits in the right panels, the dashed Hadamard gates are only present when measuring the  $X$  stabilizers, and are replaced by idling gates for the  $Z$  stabilizer circuits; the grayed out gates correspond to conditional-phase gates between the considered data qubits and ancillas belonging to other plaquettes; and the data qubits are only measured after the last round of error correction, otherwise they idle whilst the ancillas are measured.

Calculation of the compensation vectors  $\vec{m}^0(t)$  and  $\vec{m}_{\text{flag}}^0(t)$  requires knowledge of the stabilizer  $\vec{s}(t-1)$ , and the initialization of the ancilla qubits  $\vec{m}(t-1)$  and the flag qubits  $\vec{m}_{\text{flag}}(t-1)$ , being the combination of the effects of individual non-zero terms in each of these.

Note that a flag qubit being initialized in  $|1\rangle$  will cause errors to propagate onto nearby data qubits, but these errors can be predicted and removed prior to decoding with the neural network. In particular, let us concatenate  $\vec{m}(t)$ ,  $\vec{m}_{\text{flag}}(t)$  and  $\vec{s}(t)$  to form a vector  $\vec{d}(t)$ . The update may

## 6. Neural network decoder for topological color codes with circuit level noise

then be written as a matrix multiplication:

$$\vec{m}_{\text{flag}}^0(t) = M_f \vec{d}(t-1) \mod 2, \quad (6.7)$$

Where  $M_f$  is a sparse, binary matrix. The syndromes  $\vec{s}(t)$  may be updated in a similar fashion

$$\vec{s}(t) = \vec{s}(t-1) + \delta \vec{s}(t) + M_s \vec{d}(t-1) \mod 2, \quad (6.8)$$

where  $M_s$  is likewise sparse. Both  $M_f$  and  $M_s$  may be constructed by modeling the stabilizer measurement circuit in the absence of errors. The sparsity in both matrices reflect the connectivity between data and ancilla qubits; for a topological code, both  $M_f$  and  $M_s$  are local. The calculation of the syndrome increments  $\delta \vec{s}(t)$  via Eq. (6.5) does not require prior calculation of  $\vec{s}(t)$ .

## 6.B. Details of the neural network decoder

### 6.B.1. Architecture

The decoder consists of a double headed network, see Fig. 6.2, which we implement using the *TensorFlow* library [259]\*. The network maps a list of syndrome increments  $\delta \vec{s}(t)$  and flag measurements  $\vec{s}_{\text{flag}}(t)$  with  $t/t_{\text{cycle}} = 1, 2, \dots, T$  to a pair of probabilities  $p', p \in [0, 1]$ . (In what follows we measure time in units of the cycle duration  $t_{\text{cycle}} = N_0 t_{\text{step}}$ , with  $N_0 = 20$ .) The lower head gets as additional input a single final syndrome increment  $\delta \vec{f}(T)$ . The cost function  $I$  that we seek to minimize by varying the weight matrices  $\mathbf{w}$  and bias vectors  $\vec{b}$  of the network is the cross-entropy

$$H(p_1, p_2) = -p_1 \log p_2 - (1 - p_1) \log(1 - p_2) \quad (6.9)$$

between these output probabilities and the true final parity  $p_{\text{true}} \in \{0, 1\}$  of bit flip errors:

$$I = H(p_{\text{true}}, p) + \frac{1}{2} H(p_{\text{true}}, p') + c \|\mathbf{w}_{\text{EVAL}}\|^2. \quad (6.10)$$

The term  $c \|\mathbf{w}_{\text{EVAL}}\|^2$  with  $c \ll 1$  is a regularizer, where  $\mathbf{w}_{\text{EVAL}} \subset \mathbf{w}$  are the weights of the evaluation layer.

The body of the double headed network is a recurrent neural network,

---

\*The source code of the neural network decoder can be found at [https://github.com/baireuther/neural\\_network\\_decoder](https://github.com/baireuther/neural_network_decoder).

consisting of two LSTM layers [258, 260, 261]. Each of the LSTM layers has two internal states, representing the long-term memory  $\vec{c}_t^{(i)} \in \mathbb{R}^N$  and the short-term memory  $\vec{h}_t^{(i)} \in \mathbb{R}^N$ , where  $N = 32, 64, 128$  for distances  $d = 3, 5, 7$ . Internally, an LSTM layer consists of four simple neural networks that control how the short- and long-term memory are updated based on their current states and new input  $x_t$ . Mathematically, it is described by the following equations [260, 261]:

$$\vec{i}_t = \sigma(\mathbf{w}_i \vec{x}_t + \mathbf{v}_i \vec{h}_{t-1} + \vec{b}_i), \quad (6.11a)$$

$$\vec{f}_t = \sigma(\mathbf{w}_f \vec{x}_t + \mathbf{v}_f \vec{h}_{t-1} + \vec{b}_f), \quad (6.11b)$$

$$\vec{o}_t = \sigma(\mathbf{w}_o \vec{x}_t + \mathbf{v}_o \vec{h}_{t-1} + \vec{b}_o), \quad (6.11c)$$

$$\vec{m}_t = \tanh(\mathbf{w}_m \vec{x}_t + \mathbf{v}_m \vec{h}_{t-1} + \vec{b}_m), \quad (6.11d)$$

$$\vec{c}_t = \vec{f}_t \odot \vec{c}_{t-1} + \vec{i}_t \odot \vec{m}_t \quad (6.11e)$$

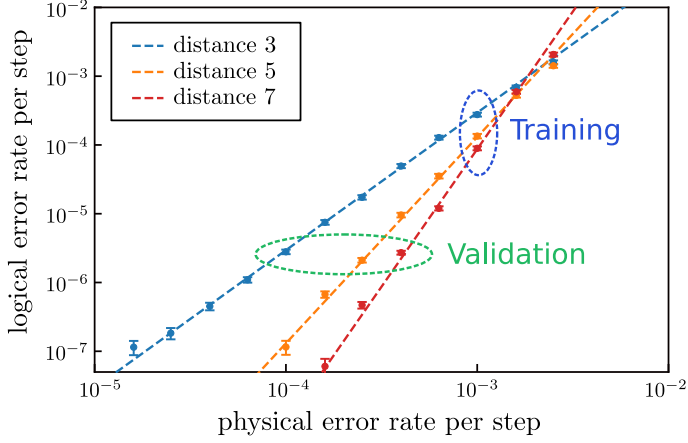
$$\vec{h}_t = \vec{o}_t \odot \tanh(\vec{c}_t). \quad (6.11f)$$

Here  $\mathbf{w}$  and  $\mathbf{v}$  are weight matrices,  $\vec{b}$  are bias vectors,  $\sigma$  is the sigmoid function, and  $\odot$  is the element-wise product between two vectors. The letters  $i$ ,  $m$ ,  $f$ , and  $o$  label the four internal neural network gates: input, input modulation, forget, and output. The first LSTM layer gets the syndrome increments  $\delta \vec{s}(t)$  and flag measurements  $\vec{s}_{\text{flag}}(t)$  as input, and outputs its short term memory states  $\vec{h}_t^{(1)}$ . These states are in turn the input to the second LSTM layer.

The heads of the network consist of a single layer of rectified linear units, whose outputs are mapped onto a single probability using a sigmoid activation function. The input of the two heads is the last short-term memory state of the second LSTM layer, subject to a rectified linear activation function  $\text{ReL}(\vec{h}_T^{(2)})$ . For the lower head we concatenate  $\text{ReL}(\vec{h}_T^{(2)})$  with the final syndrome increment  $\delta \vec{f}(T)$ .

## 6.B.2. Training and evaluation

We use three separate datasets for each code distance. The training dataset is used by the optimizer to optimize the trainable variables of the network. It consists of  $2 \cdot 10^6$  sequences of lengths between  $T = 1$  and  $T = 40$  at a large error rate of  $p = 10^{-3}$  for distances 3 and 5, and of  $5 \cdot 10^6$  sequences for distance 7. At the end of each sequence, it contains the final syndrome increment  $\delta \vec{f}(T)$  and the final parity of bit flip errors  $p_{\text{true}}$ . After each training epoch, consisting of 3000 to 5000 mini-batches of size 64, we

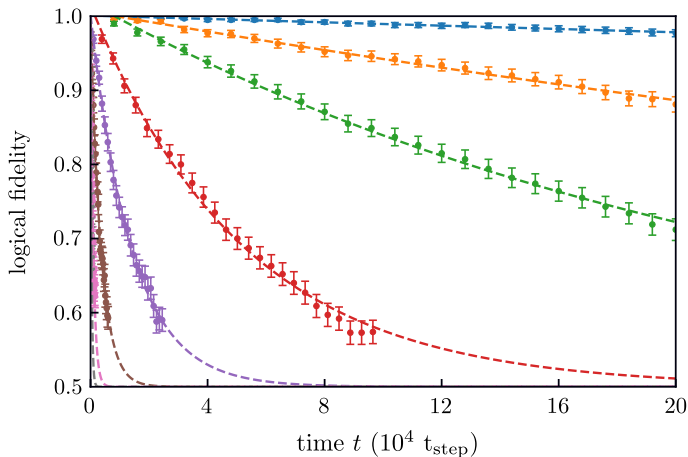


**Figure 6.7.:** Same as Fig. 6.4. The blue ellipse indicates the error rates used during training, and the green ellipse indicates the error rates used for validation.

validate the network (using only the lower head) on a validation dataset consisting of  $10^3$  sequences of 30 different lengths between 1 and  $10^4$  cycles. By validating on sequences much longer than the sequences in the training dataset, we select the instance of the decoder that generalizes best to long sequences. The error rates of the validation datasets are chosen such that they are the largest error rate for which the expected logical fidelity after  $10^4$  cycles is still larger than 0.6 (see Fig. 6.7), because if the logical fidelity approaches 0.5 a meaningful prediction is no longer possible. The error rates of the validation datasets are  $1 \cdot 10^{-4}$ ,  $2.5 \cdot 10^{-4}$ ,  $4 \cdot 10^{-4}$  for distances 3, 5, 7 respectively. To avoid unproductive fits during the early training stages, we calculate the logical error rate with a single parameter fit to Eq. (6.4) by setting  $t_0 = 0$  during validation. If the logical error rate reaches a new minimum on the validation dataset, we store this instance of the network.

We stop the training after  $10^3$  epochs. One training epoch takes about one minute for distance 3 (network size  $N = 32$ ) when training on sequences up to length  $T = 20$  and about two minutes for sequences up to length  $T = 40$  on an Intel(R) Xeon(R) CPU E3-1270 v5 @ 3.60GHz. For distance 5 ( $N = 64$ ,  $T = 1, 2, \dots, 40$ ) one epoch takes about five minutes and for distance 7 ( $N = 128$ ,  $T = 1, 2, \dots, 40$ ) about ten minutes.

To keep the computational effort of the data generation tractable, for the density matrix-based simulation (Fig. 6.5) we only train on  $10^6$  sequences of lengths between  $T = 1$  and  $T = 20$  cycles and validate on  $10^4$  sequences

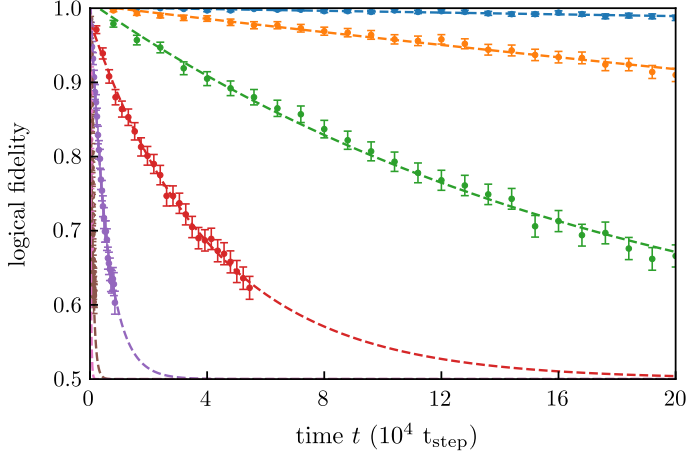


**Figure 6.8.:** Same as Fig. 6.3 for a distance-5 code; the physical error rate  $\epsilon_{\text{phys}}$  from top to bottom is:  $1.0 \cdot 10^{-4}$ ,  $1.6 \cdot 10^{-4}$ ,  $2.5 \cdot 10^{-4}$ ,  $4.0 \cdot 10^{-4}$ ,  $6.3 \cdot 10^{-4}$ ,  $1.0 \cdot 10^{-3}$ ,  $1.6 \cdot 10^{-3}$ ,  $2.5 \cdot 10^{-3}$ .

of lengths between  $T = 1$  and  $T = 30$  cycles. For the density matrix-based simulation, all datasets have the same error rate.

We train using the Adam optimizer [262] with a learning rate of  $10^{-3}$ . To avoid over-fitting and reach a better generalization of the network to unseen data, we employ two additional regularization methods: Dropout and weight regularization. Dropout with a keep probability of 0.8 is applied to the output of each LSTM layer and to the output of the hidden units of the evaluation layers. Weight regularization, with a prefactor of  $c = 10^{-5}$ , is only applied to the weights of the evaluation layers, but not to the biases. The hyperparameters for training rate, dropout, and weight regularization were taken from [211]. The network sizes were chosen by try and error to be as small as possible without fine-tuning, restricted to powers of two  $N = 2^n$ .

After training is complete we evaluate the decoder on a test dataset consisting of  $10^3$  ( $10^4$  for the density matrix-based simulation) sequences of lengths such that the logical fidelity decays to approximately 0.6, but no more than  $T = 10^4$  cycles. Unlike for the training and validation datasets, for the test dataset we sample a final syndrome increment and the corresponding final parity of bit flip errors after each cycle. We then select an evenly distributed subset of  $t_n = n\Delta T < T_{\text{max}}$  cycles, where  $\Delta T$  is the smallest integer for which the total number of points is less than 50,



**Figure 6.9.:** Same as Fig. 6.3 for a distance-7 code; the physical error rate  $\epsilon_{\text{phys}}$  from top to bottom is:  $1.6 \cdot 10^{-4}$ ,  $2.5 \cdot 10^{-4}$ ,  $4.0 \cdot 10^{-4}$ ,  $6.3 \cdot 10^{-4}$ ,  $1.0 \cdot 10^{-3}$ ,  $1.6 \cdot 10^{-3}$ ,  $2.5 \cdot 10^{-3}$ .

for evaluation. This is done in order to reduce the needed computational resources. The logical error rate  $\epsilon$  per step is determined by a fit of the fidelity to Eq. (6.4).

### 6.B.3. Pauli frame updater

We operate the neural network as a bit-flip decoder, but we could have alternatively operated it as a Pauli frame updater. We briefly discuss the connection between the two modes of operation.

Generally, a decoder executes a classical algorithm that determines the operator  $P(t) \in \Pi^n$  (the so-called Pauli frame) which transforms  $|\psi_L(t)\rangle$  back into the logical qubit space  $\mathcal{H}_{\vec{0}} = \mathcal{H}_L$ . Equivalently (with minimal overhead), a decoder may keep track of logical parity bits  $\vec{p}$  that determine whether the Pauli frame of a ‘simple decoder’ [247] commutes with a set of chosen logical operators for each logical qubit.

The second approach of bit-flip decoding has two advantages over Pauli frame updates: Firstly, it removes the gauge degree of freedom of the Pauli frame ( $SP(t)$  is an equivalent Pauli frame for any stabilizer  $S$ ). Secondly, the logical parity can be measured in an experiment, where no ‘true’ Pauli frame exists (due to the gauge degree of freedom).

Note that in the scheme where flag qubits are used without reset, the

errors from qubits initialized in  $|1\rangle$  may be removed by the simple decoder without any additional input required by the neural network.

## 6.C. Results for distance-5 and distance-7 codes

Figures 6.8 and 6.9 show the decay curves for the  $d = 5$  and  $d = 7$  color codes, similar to the  $d = 3$  decay curves shown in figure 6.3 in the main text.





## Part III.

# Quantum Algorithms for digital quantum simulation



# 7. Majorana-based fermionic quantum computation

## 7.1. Introduction

Particle exchange statistics is a fundamental quantum property that distinguishes commuting spin or qubit degrees of freedom from anticommuting fermions, despite single particles in both systems only having two quantum states. Different exchange statistics cause a different set of Hamiltonian terms to be local, or even physically possible. For example, although it is Hermitian, the linear superposition of a fermionic creation and annihilation operator  $c + c^\dagger$  never occurs as a Hamiltonian term in nature due to violating fermion parity conservation, whilst spin systems have no such restrictions. Despite these differences, it is possible to simulate fermions using qubits and vice versa [47]. Such simulation necessarily incurs overhead because of the need to transform local fermion operators into non-local qubit ones by using, for example, the Jordan-Wigner transformation. Because quantum simulation of the electronic structure of molecules is a promising application of quantum computation [263], much recent work focused on minimizing this overhead of simulating fermionic Hamiltonians with qubits [264–266].

Majorana zero modes (also Majorana modes or just Majoranas) are non-abelian particles, with two Majoranas combining to form a single fermion (see e.g. Refs. [267–269] for a review). Spatially separating two Majoranas protects this fermionic degree of freedom, and provides a natural implementation of a topological quantum computer [31, 270]. Further, conservation of fermion parity prevents creating a superposition between the two different parity states of two Majoranas, and therefore most of the existing proposals combine 4 Majoranas with a fixed fermion parity into a single qubit.

---

The contents of this chapter have been published in T. E. O’Brien, P. Rožek and A. R. Akhmerov, *Phys. Rev. Lett.* **120** (22), 220504 (2018).

## 7. Majorana-based fermionic quantum computation

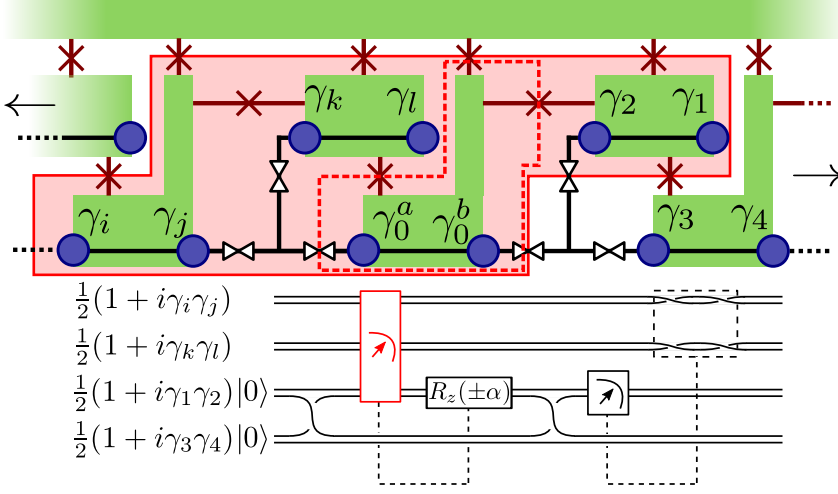
Fermionic quantum computation [47] was so far not actively pursued because of the lack of known ways to protect fermionic degrees of freedom from dephasing. In this chapter, we observe that Majoranas naturally offer this protection, while in addition providing a platform for implementing quantum chemistry algorithms. We therefore show that for the problem of simulating fermionic systems on a Majorana quantum computing architecture, it is both possible and preferable to use fermions composed from pairs of Majoranas instead of further combining pairs of these fermions to form single qubits. Formulating fermionic quantum simulation algorithms in terms of fermions imposes the fermion parity conservation at the hardware level, and prohibits a large class of errors bringing the simulator out of the physical subspace. Furthermore, working natively with fermions, we remove the need for the Jordan-Wigner (or related) transformation to map a fermionic problem to a spin system. When simulating a typical quantum chemistry Hamiltonian, our approach results in a more dense encoding of the computational degrees of freedom. The benefit from using the fermionic degrees of freedom becomes more important in simulating local fermionic Hamiltonians, such as the Hubbard model, allowing the simulation of unitary time evolution in  $\mathcal{O}(1)$  time per Trotter step, and further reducing the cost of pre-error-correction quantum simulation [271]. Finally, we show how to apply the known magic state distillation protocol in fermionic quantum computation. Combined with the recent realization of the fermionic error correction [272] this provides a fault-tolerant fermionic quantum computer.

### 7.2. Description of the architecture

Our approach relies on the known set of ingredients to perform universal operations with Majorana states [273]: controllable Josephson junctions, direct Majorana coupling, and Coulomb energy. A possible architecture implementing a Majorana-based fermionic quantum processor is shown in Fig. 7.1. Because our system cannot be separated into blocks with a fixed fermion parity, the protection of the quantum degrees of freedom is only possible if different parts of the system are connected to a common superconducting ground \*. Turning off some of the Josephson junctions (these may be either flux-controlled SQUIDs or gate-controlled [274, 275]) then isolates a part of the system, and generates a Coulomb interaction

---

\*The need to use a common superconducting ground makes it impossible to utilize the partial protection from quasiparticle poisoning by applying Coulomb blockade to superconducting islands containing individual qubits [288].



**Figure 7.1.:** Top: a 1D implementation of a Majorana circuit. Majoranas (blue dots) occur at either the edge of a nanowire (black line) or as it crosses the boundary of a superconductor (light green). Josephson junctions (red crossed lines) connect superconducting islands to a common base, allowing for parallel joint parity measurements. Fully-tunable T-junctions (valve symbols) allow for a computational Majorana to be shifted from one end of any coupled set of itself and two braiding ancillas (prepared in a known state) to another end. Bottom: an implementation of a weight-four Majorana rotation (Eq. 7.6) using the labeled qubits in the design. The operation of individual circuit elements is listed in Table. 7.1. The highlighted parity measurement is performed by isolating the highlighted area of the architecture via tunable Josephson junctions, and measuring the total charge parity. This requires a separate preparation of the Majoranas  $\gamma_0^a$  and  $\gamma_0^b$  (dashed red box) in the  $i\gamma_0^a\gamma_0^b = 1$  state (which is also required to use these as spare sites for braiding).

[276, 277]

$$H_C = i^{N/2} E_C \prod_k^N \gamma_k, \quad (7.1)$$


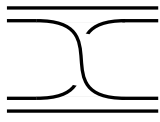
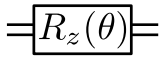
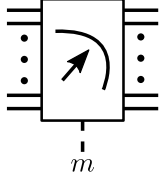
that couples all the Majorana modes  $\gamma_i$  belonging to the isolated part of the system with the charging energy  $E_C$ . An example of such coupling acting on 8 Majorana modes is shown by a red box in Fig. 7.1. Finally, gate-controlled T-junctions exert the interaction

$$H_M = iE_M \gamma_j \gamma_k, \quad (7.2)$$

## 7. Majorana-based fermionic quantum computation

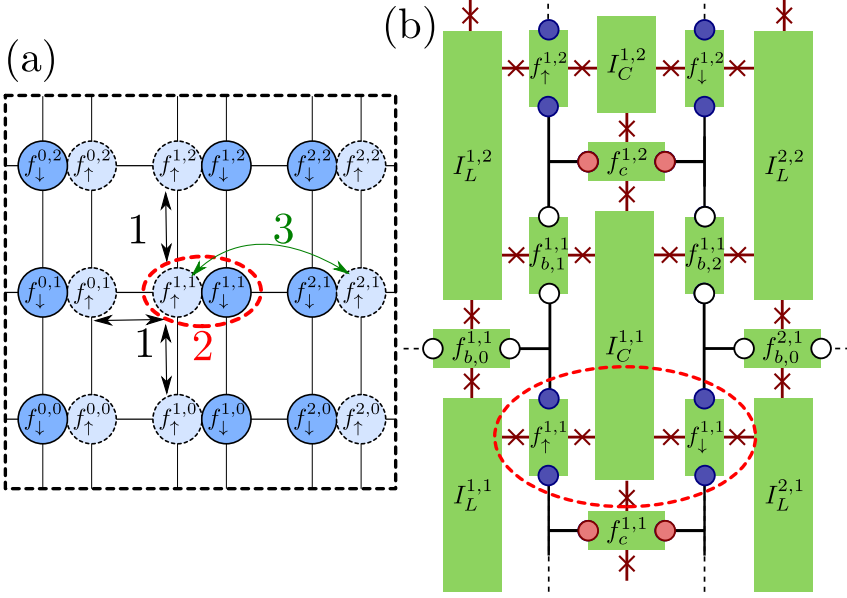
on any two Majorana modes coupled by a T-junction, with  $E_M$  the Majorana coupling energy.

Controllable pairwise interactions between Majorana modes [278, 279] or two-Majorana parity measurements [280] allow the implementation of braiding, while the joint readout of the fermionic parity of more than 2 Majorana modes generates the rest of the Clifford group [273]. Finally, a diabatic pulse of a two-Majorana coupling implements an unprotected phase gate  $e^{\theta\gamma_i\gamma_j}$ . We summarize these elementary gates that serve as a basis of our protocol in Table 7.1. This gate set is computationally universal within a fixed fermion parity sector [47].

Name	Element	Operation
Preparation	$i\gamma_1\gamma_2  0\rangle \equiv$	Prepare $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
Braiding		$\begin{pmatrix} e^{i\pi/4} & 0 \\ 0 & e^{-i\pi/4} \end{pmatrix}$
Braiding		$\begin{pmatrix} 1 & 0 & 0 & -i \\ 0 & 1 & -i & 0 \\ 0 & -i & 1 & 0 \\ -i & 0 & 0 & 1 \end{pmatrix}$
Rotation		$\begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}$
Measurement		$\sum_{P(\phi)=m}  \phi\rangle\langle\phi $

**Table 7.1.:** Basic circuit elements we allow in our computation scheme. The above is sufficient to generate universal quantum computation in the single-parity sector. Computational degrees of freedom are formed by two Majoranas, and are therefore represented as a double line. Preparation, braiding, and measurement gates are assumed to be topologically protected. The  $R_z(\theta)$  rotation is not topologically protected, but may be distilled via our magic state distillation protocol. The measurement projects our system onto a state of definite parity  $P(\phi)$ , being the sum  $\sum_{i,j} \frac{1}{2}(1 + i\gamma_i\gamma_j)$  of the pairs of Majoranas  $\gamma_i, \gamma_j$  on islands connected to ground via Josephson junctions.

### 7.3. Quantum algorithms



**Figure 7.2.:** A 2d Majorana architecture to implement the Hubbard model on a square lattice. (a) A schematic description of the initial layout of the fermions (each of which is made of two Majoranas). Lines denote fermions separated by ancilla Majoranas only. Our scheme groups the 11 Trotter steps into three stages as numbered, which are performed in series. (b) A physical architecture to support the schematic of (a). Wires on superconducting islands and T-junction symbols from Fig. 7.1 have been removed to prevent cluttering; it is still assumed that all T-junctions are fully tunable. Majoranas are colored according to their designation; blue for system fermions, red for control ancillas, and white for braiding and phase ancillas. An example spin-1/2 fermion supported on four Majoranas (the minimum possible) is matched to (a)

The above gate set is sufficient to construct circuits for time evolution, quantum phase estimation (QPE), and a variational quantum eigensolver—the unitary coupled cluster ansatz (UCC). Most fermionic systems have Hamiltonians constructed from twofold and fourfold fermionic terms:

$$H = \sum_{i,j} h_{i,j} \hat{f}_i^\dagger \hat{f}_j + \sum_{i,j,k,l} \hat{f}_i^\dagger \hat{f}_j^\dagger \hat{f}_k \hat{f}_l. \quad (7.3)$$

## 7. Majorana-based fermionic quantum computation

Here,  $\hat{f}_i^\dagger$  ( $\hat{f}_i$ ) is the creation (annihilation) operator for an electron. This is equivalent to a sum over 2 and 4-fold Majorana terms:

$$H = \sum_{i,j} i g_{i,j} \gamma_i \gamma_j + \sum_{i,j,k,l} g_{i,j,k,l} \gamma_i \gamma_j \gamma_k \gamma_l. \quad (7.4)$$

Time evolution is performed by applying the Trotter expansion of the evolution operator  $e^{iHt}$  \*:

$$e^{iH\Delta t} \underset{\Delta t \rightarrow 0}{\approx} \prod_{i,j} e^{-g_{i,j} \gamma_i \gamma_j \Delta t} \prod_{i,j,k,l} e^{i g_{i,j,k,l} \gamma_i \gamma_j \gamma_k \gamma_l \Delta t}, \quad (7.5)$$

and thus requires consecutive application of the unitary operators  $e^{\theta \gamma_i \gamma_j}$  and  $e^{i\theta \gamma_i \gamma_j \gamma_k \gamma_l}$ . We therefore introduce the weight- $2N$  Majorana rotation operator

$$\exp \left\{ i\theta \prod_{n=1}^N i \gamma_{2n-1} \gamma_{2n} \right\}, \quad (7.6)$$

that forms the basis of all the algorithms we consider.

A Majorana rotation may be performed using a generic circuit with an additional four-Majorana ancilla qubit. To demonstrate, the circuit of Fig. 7.1 applies a Majorana rotation  $e^{i\theta \gamma_i \gamma_j \gamma_k \gamma_l}$ . The same scheme implements weight-two Majorana rotations by removing Majoranas  $\gamma_k$  and  $\gamma_l$ , and higher weight- $2N$  Majorana rotations by adding  $2N - 4$  more Majoranas to the correlated parity check and conditional final braiding. The ancillary Majoranas  $\gamma_0^a$  and  $\gamma_0^b$  used for the braiding begin in the parity eigenstate  $i\gamma_0^a \gamma_0^b = 1$ . The eight-Majorana charge parity measurement  $\gamma_i \gamma_j \gamma_k \gamma_l \gamma_0^a \gamma_0^b \gamma_2 \gamma_1$  (implemented by isolating the circled superconducting islands in Fig. 7.1) therefore reduces to the 6-Majorana measurement highlighted in the circuit. The unprotected rotation by the angle  $\alpha = \theta + \frac{\pi}{2}$  both corrects an unwanted phase from the braiding of  $\gamma_2$  and  $\gamma_3$ , and applies the non-Clifford rotation by  $\theta$ .

Quantum phase estimation requires the unitary evolution of a state (prepared across a set of system qubits) conditional on a set of ancilla qubits, which then have the eigenphases of the unitary operator encoded upon them [22]. For the purposes of simulating quantum chemistry, a common choice of this operator is the time evolution operator, approximated by the

---

\*We have not discussed post-Trotter methods such as [289–291] in this chapter. However, these methods still require the Jordan-Wigner transformation or equivalent to represent a fermionic Hamiltonian on a qubit architecture. As such, they gain a similar advantage to the studied Trotterized evolution of  $e^{iHt}$  from a Majorana-based fermion implementation.



Trotter expansion. In App. 7.A, we show how to encode the ancilla qubit non-locally across an array of fermions, each of those controlling the unitary evolution of a local Hamiltonian term. This reduces the requirements for QPE to consecutive operations of weight-four and weight-six Majorana rotations, with two Majoranas in each rotation belonging to an ancilla fermion. In App. 7.B we show how this circuit is used to execute a single Trotter step for a fully-connected fourth-order Hamiltonian in  $O(N^3)$  time.

Variational quantum eigensolvers prepare a trial state  $|\psi(\vec{\theta})\rangle$  from a circuit depending on a set of variational parameters  $\vec{\theta}$ , which are then tuned to minimize the energy  $\langle\psi(\vec{\theta})|H|\psi(\vec{\theta})\rangle$  [281]. One example of such ansatz is the UCC-2, which uses the exponential of the second order expansion of the cluster operator:

$$|\psi(t_p^r, t_{pq}^{rs})\rangle = e^{T^{(2)} - T^{(2)\dagger}} |\Phi_{\text{ref}}\rangle,$$

$$T^{(2)} = \sum_{p,r} t_p^r \hat{f}_p^\dagger \hat{f}_r + \sum_{p,q,r,s} t_{pq}^{rs} \hat{f}_p^\dagger \hat{f}_q^\dagger \hat{f}_r \hat{f}_s.$$

After Trotterizing, this requires only weight-two or -four Majorana rotations to prepare.

When the Hamiltonian contains a small fraction of all possible second- or fourth-order terms, the lack of Jordan-Wigner strings gives our fermionic architecture an advantage over qubit-based implementations. As an example, we consider the Hubbard model on a square lattice, with Hamiltonian

$$H = -t \sum_{\langle i,j \rangle, \sigma} \hat{f}_{i,\sigma}^\dagger \hat{f}_{j,\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} - \mu \sum_{i\sigma} \hat{n}_{i\sigma}. \quad (7.7)$$

Here  $\sigma$  is a spin index, and the first sum is goes over the pairs of nearest neighbor lattice sites, while  $t$ ,  $\mu$ , and  $U$  are the model parameters [282]. Rewriting the Hubbard model Hamiltonian in terms of Majorana operators  $\hat{f}_{i\sigma}^\dagger = \frac{1}{2}(\gamma_{\sigma,1}^i + i\gamma_{\sigma,2}^i)$  gives:

$$H = \frac{t}{2} \sum_{\langle i,j \rangle, \sigma} i\gamma_{\sigma,1}^i \gamma_{\sigma,2}^j + N(\frac{U}{4} - \mu)$$

$$+ \frac{i}{4}(U - 2\mu) \sum_{i,\sigma} \gamma_{\sigma,1}^i \gamma_{\sigma,2}^i - \frac{U}{4} \sum_i \gamma_{\uparrow,1}^i \gamma_{\uparrow,2}^i \gamma_{\downarrow,1}^i \gamma_{\downarrow,2}^i. \quad (7.8)$$

This gives in total 11 terms per site  $i$  that need to be simulated for quantum phase estimation or unitary time evolution. In Fig. 7.2 we show



a 2d architecture that implements parallel application of Trotter steps across the entire lattice. For unitary evolution, this scheme is 33% dense, with 12 Majoranas used per site with 2 fermions. For parallel QPE we use an additional ancilla per site (following App. 7.A), making the scheme 50% dense. We detail the computation scheme for QPE in App. 7.C, achieving a  $O(1)$  circuit depth per controlled unitary evolution step. This should be compared first to the  $O(N^{1/2})$  circuit depth in the case of a qubit implementation via a parallelized Jordan-Wigner transformation [283]. This circuit depth can be reduced to  $O(\log(N))$  if the Bravyi-Kitaev transformation [47] is used instead, but at the cost of requiring dense qubit connectivity. Separate encodings [284, 285] also exist to reduce the circuit depth to  $O(1)$ , at a cost of doubling the required number of qubits. It is likewise possible to achieve a similar  $O(1)$  circuit depth, assuming the ability to couple a global resonator to every qubit in a superconducting architecture [286].

The required ingredient for universal fermionic quantum computation—a Majorana rotation by an arbitrary angle  $\theta$ —is most simply implemented using an unprotected coupling between two Majoranas. In a scalable architecture this gate needs to have increasingly higher fidelity so that it may be applied an arbitrary number of times without failure. In Fig. 7.3 we develop a high fidelity Majorana rotation using the magic state distillation protocol of [48] to perform fermionic gates. In this procedure, we generate 5 low-fidelity  $|T\rangle = \cos(\beta)|0\rangle + e^{i\pi/4}\sin(\beta)|1\rangle$  states ( $\cos(2\beta) = \frac{1}{\sqrt{3}}$ ) on four-Majorana qubits, then combine them to obtain a single higher fidelity  $|T\rangle$  state on a qubit (assuming topologically-protected Clifford gates). We then use an average of 3 distilled  $|T\rangle$  states to perform a  $\theta = \pm \frac{\pi}{12}$  Majorana rotation. On average, this procedure requires 15 noisy  $|T\rangle$  states, 225 braidings and 66 measurements. We furthermore use 20 Majoranas to make the 5 noisy  $|T\rangle$  qubit states, due to the  $|T\rangle$  state of a single fermion breaking parity conservation.

## 7.4. Conclusion

In summary, we have demonstrated a Majorana-based scheme for fermionic quantum computation. We then adapted this scheme to simulate interacting fermionic Hamiltonians using both the QPE and VQC algorithms, and modified it to simulate the Hubbard model using a constant-depth circuit per time-evolution step. While our fermionic scheme has advantages compared to using qubits, finding optimal circuit layouts for both a general purpose fermionic quantum computation and problem-specific

## 7. Majorana-based fermionic quantum computation

ones, like the Hubbard model simulator remain an obvious point for further research. Further, our implementation of magic state distillation is a direct translation of the original scheme, and it should be possible to find a smaller circuit operating only on fermions, for example using the minimal fermionic error correcting circuit of [287]. A final open direction of further research is combining our circuits with quantum error correction [272, 287], which would enable fault-tolerant fermionic quantum computation.

### 7.A. Preparing extended ancilla qubits for quantum phase estimation

The QPE algorithm requires the application of a unitary operator conditional on an ancilla qubit, which naively would require each Trotter step to be performed in series as the ancilla qubit is passed through the system. The following method parallelizes the QPE algorithm at a cost of  $O(N)$  ancilla qubits and a constant depth preparation circuit, which may well be preferable. We make this trade by preparing a large cat state on  $4n$  Majoranas by the circuit in Fig. 7.4. First, we prepare  $n \times 4$  Majoranas in the  $\frac{1}{2}(|00\rangle + |11\rangle)$  state on Majoranas  $\gamma_{4j}\gamma_{4j+1}\gamma_{4j+2}\gamma_{4j+3}$  for  $j = 0, \dots, n-1$ . Then, making the joint parity measurements  $\gamma_{4j+2}\gamma_{4j+3}\gamma_{4j+4}\gamma_{4j+5}$  for  $j = 0, \dots, n-2$  forces our system into an equal superposition of

$$\frac{1}{\sqrt{2}} \left( \left| \prod_{j=0}^{n-1} x_{2j} x_{2j+1} \right\rangle + \left| \prod_{j=0}^{n-1} \bar{x}_{2j} \bar{x}_{2j+1} \right\rangle \right), \quad (7.9)$$

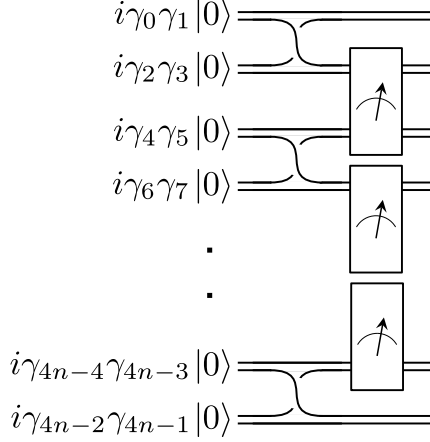
where  $x_j \in \{0,1\}$  is the parity on the  $j$ th fermion ( $\bar{x} = 1 - x$ ), and  $x_{2j} \oplus x_{2j-1}$  is determined by the outcome of the joint parity measurement. This can then be converted to the GHZ state  $\frac{1}{\sqrt{2}}(|00\dots 0\rangle + |11\dots 1\rangle)$  by braiding (or the value of  $x_j$  can be stored and used to decide whether to rotate by  $\theta$  or  $-\theta$ ). The rotations to be performed for QPE may then be controlled by *any* of the pairs of Majoranas defining a single fermion, and so we may spread this correlated ancilla over our system as required to perform rotations. As the interaction between ancilla qubits and system qubits is limited to a single joint parity measurement per Trotter step, we expect that although  $n$  should scale as  $O(N)$  to allow for parallelizing the circuit, the prefactor will be quite small. At the end of the QPE circuit, we recover the required phase by rotating  $\exp(i\frac{\pi}{4}\gamma_{4j+1}\gamma_{4j+2})$  for  $j = 0, \dots, n-1$  and reading out the parity of all Fermions individually.

### 7.B. Algorithm to perform Trotter steps in $O(N^3)$ time

Starting from the state

$$\frac{1}{\sqrt{2}} (|00 \dots 0\rangle + e^{i\phi} |11 \dots 1\rangle), \quad (7.10)$$

this prescription yields a  $\cos^2(\phi/2)$  probability for the sum of all parities to be 0 mod 4.



**Figure 7.4.:** Circuit for preparing an extended cat state on a set of ancilla qubits with constant depth. The circuit need only be as local as the weight-four parity checks allow. Afterwards, any pair  $\{\gamma_{2j}, \gamma_{2j+1}\}$  of Majoranas may be used equivalently to perform a conditional Trotter step in QPE.

### 7.B. An algorithm to perform a Trotter step for a fully-connected fourth-order Hamiltonian in $O(N^3)$ time.

We showed in the main text a compact circuit for a four-Majorana Trotter step that does not require Jordan-Wigner strings, and in App. 7.A we suggested a method to perform conditional evolution in parallel by using a large GHZ state for an ancilla qubit. Assuming a Fermionic Hamiltonian on  $N$  spin-orbitals with 4th order terms, this would imply an  $O(N^3)$  circuit depth for our QPE algorithm per Trotter step. However, there is an additional complication; we need to ensure that we do not gain

## 7. Majorana-based fermionic quantum computation

additional circuit depth from the requirement to bring sets of 4 Majoranas close enough to perform this conditional evolution. To show this, we consider a line of  $N$  Majoranas  $\gamma_1, \dots, \gamma_N$ . We allow ourselves at each timestep  $t$  to swap a Majorana with its neighbour on the left or the right. (Note that this is a simplification from our architecture where we may not directly swap initialized Majoranas, but this brings only an additional constant time cost.) We wish to give an algorithm of length  $O(N^3)$  such that for any set of four Majoranas  $\{\gamma_i, \gamma_j, \gamma_k, \gamma_l\}$ , there exists a timestep  $t$  where these are placed consecutively along the line. As demonstrated in [265], inverting the line by a bubblesort solves the equivalent problem for pairs  $\{\gamma_i, \gamma_j\}$  in  $O(N)$  time, and this may be quickly extended to the case of sets of four. Let us consider the problem of forming all groups of 3 Majoranas. We divide our line into the sets  $\Gamma_0 = \{\gamma_i, i \leq N/2\}$ , and  $\Gamma_1 = \{\gamma_i, i > N/2\}$ . We then group neighboring pairs of elements in  $\Gamma_1$  to form subsets, which we then pair with all elements in  $\Gamma_0$  in  $O(N)$  time by a reverse bubblesort. Then, upon restoring to our previous position, we fix the position of elements of  $\Gamma_0$ , and perform a single iteration of the reverse bubblesort on the elements of  $\Gamma_1$  to form new subsets of pairs. Repeating this procedure until the second bubblesort has finished generates all subsets consisting of 2 Majoranas in  $\Gamma_1$  and 1 from  $\Gamma_0$  in  $O(N^2)$  time. All groups of 2 Majoranas from  $\Gamma_0$  and 1 from  $\Gamma_1$  may be given in the same manner. Then, we may split the line in 2, and reapply the above method on  $\Gamma_0$  and  $\Gamma_1$  separately to obtain all groups consisting of 3 Majoranas within. This final step takes  $O((N/2)^2 + (N/4)^2 + (N/8)^2 + \dots) = O(N^2)$  time. From here, it is clear how to proceed for groups of 4. We again divide our line into the sets  $\Gamma_0$  and  $\Gamma_1$ , and split our problem into that of making all groups of  $(m, 4 - m)$  Majoranas, where the first index denotes the number from  $\Gamma_0$  and the second from  $\Gamma_1$ . For  $1 \leq m \leq 3$ , we have an  $O(N^{m-1})$  circuit to prepare all groups of  $m$  Majoranas in  $\Gamma_0$ , a  $O(N^{3-m})$  circuit to prepare groups of  $4 - m$  Majoranas in  $\Gamma_1$ , and an  $O(N)$  bubblesort to pair all groups from  $\Gamma_0$  and  $\Gamma_1$ . These three steps must be looped within each other, giving a total time of  $O(N^{m-1}N^{3-m}N) = O(N^3)$ . Finally, we perform the  $m = 0$  and  $m = 4$  case simultaneously by repeating this procedure on the sets  $\Gamma_0$ , which takes again  $O(N^3)$  time by the arguments above.

## 7.C. Details of parallel circuit for Hubbard model

In this section we expand upon the proposal in Fig. 7.2 to perform QPE for the Hubbard model in constant time. This is a key feature of proposals for pre-error correcting quantum simulation [271], and as such bears further detail. There are 11 terms in equation 7.8 per site of our lattice, corresponding to 11 Trotter steps that must be performed in series (as each circuit piece requires accessing a prepared ancilla and additional Majoranas for the controlled braiding). As part of these Trotter steps, we must move Majoranas to their appropriate islands for parity measurements, and leave sufficient space for the preparation of the controlled rotation gate. We split the 11 Trotter steps into 3 stages, as indicated in Fig. 7.2(a). In the first stage, the Trotter steps corresponding to hopping terms between nearest neighbour fermions of the same spin are implemented, but only for those neighbours that are directly connected on the graph of Fig. 7.2(a) (i.e. those separated by a single braiding ancilla fermion). In the second stage, the steps for onsite two and four fermion interactions are implemented. From stage 2, as the qubits are being brought back to their resting position, the spin up and spin down fermions on each site have their locations exchanged. This allows for the final two Trotter steps to be applied between fermions that are now locally connected, without the large overhead of bringing distant fermions together and then apart. At the end of the unitary, the system is in a spin-rotated version of itself, and the order of Trotter steps for a second unitary evolution should be changed slightly to minimize braiding overhead. In Table 7.2, we detail these three stages further. In particular, we focus on the 10 terms involving the fermion  $f_{\uparrow}^{1,1}$ , and the onsite interaction term for the fermion  $f_{\downarrow}^{1,1}$ . For each term, we specify the location of all involved system Majoranas, parking spots for unused system Majoranas, the control ancilla, three braiding ancillas (for the implementation of the phase gate of Fig. 7.1), and which islands are involved in the parity measurement. Each such set of operations should then be tessellated across the lattice by a translation of a unit cell and a spin rotation to generate 10 parallelized Trotter steps for all fermions. (For example, the hopping steps involving  $f_{\downarrow}^{1,1}$  or  $f_{\uparrow}^{1,2}$  are implemented in the operations from neighboring cells, and the hopping steps of  $f_{\sigma}^{1,2}$  are reflected compared to those of  $f_{\sigma}^{1,1}$ , but those of  $f_{\sigma}^{2,1}$  are not). One should be careful then that this tessellation does not self-intersect, that all required qubits are connected to an island being measured, that the three braiding ancillas are connected in a way that allows for braiding, and

## 7. Majorana-based fermionic quantum computation

that the measurement circuit does not isolate individual islands (which would cause them to dephase). We assume that the conditional braidings on system Majoranas is performed as they move between configurations (or potentially cancelled), and so we do not account for these. We also assume that our finite-sized lattice is surrounded by a common ground, and so parallel lines of coupled islands will maintain a common phase by connecting to this. We have further found paths to hop Majoranas between their needed configurations and costed them in terms of the number of hoppings. We make no claim that the found arrangement is optimal, and invite any interested readers to attempt to beat our score for an optimal braiding pattern.



Stage	Hamiltonian term	System fermions	Parking sites	Control ancilla
1a	$\frac{it}{2} \gamma_{\uparrow,1}^{1,1} \gamma_{\uparrow,2}^{0,1}$	$f_{b,1}^{1,0}$	$f_{b,0}^{1,0}$	$f_{\uparrow}^{1,1}$
1b	$\frac{it}{2} \gamma_{\uparrow,1}^{0,1} \gamma_{\uparrow,2}^{1,1}$	$f_{b,0}^{1,0}$	$f_{b,1}^{1,0}$	$f_{\uparrow}^{1,1}$
1c	$\frac{it}{2} \gamma_{\uparrow,1}^{1,1} \gamma_{\uparrow,2}^{1,2}$	$f_{\uparrow}^{1,1}$	$f_{b,1}^{1,1}$	$f_c^{1,2}$
1d	$\frac{it}{2} \gamma_{\uparrow,1}^{1,2} \gamma_{\uparrow,2}^{1,1}$	$f_{b,1}^{1,1}$	$f_{\uparrow}^{1,1}$	$f_c^{1,2}$
1e	$\frac{it}{2} \gamma_{\uparrow,1}^{1,1} \gamma_{\uparrow,2}^{1,0}$	$f_{b,1}^{1,0}$	$f_{\uparrow}^{1,0}$	$f_c^{1,1}$
1f	$\frac{it}{2} \gamma_{\uparrow,1}^{1,0} \gamma_{\uparrow,2}^{1,1}$	$f_{\uparrow}^{1,0}$	$f_{b,1}^{1,0}$	$f_c^{1,1}$
2a	$\frac{i}{4} (U - 2\mu) \gamma_{\uparrow,1}^{1,1} \gamma_{\uparrow,2}^{1,1}$	$f_c^{1,1}$	$f_{b,2}^{1,1}$	$f_{\downarrow}^{1,1}$
2b	$\frac{i}{4} (U - 2\mu) \gamma_{\downarrow,1}^{1,1} \gamma_{\downarrow,2}^{1,1}$	$f_{b,2}^{1,1}$	$f_c^{1,1}$	$f_{\downarrow}^{1,1}$
2c	$-\frac{U}{4} \gamma_{\uparrow,1}^{1,1} \gamma_{\uparrow,2}^{1,1} \gamma_{\downarrow,1}^{1,1} \gamma_{\downarrow,2}^{1,1}$	$f_{b,2}^{1,1}, f_c^{1,1}$		$f_{\downarrow}^{1,1}$
3a	$\frac{it}{2} \gamma_{\uparrow,1}^{1,1} \gamma_{\uparrow,2}^{2,1}$	$f_{b,1}^{1,0}$	$f_{b,0}^{1,0}$	$f_{\uparrow}^{1,1}$
3b	$\frac{it}{2} \gamma_{\uparrow,1}^{2,1} \gamma_{\uparrow,2}^{1,1}$	$f_{b,0}^{2,0}$	$f_{b,2}^{1,0}$	$f_{\downarrow}^{1,1}$

**Table 7.2.:** (split with Table 7.3) A three-stage implementation of QPE on the Hubbard model (Eq. (7.8)), using the architecture in Fig. 7.2. We specify a translatable layout for each Trotter step to be performed simultaneously, by specifying which sites should be used to store system fermions, control ancilla fermions, braiding ancilla fermions (on Table 7.3), and any additional fermions not used in this rotation (parking sites). We further specify the island to be used for any joint parity readout (on Table 7.3), and the cost of shuffling the Majoranas around the layout (on Table 7.3).

## 7. Majorana-based fermionic quantum computation

Stage	Braiding ancillas	Measurement island	Cost
1a	$f_{\uparrow}^{0,1}, f_c^{0,1}, f_{b,2}^{0,0}$	$I_L^{1,1}$	(11)
1b	$f_{\uparrow}^{0,1}, f_c^{0,1}, f_{b,2}^{0,0}$	$I_L^{1,1}$	0
1c	$f_{b,2}^{1,1}, f_{b,0}^{2,1}, f_{\downarrow}^{1,1}$	$I_C^{1,1}$	11+7
1d	$f_{b,2}^{1,1}, f_{b,0}^{2,1}, f_{\downarrow}^{1,1}$	$I_C^{1,1}$	0
1e	$f_{b,2}^{1,0}, f_{b,0}^{2,0}, f_{\downarrow}^{1,0}$	$I_C^{1,0}$	7+7
1f	$f_{b,2}^{1,0}, f_{b,0}^{2,0}, f_{\downarrow}^{1,0}$	$I_C^{1,0}$	0
2a	$f_{b,1}^{1,1}, f_{b,0}^{1,1}, f_{\uparrow}^{1,1}$	$I_C^{1,1}$	7+6
2b	$f_{b,1}^{1,1}, f_{b,0}^{1,1}, f_{\uparrow}^{1,1}$	$I_C^{1,1}$	0
2c	$f_{b,1}^{1,1}, f_{b,0}^{1,1}, f_{\uparrow}^{1,1}$	$I_C^{1,1}$	0
3a	$f_{\uparrow}^{0,1}, f_c^{0,1}, f_{b,2}^{0,0}$	$I_L^{2,1}$	28+11
3b	$f_{\downarrow}^{2,1}, f_c^{2,1}, f_{b,1}^{2,0}$	$I_L^{2,1}$	0 (+11)

**Table 7.3.:** (Split with Table 7.2) A three-stage implementation of QPE on the Hubbard model (Eq. (7.8)), using the architecture in Fig. 7.2. Description of the Hamiltonian term to be simulated, and the use of each Majorana given in Table 7.2. For each Trotter step we have costed the number of Majorana hoppings required to rearrange the system from its previous state. When these are written as a sum, the first term refers to restoring the configuration of Fig. 7.2 from the configuration required for the previous step, and the second to obtaining the configuration needed for the current step. Some steps require the same configuration as the previous step, and as such incur a 0 rearrangement cost. The cost in brackets for the final step is the requirement to return the system to its shifted initial state (where up-spins and down-spins have been swapped). This may not be required, especially as the configuration for the final step and the initial steps are the same (modulo the swapping of the spins), and so repeated unitary evolution would not need this nor the rearrangement cost of the first step. This reduces the total rearrangement cost of the circuit to 85 Majorana hoppings.

## 8. Quantum phase estimation for noisy, small-scale experiments

### 8.1. Introduction

It is known that any problem efficiently solvable on a quantum computer can be formulated as eigenvalue sampling of a Hamiltonian or eigenvalue sampling of a sparse unitary matrix [292]. In this sense the algorithm of quantum phase estimation is the only quantum algorithm which can give rise to solving problems with an exponential quantum speed-up. Despite it being such a central component of many quantum algorithms, very little work has been done so far to understand what quantum phase estimation offers in the current NISQ (Noisy Intermediate Scale Quantum) era of quantum computing [293] where quantum devices are strongly coherence-limited. Quantum phase estimation comes in many variants, but a large subclass of these algorithms (e.g. the semi-classical version of textbook phase estimation [23, 294], Kitaev’s phase estimation [22], Heisenberg-optimized versions [295]), are executed in an iterative sequential form using controlled- $U^k$  gates with a single ancilla qubit [296, 297] (see Fig. 8.1), or by direct measurement of the system register itself [295]. Such circuits are of practical interest in the near term when every additional qubit requires a larger chip and brings in additional experimental complexity and incoherence.

Some of the current literature on quantum phase estimation works under limiting assumptions. The first is that one does not start in an eigenstate of the Hamiltonian [298, 299]. A second limitation is that one does not take into account the (high) temporal cost of running  $U^k$  [297] for large  $k$  when optimizing phase estimation. The size and shallowness of the

---

The contents of this chapter has been accepted for publication as T. E. O’Brien, B. Tarasinski and B. M. Terhal, New J. Phys. 21, 023022 (2019).

## 8. Quantum phase estimation for noisy, small-scale experiments

quantum phase estimation circuit is important since, in the absence of error correction or error mitigation, one expects entropy build-up during computation. This means that circuits with large  $k$  may not be of any practical interest.

The scenario where the input state is not an eigenstate of the unitary matrix used in phase estimation is the most interesting one from the perspective of applications, and we will consider it in this chapter. Such an input state can be gradually projected onto an eigenstate by the phase estimation algorithm and the corresponding eigenvalue can be inferred. However, for coherence-limited low-depth circuits one may not be able to evolve sufficiently long to project well onto one of the eigenstates. This poses the question what one can still learn about eigenvalues using low-depth circuits. An important point is that it is experimentally feasible to repeat many relatively shallow experiments (or perform them in parallel on different machines). Hence we ask what the spectral-resolving power of such phase estimation circuits is, both in terms of the number of applications of the controlled- $U$  circuit in a single experiment, and the number of times the experiment is repeated. Such repeated phase estimation experiments require classical post-processing of measurement outcomes, and we study two such algorithms for doing this. One is our adaptation of the Bayesian estimator of [299] to the multiple-eigenvalue scenario. A second is a new estimator based on a treatment of the observed measurements as a time-series, and construction of the resultant time-shift operator. This latter method is very natural for phase estimation, as one interprets the goal of phase estimation as the reconstruction of frequencies present in the output of a temporal sound signal. In fact, the time-series analysis that we develop is directly related to what are called Prony-like methods in the signal-processing literature, see e.g. [300]. The use of this classical method in quantum signal processing, including in quantum tomography [301], seems to hold great promise.

One can interpret our results as presenting a new hybrid classical-quantum algorithm for quantum phase estimation. Namely, when the number of eigenstates in an input state is small, i.e. scaling polynomially with the number of qubits  $N_{\text{sys}}$ , the use of our classical post-processing method shows that there is no need to run a quantum algorithm which projects onto an eigenstate to learn the eigenvalues. We show that one can extract these eigenvalues efficiently by classically post-processing the data from experiments using a single-round quantum phase estimation circuits (see Section 8.2) and classically handling  $\text{poly}(N_{\text{sys}}) \times \text{poly}(N_{\text{sys}})$  matrices. This constitutes a saving in the required depth of the quantum circuits.

The spectral-resolution power of quantum phase estimation can be

defined by its scaling with parameters of the experiment and the studied system. We are able to derive analytic scaling laws for the problem of estimating single eigenvalues with the time-series estimator. We find these to agree with the numerically-observed scaling of both studied estimators. For the more general situation, with multiple eigenvalues and experimental error, we study the error in estimating the lowest eigenvalue numerically. This is assisted by the low classical computation cost of both estimators. We observe scaling laws for this error in terms of the overlap between the ground and starting state (i.e. the input state of the circuit), the gap between the ground and excited states, and the coherence length of the system. In the presence of experimental noise we attempt to adjust our estimators to mitigate the induced estimation error. For depolarizing-type noise we find such compensation easy to come by, whilst for a realistic circuit-level simulation we find smaller improvements using similar techniques.

Even though this chapter focuses on quantum phase estimation where the phases corresponds to eigenvalues of a unitary matrix, our post-processing techniques may also be applicable to multi-parameter estimation problems in quantum optical settings. In these settings the focus is on determining an optical phase-shift [302–304] through an interferometric set-up. There is experimental work on (silicon) quantum photonic processors [305–307] on multiple-eigenvalue estimation for Hamiltonians which could also benefit from using the classical post-processing techniques that we develop in this chapter.

## 8.2. Quantum phase estimation

Quantum phase estimation (QPE) covers a family of quantum algorithms which measure a system register of  $N_{\text{sys}}$  qubits in the eigenbasis of a unitary operator  $U$  [22, 308]

$$U|\phi_j\rangle = e^{i\phi_j}|\phi_j\rangle, \quad (8.1)$$

to estimate one or many phases  $\phi_j$ . Quantum phase estimation algorithms assume access to a noise-free quantum circuit which implements  $U$  on our system register conditioned on the state of an ancilla qubit. Explicitly, we require the ability to implement

$$\mathcal{U}_c = |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes U, \quad (8.2)$$

## 8. Quantum phase estimation for noisy, small-scale experiments

where  $|0\rangle$  and  $|1\rangle$  are the computational basis states of the ancilla qubit, and  $\mathbb{I}$  is the identity operator on the system register.

In many problems in condensed matter physics, materials science, or computational chemistry, the object of interest is the estimation of spectral properties or the lowest eigenvalue of a Hamiltonian  $\mathcal{H}$ . The eigenvalue estimation problem for  $\mathcal{H}$  can be mapped to phase estimation for a unitary  $U_\tau = \exp(-i\tau\mathcal{H})$  with a  $\tau$  chosen such that the relevant part of the eigenvalue spectrum induces phases within  $[-\pi, \pi)$ . Much work has been devoted to determining the most efficient implementation of the (controlled)- $\exp(-i\tau\mathcal{H})$  operation, using exact or approximate methods [308–311]. Alternatively, one may simulate  $\mathcal{H}$  via a quantum walk, mapping the problem to phase estimating the unitary  $\exp(-i \arcsin(\mathcal{H})/\lambda)$  for some  $\lambda$ , which may be implemented exactly [289–291, 312]. In this chapter we do not consider such variations, but rather focus on the error in estimating the eigenvalue phases of the unitary  $U$  that is actually implemented on the quantum computer. In particular, we focus on the problem of determining the value of a single phase  $\phi_0$  to high precision (this phase could correspond, for example, to the ground state energy of some Hamiltonian  $\mathcal{H}$ ).

Phase estimation requires the ability to prepare an input, or *starting state*

$$|\Psi\rangle = \sum_j a_j |\phi_j\rangle, A_j \equiv |a_j|^2, \quad (8.3)$$

with good overlap with the ground state;  $A_0 \gg 0$ . Note here that the spectrum of  $U$  may have exact degeneracies (e.g. those enforced by symmetry) which phase estimation does not distinguish; we count degenerate eigenvalues as a single  $\phi_j$  throughout this chapter. The ability to start quantum phase estimation in a state which already has good overlap with the ground state is a nontrivial requirement for the applicability of the quantum phase estimation algorithm. On the other hand, it is a well-known necessity given the QMA-completeness [313] of the lowest eigenvalue problem\*. For many quantum chemistry and materials science problems it is known or expected that the Hartree-Fock state has good overlap with the ground state, although rigorous results beyond perturbation theory

---

\*QMA stands for Quantum Merlin Arthur, which is a complexity class which contains decision problems which are easy to verify on a quantum computer, though not necessarily easy to solve. This class is the natural quantum counterpart to the complexity class NP of problems that may be verified easily on a classical computer. A QMA-complete problem is one of the ‘hardest possible’ such problems (in analogy with NP-complete problems); the ability to solve these problems in polynomial time would allow polynomial-time solving of any other problem in QMA.

are far and few between (see e.g. [314]). Beyond this, either adiabatic evolution [263, 309] or variational quantum eigensolvers [281] can provide an approximate starting state to improve on via phase estimation.

Phase estimation is not limited to simply learning the value of  $\phi_0$ ; it may obtain information about all phases  $\phi_j$  as long as  $A_j > 0$ . However, the resources required to estimate  $\phi_j$  are bounded below by  $1/A_j$ . To see this, note that the controlled-unitary  $\mathcal{U}_c$  does not mix eigenstates, and so there is no difference (in the absence of error) between starting with  $|\Psi\rangle$  and the mixed state

$$\rho_\Psi = \sum_j A_j |\phi_j\rangle\langle\phi_j|. \quad (8.4)$$

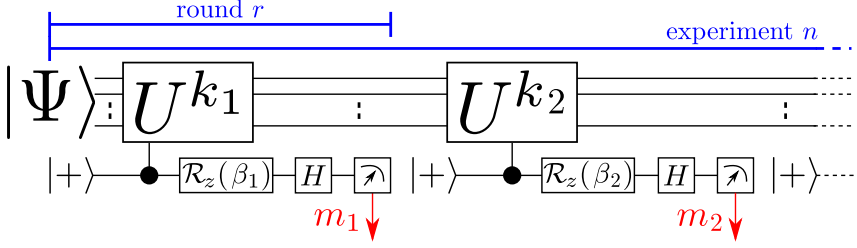
The latter is then equivalent to preparing the pure state  $|\phi_j\rangle$  with probability  $A_j$ , so if  $N$  preparations of  $|\phi_j\rangle$  are required to estimate  $\phi_j$  to an error  $\epsilon$ , the same error margin requires at least  $N/A_j$  preparations of the state  $|\Psi\rangle$ . As the number of eigenstates  $N_{\text{eig}}$  with non-zero contribution to  $|\Psi\rangle$  generally scales exponentially with the system size  $N_{\text{sys}}$ , estimating more than the first few  $\phi_j$  (ordered by the magnitude  $A_j$ ) will be unfeasible.

Low-cost (in terms of number of qubits) quantum phase estimation may be performed by entangling the system register with a single ancilla qubit [22, 297, 299, 313]. In Fig. 8.1, we give the general form of the quantum circuit to be used throughout this chapter. An experiment, labeled by a number  $n = 1, \dots, N$ , can be split into one or multiple rounds  $r = 1, \dots, R_n$ , following the preparation of the starting state  $|\Psi\rangle$ . In each round a single ancilla qubit prepared in the  $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$  state controls  $\mathcal{U}_c^{k_r}$  where the integer  $k_r$  can vary per round. The ancilla qubit is then rotated by  $\mathcal{R}_z(\beta_r) = \exp(-i\beta_r Z/2)$  (with the phase  $\beta_r$  possibly depending on other rounds in the same experiment) and read out in the  $X$ -basis, returning a measurement outcome  $m_r \in \{0, 1\}$ . We denote the chosen strings of integers and phases of a single multi-round experiment by  $\mathbf{k}$  and  $\boldsymbol{\beta}$  respectively. We denote the number of controlled- $U$  iterations per experiment as  $K = \sum_{r=1}^{R_n} k_r$ . We denote the total number of controlled- $U$  iterations over all experiments as

$$K_{\text{tot}} = \sum_{n=1}^N \sum_{r=1}^{R_n} k_r. \quad (8.5)$$

As the system register is held in memory during the entire time of the experiment, the choice of  $K$  is dictated by the coherence time of the underlying quantum hardware. Hence, we introduce a dimensionless

## 8. Quantum phase estimation for noisy, small-scale experiments



**Figure 8.1.:** Circuit for the QPE experiments described in this chapter. The state  $|\Psi\rangle$  is defined in Eq. (8.3). The probability for the ancilla qubit to return the vector  $\mathbf{m}$  of results in the absence of error is given by Eq. (8.10). The single-qubit rotation equals  $\mathcal{R}_z(\beta) = \exp(-i\beta Z/2)$  while  $H$  is the Hadamard gate.

coherence length

$$K_{\text{err}} = \frac{T_{\text{err}}}{N_{\text{sys}} T_U}. \quad (8.6)$$

Here  $T_U$  is the time required to implement a single application of controlled- $U$  in Eq. (8.7), and  $T_{\text{err}}$  is the time-to-error of a single qubit, so that  $T_{\text{err}}/N_{\text{sys}}$  is the time-to-failure of  $N_{\text{sys}}$  qubits. The idea is that  $K_{\text{err}}$  bounds the maximal number of applications of  $U$  in an experiment, namely  $K \leq K_{\text{err}}$ .

A new experiment starts with the same starting state  $|\Psi\rangle$ . Values of  $k_r$  and  $\beta_r$  may be chosen independently for separate experiments  $n$ , i.e. we drop the label  $n$  for convenience. We further drop the subscript  $r$  from single-round experiments (with  $R = 1$ ).

In the absence of error, one may calculate the action of the QPE circuit on the starting state (defined in Eq. (8.3)). Working in the eigenbasis of  $U$  on the system register, and the computational basis on the ancilla qubit, we calculate the state following the controlled-rotation  $\mathcal{U}_c^{k_1}$ , and the rotation  $\mathcal{R}_z(\beta_1)$  on the ancilla qubit to be

$$\frac{1}{\sqrt{2}} \sum_j a_j \left( |0\rangle + e^{i(k_1 \phi_j + \beta_1)} |1\rangle \right) |\phi_j\rangle. \quad (8.7)$$

The probability to measure the ancilla qubit in the  $X$ -basis as  $m_1 \in \{0, 1\}$



is then

$$\sum_j A_j \cos^2 \left( \frac{k_1 \phi_j}{2} + \frac{\beta_1 - m_1 \pi}{2} \right), \quad (8.8)$$

and the unnormalized post-selected state of the system register is

$$\sum_j a_j e^{\frac{i}{2}(k_1 \phi_j + \beta_1)} \cos \left( \frac{k_1 \phi_j}{2} + \frac{\beta_1 - m_1 \pi}{2} \right) |\phi_j\rangle. \quad (8.9)$$

The above procedure may then be repeated for  $r$  rounds to obtain the probability of a string  $\mathbf{m}$  of measurement outcomes of one experiment as

$$P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A}) = \sum_j A_j \prod_{r=1}^R \cos^2 \left( \frac{k_r \phi_j}{2} + \frac{\beta_r - m_r \pi}{2} \right). \quad (8.10)$$

Here,  $\phi$  is the vector of phases  $\phi_j$  and  $\mathbf{A}$  the vector of probabilities for different eigenstates. We note that  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$  is independent of the order in which the rounds occur in the experiment. Furthermore, when  $N_{\text{eig}} = 1$ ,  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi) = P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$  is equal to the product of the single-round probabilities  $P_{k_r,\beta_r}(m_r|\phi)$ , as there is no difference between a multi-round experiment and the same rounds repeated across individual experiments.

One can make a direct connection with parameter estimation work by considering the single-round experiment scenario in Fig. 8.1. The Hadamard gate putting the ancilla qubit in  $|+\rangle$  and measuring the qubit in the  $X$ -basis are, in the optical setting, realized by beam-splitters, so that only the path denoted by the state  $|1\rangle$  will pick up an unknown phase-shift. When the induced phase-shift is not unique but depends, say, on the state of another quantum system, we may like to estimate all such possible phases corresponding to our scenario of wishing to estimate multiple eigenvalues. Another physical example is a dispersively coupled qubit-cavity mode system where the cavity mode occupation number will determine the phase accumulation of the coupled qubit [315].

### 8.3. Classical data analysis

Two challenges are present in determining  $\phi_0$  from QPE experiments. First, we only ever have inexact sampling knowledge of  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$ . That is, repeated experiments at fixed  $\mathbf{k}, \beta$  do not directly determine  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$ ,

## 8. Quantum phase estimation for noisy, small-scale experiments

but rather sample from the multinomial distribution  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$ . From the measurement outcomes we can try to estimate  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$  (and from this  $\phi_0$ ) as a hidden variable. Secondly, when  $N_{\text{eig}} > 1$  determining  $\phi_0$  from  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$  poses a non-trivial problem.

Let us first consider the case  $N_{\text{eig}} = 1$ . Let us assume that we do single-round experiments with a fixed  $k$  for each experiment. Naturally, taking  $k = 1$  would give rise to the lowest-depth experiments. If we start these experiments with  $k = 1$  in the eigenstate  $|\phi_0\rangle$ , then one can easily prove that taking  $\beta = 0$  or  $\beta = \frac{\pi}{2}$  for half of the experiments, suffices to estimate  $\phi_0$  with variance scaling as  $\sim 1/N = 1/K_{\text{tot}}$ . This result can be derived using standard Chernoff bounds, see e.g. [316, 317], and represent standard sampling or shot noise behavior. When  $N_{\text{eig}} = 1$ ,  $N$   $K$ -round experiments each with  $k = 1$  are indistinguishable from  $N \times K$  single-round experiments with  $k = 1$ . This implies that the same scaling holds for such multi-round experiments, i.e. the variance scales as  $1/(NK) = 1/K_{\text{tot}}$ .

Once the phase  $\phi_0$  is known to sufficient accuracy, performing QPE experiments with  $k > 1$  is instrumental in resolving  $\phi_0$  in more detail, since the probability of a single-round outcome depends on  $k\phi_0$  [295]. Once one knows with sufficient certainty that  $\phi_0 \in [(2m-1)\pi/k, (2m+1)\pi/k]$  (for integer  $m$ ), one can achieve variance scaling as  $O(\frac{1}{k^2N})$  (conforming to so-called local estimation Cramer-Rao bounds suggested in [299, 318]). A method achieving Heisenberg scaling, where the variance scales as  $\sim 1/K_{\text{tot}}^2$  (see Eq. (8.5)), was analyzed in [295, 316]. This QPE method can also be compared with the information-theoretic optimal maximum-likelihood phase estimation method of [297] where  $N \sim \log K$  experiments are performed, each choosing a random  $k \in \{1, \dots, K\}$  to resolve  $\phi_0$  with error  $\sim 1/K$ . The upshot of these previous results is that, while the variance scaling in terms of the total number of unitaries goes like  $1/K_{\text{tot}}$  when using  $k = 1$ , clever usage of  $k > 1$  data can lead to  $1/K_{\text{tot}}^2$  scaling. However, as  $K$  is limited by  $K_{\text{err}}$  in near-term experiments, this optimal Heisenberg scaling may not be accessible.

When  $N_{\text{eig}} > 1$ , the above challenge is complicated by the need to resolve the phase  $\phi_0$  from the other  $\phi_j$ . This is analogous to the problem of resolving a single note from a chord. Repeated single-round experiments at fixed  $k$  and varying  $\beta$  can only give information about the value of the function:

$$g(k) = \sum_j A_j e^{ik\phi_j}, \quad (8.11)$$

at this fixed  $k$ , since

$$\begin{aligned} P_{k,\beta}(m|\phi) &= \frac{1}{2} + \frac{1}{2} \cos(\beta + m\pi) \operatorname{Re}[g(k)] \\ &\quad - \frac{1}{2} \sin(\beta + m\pi) \operatorname{Im}[g(k)]. \end{aligned} \quad (8.12)$$

This implies that information from single-round experiments at fixed  $k$  is insufficient to resolve  $\phi_0$  when  $N_{\text{eig}} > 1$ , as  $g(k)$  is then not an invertible function of  $\phi_0$  (Try to recover a frequency from a sound signal at a single point in time!). In general, for multi-round experiments using a maximum of  $K$  total applications of  $\mathcal{U}_c$ , we may only ever recover  $g(k)$  for  $k \leq K$ . This can be seen from expanding  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$  as a sum of  $\sum_j A_j \cos^m(\phi_j) \sin^n(\phi_j)$  terms with  $m + n \leq K$ , which are in turn linear combinations of  $g(k)$  for  $k \leq K$ . As we will show explicitly in the next Section 8.3.1 this allows us to recover up to  $K$   $\phi_j$ . However, when  $N_{\text{eig}} > K$ , these arguments imply that we cannot recover any phases exactly. In this case, the accuracy to which we can estimate our target  $\phi_0$  is determined by the magnitude of the amplitude  $A_0$  in the initial state  $|\Psi\rangle$  as well as the gap towards the other eigenvalues. For example, in the limit  $A_0 \rightarrow 1$ , an unbiased estimation of  $\phi_0$  using data from  $k = 1$  would be

$$\operatorname{Arg}[g(1)] = \operatorname{Im}[\ln(\sum_j A_j e^{i\phi_j})], \quad (8.13)$$

and the error in such estimation is

$$\begin{aligned} |\operatorname{Arg}[g(1)] - \phi_0| &= \left| \frac{1}{A_0} \sum_{j=1}^{N_{\text{eig}}-1} A_j \sin(\phi_j - \phi_0) + O(A_0^{-2}) \right| \\ &\leq \frac{1 - A_0}{A_0}, \end{aligned}$$

with our bound being independent of  $N_{\text{eig}}$ . We are unable to extend this analysis beyond the  $k = 1$  scenario, and instead we study the scaling in this estimation numerically in Sec. 8.4. In the remainder of this section, we present two estimators for multi-round QPE. The first is an estimator based on a time-series analysis of the function  $g(k)$  using Prony-like [300] methods that has a low computation overhead. The second is a Bayesian estimator similar to that of [299], but adapted for multiple eigenphases  $\phi_j$ .

### 8.3.1. Time-series analysis

Let us assume that the function  $g(k)$  in Eq. (8.11) is a well-estimated function at all points  $0 \leq k \leq K$ , since the number of experiments  $N$  is sufficiently large. We may extend this function to all points  $-K \leq k \leq K$  using the identity  $g(-k) = g^*(k)$  to obtain a longer signal  $*$ . We wish to determine the dominant frequencies  $\phi_j$  in the signal  $g(k)$  as a function of ‘time’  $k$ . This can be done by constructing and diagonalizing a time-shift matrix  $\mathfrak{T}$  whose eigenvalues are the relevant frequencies in the signal, as follows.

We first demonstrate the existence of the time-shift matrix  $\mathfrak{T}$  in the presence of  $N_{\text{eig}} < K$  separate frequencies. Since we may not know  $N_{\text{eig}}$ , let us first estimate it as  $l$ . We then define the vectors  $\mathbf{g}(k) = (g(k), g(k+1), \dots, g(k+l))^T$ ,  $k = -K, \dots, K$ . These vectors can be decomposed in terms of single-frequency vectors  $\mathbf{b}_j = (1, e^{i\phi_j}, \dots, e^{il\phi_j})^T$

$$\mathbf{g}(k) = \sum_j A_j e^{ik\phi_j} \mathbf{b}_j. \quad (8.14)$$

We can make a  $l \times N_{\text{eig}}$  matrix  $B$  with the components  $\mathbf{b}_j$  as columns

$$B_{k,j} = e^{ik\phi_j}. \quad (8.15)$$

When  $N_{\text{eig}} \leq l$ , the columns of  $B$  are typically linearly independent <sup>†</sup>, hence the non-square matrix  $B$  is invertible and has a (left)-pseudoinverse  $B^{-1}$  such that  $B^{-1}B = \mathbf{1}$ . Note however, when  $N_{\text{eig}} > l$  the columns of  $B$  are linearly-dependent, so  $B$  cannot be inverted. If  $B$  is invertible, we can construct the shift matrix  $\mathfrak{T} = BDB^{-1}$  with  $D_{i,j} = \delta_{i,j}e^{i\phi_j}$ . By construction,  $\mathfrak{T}\mathbf{b}_j = e^{i\phi_j}\mathbf{b}_j$  (as  $\mathfrak{T}B = BD$ ), and thus

$$\begin{aligned} \mathfrak{T}\mathbf{g}(k) &= \sum_j A_j e^{ik\phi_j} \mathfrak{T}\mathbf{b}_j \\ &= \sum_j A_j e^{i(k+1)\phi_j} = \mathbf{g}(k+1). \end{aligned} \quad (8.16)$$

This implies that  $\mathfrak{T}$  acts as the time-shift operator mapping  $\mathbf{g}(k)$  to  $\mathbf{g}(k+1)$ . As the eigenvalues of  $\mathfrak{T}$  are precisely the required phases  $e^{i\phi_j}$

---

\*Extending  $g(k)$  from  $0 \leq k \leq K$  to  $-K \leq k \leq K$  is not required to perform a time-series analysis, however numerically we observe that this obtains up to order of magnitude improvement in estimating  $\phi_0$ .

<sup>†</sup>Counterexamples may exist, but are hard to construct and have not occurred in any numerics.

in case  $N_{\text{eig}} \leq l$ , constructing and diagonalizing  $\mathfrak{T}$  will obtain our desired phases including  $\phi_0$ . When  $N_{\text{eig}} > l$ , the eigen-equation for  $\mathfrak{T}$  cannot have the solution  $\mathbf{b}_j$  since these are not linearly independent.

The above proof of existence does not give a method of constructing the time-shift operator  $\mathfrak{T}$ , as we do not have access to the matrices  $B$  or  $D$ . To construct  $\mathfrak{T}$  from the data that we do have access to, we construct the  $l \times (2K + 1 - l)$  Hankel matrices  $G^{(0)}, G^{(1)}$  by

$$G_{i,j}^{(a)} = g(i + j + a - K), \quad (8.17)$$

indexing  $0 \leq i \leq l - 1, 0 \leq j \leq 2K - l$ . The  $k$ -th column of  $G^{(a)}$  is the vector  $\mathbf{g}(k + a - K)$ , and so  $\mathfrak{T}G^{(0)} = G^{(1)}$ . We can thus attempt to find  $\mathfrak{T}$  as a solution of the (least-squares) problem of minimizing  $\|\mathfrak{T}G^{(0)} - G^{(1)}\|$ . The rank of the obtained  $\mathfrak{T}$  is bounded by the rank of  $G^{(0)}$ . We have that  $\text{rank}(G^{(0)})$  is at most  $N_{\text{eig}}$  since it is a sum over rank-1 matrices. At the same time  $\text{rank}(G^{(0)}) \leq \min(l, 2K + 1 - l)$ . This implies that we require both  $l \geq N_{\text{eig}}$  and  $2K + 1 - l \geq N_{\text{eig}}$  to obtain a shift matrix  $\mathfrak{T}$  with  $N_{\text{eig}}$  eigenvalues. This is only possible when  $K \geq N_{\text{eig}}$ , giving an upper bound for the number of frequencies obtainable. When  $G^{(0)}$  is not full rank (because  $N_{\text{eig}} < l$ ), this problem may have multiple zeros  $\tilde{\mathfrak{T}}$ . However, when  $N_{\text{eig}} < l$  each of these must satisfy  $\tilde{\mathfrak{T}}\mathbf{g}(k) = \mathbf{g}(k + 1)$  for  $-K < k < K - l$ .

Then, as long as  $\text{rank}(G^{(0)}) \geq N_{\text{eig}}$ , Eq. (8.14) is invertible by an operator  $C$

$$\sum_k C_{i,k} A_j e^{ik\phi_j} = \delta_{i,j} \rightarrow \mathbf{b}_j = \sum_k C_{j,k} \mathbf{g}(k). \quad (8.18)$$

It follows that

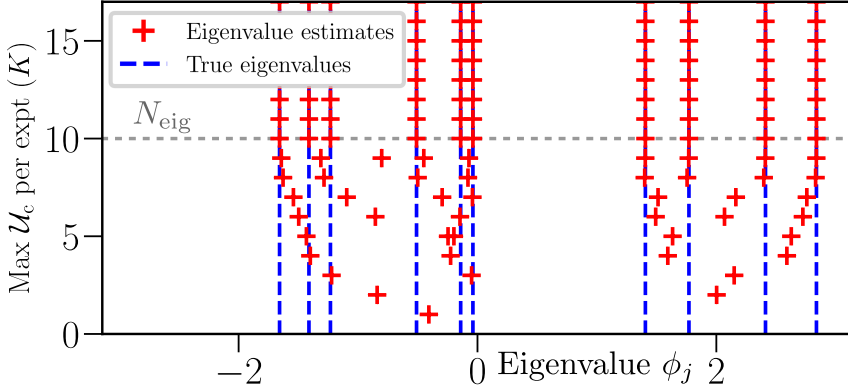
$$\sum_k C_{j,k} \mathbf{g}(k + 1) = \sum_{k,l} C_{j,k} A_l e^{ik\phi_l} (e^{i\phi_l} \mathbf{b}_l) = e^{i\phi_j} \mathbf{b}_j, \quad (8.19)$$

and then

$$\tilde{\mathfrak{T}}\mathbf{b}_j = \sum_k C_{k,j} \tilde{\mathfrak{T}}\mathbf{g}(k) = \sum_k C_{k,j} \mathbf{g}(k + 1) = e^{i\phi_j} \mathbf{b}_j, \quad (8.20)$$

so every  $\tilde{\mathfrak{T}}$  obtained in this way must have eigenvalues  $e^{i\phi_j}$ .

The above analysis is completely independent of the coefficients  $A_j$ . However, once the eigenvalues  $\phi_j$  are known, the matrix  $B$  (eq. 8.15) may be constructed, and the  $A_j$  may be recovered by a subsequent least-squares



**Figure 8.2.:** Convergence of the time-series estimator in the estimation of  $N_{\text{eig}} = 10$  eigenvalues (chosen at random with equally sized amplitudes  $A_j = 1/10$ ) when the exact function  $g(k)$  is known at points  $0, \dots, K$ . The estimator constructs and calculates the eigenvalues of  $K \times K$  matrix which are shown as the red plusses in the Figure. When  $K \geq N_{\text{eig}}$  (gray dashed line), the frequencies are attained to within machine precision. When  $K < N_{\text{eig}}$ , it is clear from the Figure that the found eigenvalues provide some form of binning approximation of the spectrum.

minimization of

$$\|B\mathbf{A} - \mathbf{g}(0)\|. \quad (8.21)$$

This allows us to identify spurious eigenvalues if  $l > N_{\text{eig}}$  (as these will have a corresponding zero amplitude). Numerically, we find no disadvantage to then choosing the largest  $l$  permitted by our data, namely  $l = K$ .

Assuming a sufficient number of repetitions  $N$  these arguments imply that this strategy requires that  $K \geq N_{\text{eig}}$  to determine all eigenvalues accurately. However, when  $K < N_{\text{eig}}$  there still exists a least-squares solution  $\tilde{\mathfrak{T}}$  that minimizes  $\|\tilde{\mathfrak{T}}G^{(0)} - G^{(1)}\|$ . When  $A_0 \gg 0$ , we expect that  $\tilde{\mathfrak{T}}$  should have eigenvalues  $e^{i\tilde{\phi}_0} \approx e^{i\phi_0}$  that we can take as the estimator for  $\phi_0$ ; the same is true for any other  $\phi_j$  with sufficiently large  $A_j$ . In Fig. 8.2 we show an example of convergence of this estimation for multiple eigenvalues  $\phi_j$  as  $K \rightarrow N_{\text{eig}}$  in the case where  $g(k)$  is known precisely (i.e. in the absence of sampling noise). The error  $|\tilde{\phi}_0 - \phi_0|$  when  $K < N_{\text{eig}}$  depends on the eigenvalue gap above  $\phi_0$ , as well as the relative weights  $A_j$ , as we will see in Section 8.4.3.

In App.8.B we derive what variance can be obtained with this time-series method in the case  $l = N_{\text{eig}} = 1$ , using single-round circuits with  $k = 1$  up

to  $K$ . Our analysis leads to the following scaling in  $N$  and  $K$ :

$$\text{Var}(\phi) \propto \frac{1}{K^2 N}. \quad (8.22)$$

We will compare these results to numerical simulations in Sec. 8.4.1.

### Estimating $g(k)$

The function  $g(k)$  cannot be estimated directly from experiments, but may instead be created as a linear combination of  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$  for different values of  $\mathbf{k}$  and  $\beta$ . For single-round experiments, this combination is simple to construct:

$$\begin{aligned} g(k) = & P_{k,0}(0|\phi, \mathbf{A}) - P_{k,0}(1|\phi, \mathbf{A}) \\ & - iP_{k,\frac{\pi}{2}}(0|\phi, \mathbf{A}) + iP_{k,\frac{\pi}{2}}(1|\phi, \mathbf{A}). \end{aligned} \quad (8.23)$$

For multi-round experiments, the combination is more complicated. In general,  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$  is a linear combination of real and imaginary parts of  $g(l)$  with  $l < K = \sum_r k_r$ . This combination may be constructed by writing  $\cos^2(k\phi_j/2 + \beta/2)$  and  $\sin^2(k\phi_j/2 + \beta/2)$  in terms of exponentials, and expanding. However, inverting this linear equation is a difficult task and subject to numerical imprecision. For some fixed choices of experiments, it is possible to provide an explicit expansion. Here we focus on  $K$ -round  $k = 1$  experiments with  $K/2$   $\beta = 0$  and  $K/2$   $\beta = \frac{\pi}{2}$  final rotations during each experiment (choosing  $K$  even). The formula for  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$  is independent of the order in which these rounds occur. Let us write  $\mathbb{P}(\mathbf{m}, \mathbf{n}|\phi, \mathbf{A})$  as the probability of seeing both  $\mathbf{m} \in \{0, \dots, K/2\}$  outcomes with  $m_r = 1$  in the  $K/2$  rounds with  $\beta_r = 0$  and  $\mathbf{n} \in \{0, \dots, K/2\}$  outcomes with  $n_r = 1$  in the  $K/2$  rounds with  $\beta_r = \pi/2$ . In other words,  $\mathbf{m}, \mathbf{n}$  are the Hamming weights of the measurement vectors split into the two types of rounds described above. Then, one can prove that, for  $0 \leq k \leq K/2$ :

$$g(k) = \sum_{m=0}^{K/2} \sum_{n=0}^{K/2} \chi_k(\mathbf{m}, \mathbf{n}) \mathbb{P}(\mathbf{m}, \mathbf{n}|\phi, \mathbf{A}) \quad (8.24)$$

## 8. Quantum phase estimation for noisy, small-scale experiments

where

$$\begin{aligned} \chi_k(\mathbf{m}, \mathbf{n}) &= \sum_{l=0}^k (-i)^{k-l} \binom{k}{l} \\ &\times \left[ \sum_{p_1=0}^{\lfloor l/2 \rfloor} \frac{\binom{m}{2p_1} \binom{K/2-m}{l-2p_1}}{\binom{K/2}{l}} - 1 \right] \\ &\times \left[ \sum_{p_2=0}^{\lfloor (k-l)/2 \rfloor} \frac{\binom{n}{2p_2} \binom{K/2-n}{k-l-2p_2}}{\binom{K/2}{k-l}} - 1 \right]. \end{aligned} \quad (8.25)$$

The proof of this equality can be found in App.8.A.

Calculating  $g(k)$  from multi-round ( $k = 1$ ) experiments contains an additional cost: combinatorial factors in Eq. (8.24) relate the variance in  $g(k)$  to the variance in  $\mathbb{P}(\mathbf{m}, \mathbf{n} | \phi, \mathbf{A})$  but the combinatorial pre-factor  $\binom{k}{l}$  can increase exponentially in  $k$ . This can be accounted for by replacing the least squares fit used above with a weighted least squares fit, so that one effectively relies less on the correctness of  $g(k)$  for large  $k$ . To do this, we construct the matrix  $\mathfrak{T}$  row-wise from the rows  $\mathbf{g}_i^{(1)}$  of  $G^{(1)}$ . That is, for the  $i$ th row  $\mathbf{t}_i$  we minimize

$$\|\mathbf{t}_i G^{(0)} - \mathbf{g}_i^{(1)}\|. \quad (8.26)$$

This equation may be weighted by multiplying  $G^{(0)}$  and  $\mathbf{g}_i^{(1)}$  by the weight matrix

$$w_{j,k}^{(i)} = \delta_{j,k} \frac{1}{\sigma_{G_{i,j}^{(1)}}}, \quad (8.27)$$

where  $\sigma_{G_{i,j}^{(1)}}$  is the standard deviation in our estimate of  $G_{i,j}^{(1)}$ . Note that the method of weighted least-squares is only designed to account for error in the independent variable of a least squares fit, in our case this is  $G^{(1)}$ . This enhanced effect of the sampling error makes the time-series analysis unstable for large  $K$ . We can analyze how this weighting alters the previous variance analysis when  $N_{\text{eig}} = 1$ . If we take this into account (see derivation in App.8.B), we find that

$$\text{Var}(\phi) \propto \frac{1}{KN}, \quad (8.28)$$

for a time-series analysis applied to multi-round  $k = 1$  experiments.



### Classical computation cost

In practice, the time-series analysis can be split into three calculations; (1) estimation of  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$  or  $\mathbb{P}(\mathbf{m}, \mathbf{n}|\phi, \mathbf{A})$ , (2) calculation of  $g(k)$  from these probabilities via Eq. (8.23) or Eq. (8.24), and (3) estimation of the phases  $\phi$  from  $g(k)$ . Clearly (2) and (3) only need to be done once for the entire set of experiments.

The estimation of the phases  $\phi$  requires solving two least squares equations, with cost  $O(l^2 K)$  (recalling that  $l$  is the number of frequencies to estimate, and  $K$  is the maximum known value of  $g(k)$ ), and diagonalizing the time-shift matrix  $\mathfrak{T}$  with cost  $O(l^3)$ . For single-round phase estimation this is the dominant calculation, as calculating  $g(k)$  from Eq. (8.23) requires simply  $K$  additions. As a result this estimator proves to be incredibly fast, able to estimate one frequency from a set of  $N = 10^6$  experiments of up to  $K = 10000$  in  $< 100$  ms, and  $l = 1000$  frequencies from  $N = 10^6$  experiments with  $K = 1000$  in  $< 1$  min. However, for multi-round phase estimation the calculation of  $g(k)$  in Eq. (8.24) scales as  $O(K^4)$ . This then dominates the calculation, requiring 30 s to calculate 50 points of  $g(k)$ . (All calculations performed on a 2.4 GHz Intel i3 processor.) We note that all the above times are small fractions of the time required to generate the experimental data when  $N \gg K$ , making this a very practical estimator for near-term experiments.

#### 8.3.2. Efficient Bayesian analysis

When the starting state is the eigenstate  $|\phi_0\rangle$ , the problem of determining  $\phi_0$  based on the obtained multi-experiment data has a natural solution via Bayesian methods [299, 319]. Here we extend such Bayesian methodology to a general starting state. For computational efficiency we store a probability distribution over phases  $P(\phi)$  using a Fourier representation of this periodic function  $P(\phi)$  (see 8.C). This technique can also readily be applied to the case of Bayesian phase estimation applied to a single eigenstate.

A clearly information-theoretic optimal Bayesian strategy is to choose the  $\phi$  and  $\mathbf{A}$  based on the data obtained in some  $N$  experiments [297]. After these  $N$  experiments, leading to qubit measurement outcomes  $\{\mathbf{m}_i\}_{i=1}^N$ , one can simply choose  $\mathbf{A}, \phi$  which maximizes the posterior distribution:

$$P_{\text{post}}(\phi, \mathbf{A}) = \frac{P_{\{\mathbf{k}_i\}, \{\beta_i\}}(\{\mathbf{m}_i\}|\phi, \mathbf{A})}{P(\{\mathbf{m}_i\})} P_{\text{prior}}(\phi, \mathbf{A}), \quad (8.29)$$

## 8. Quantum phase estimation for noisy, small-scale experiments

In other words, one chooses

$$\begin{aligned} (\phi_{\text{opt}}, \mathbf{A}_{\text{opt}}) &= \arg \max_{\phi, \mathbf{A}} \log P_{\text{post}}(\phi, \mathbf{A}) \\ &= \arg \max_{\phi, \mathbf{A}} [\log P_{\{\mathbf{k}_i\}, \{\beta_i\}}(\{\mathbf{m}_i\} | \phi, \mathbf{A}) + \log P_{\text{prior}}(\phi, \mathbf{A})]. \end{aligned}$$

A possible way of implementing this strategy is to (1) assume the prior distribution to be independent of  $\mathbf{A}$  and  $\phi$  and (2) estimate the maximum by assuming that the derivative with respect to  $\mathbf{A}$  and  $\phi$  vanishes at this maximum.

Instead of this method we update our probability distribution over  $\phi$  and  $\mathbf{A}$  after each experiment. After experiment  $n$  the posterior distribution  $P_n(\phi, \mathbf{A})$  via Bayes' rule reads

$$P_n(\phi, \mathbf{A}) = \frac{P_{\mathbf{k}, \beta}(\mathbf{m} | \phi, \mathbf{A})}{P(\mathbf{m})} P_{n-1}(\phi, \mathbf{A}). \quad (8.30)$$

To calculate the updates we will assume that the distribution over the phases  $\phi_j$  and probabilities  $A_j$  are independent, that is,

$$P_n(\phi, \mathbf{A}) = P_n^{\text{red}}(\mathbf{A}) \prod_{j=0}^{N_{\text{eig}}-1} P_n^j(\phi_j). \quad (8.31)$$

As prior distribution we take  $P_0(\phi, \mathbf{A}) = P_{\text{prior}}(\mathbf{A})P_{\text{prior}}(\phi)$  with a flat prior  $P_{\text{prior}}(\phi) = (\frac{1}{2\pi})^{N_{\text{eig}}}$ , given the absence of a more informed choice. We take  $P_{\text{prior}}(\mathbf{A}) = e^{-(\mathbf{A}-\mathbf{A}_0)^2/2\Sigma^2}$ , with  $\mathbf{A}_0$  and  $\Sigma^2$  approximate mean and covariance matrices. We need to do this to break the symmetry of the problem, so that  $\tilde{\phi}_0$  is estimating  $\phi_0$  and not any of the other  $\phi$ s. We numerically find that the estimator convergence is relatively independent of our choice of  $\mathbf{A}_0$  and  $\Sigma^2$ .

The approximation in Eq. (8.31) allows for relatively fast calculations of the Bayesian update of  $P_n^j(\phi_j)$ , and an approximation to the maximum-likelihood estimation of  $P_n^{\text{red}}(\mathbf{A})$ . Details of this computational implementation are given in 8.C.1.

### Classical computation cost

In contrast to the time-series estimator, the Bayesian estimator incurs a computational cost in processing the data from each individual experiment. On the other hand, obtaining the estimate  $\tilde{\phi}_0$  for  $\phi_0$  is simple, once one

has the probability distribution  $P^{j=0}(\phi)$ :

$$\tilde{\phi}_0 = \arg\left(\int d\phi P^{j=0}(\phi)e^{i\phi}\right).$$

A key parameter here is the number of frequencies  $\#\text{freq}$  stored in the Fourier representation of  $P(\phi)$ ; each update requires multiplying a vector of length  $N_{\text{freq}}$  by a sparse matrix. Our approximation scheme for calculating the update to  $\mathbf{A}$  makes this multiplication the dominant time cost of the estimation. As we argue in 8.C.1 one requires  $\#\text{freq} \geq K_{\text{tot}}$  to store a fully accurate representation of the probability vector. For the single-round scenario with  $k_r = 1$ , hence  $K_{\text{tot}} = N$ , we find a large truncation error when  $\#\text{freq} \ll N$ , and so the computation cost scales as  $N^2$ . In practice we find that processing the data from  $N < 10^4$  experiments takes seconds on a classical computer, but processing more than  $10^5$  experiments becomes rapidly unfeasible.

### 8.3.3. Experiment design

Based on the considerations above we seek to compare some choices for the meta-parameters in each experiment, namely the number of rounds, and the input parameters  $k_r$  and  $\beta_r$  for each round.

Previous work [299, 320], which took as a starting state the eigenstate  $|\phi_0\rangle$ , formulated a choice of  $k$  and  $\beta$ , using single-round experiments and Bayesian processing, namely

$$k = \min\left(\left\lceil \frac{1.25}{\sigma_{P_n^{j=0}(\phi_0)}} \right\rceil, K_{\text{err}}\right), \quad \beta \sim P_n^{j=0}(\phi_0 = \beta), \quad (8.32)$$

Roughly, this heuristic adapts to the expected noise in the circuit by not using any  $k$  such that the implementation of  $U^k$  takes longer than  $T_{\text{err}}/N_{\text{sys}}$ . It also adapts  $k$  to the standard-deviation of the current posterior probability distribution over  $\phi_0$ : a small standard-deviation after the  $n$ th experiment implies that  $k$  should be chosen large to resolve the remaining bits in the binary expansion of  $\phi_0$  \*.

In this chapter we use a starting state which is not an eigenstate, and as such we must adjust the choice in Eq. (8.32). As noted in Sec. 8.3, to

---

\*Note that this strategy is the opposite of textbook phase estimation in which one necessarily learns the least-significant bit of  $\phi_0$  *first* by choosing the largest  $k$ . One chooses the next smallest  $k$  and  $\beta$  so that the next measurement outcome gives the next more-significant bit etc.

## 8. Quantum phase estimation for noisy, small-scale experiments

separate different frequency contributions to  $g(k)$  we need good accuracy beyond that at a single value of  $k$ . The optimal choice of the number of frequencies to estimate depends on the distribution of the  $A_j$ , which may not be well known in advance. Following the inspiration of [299], we choose for the Bayesian estimator

$$k \in \{1, \dots, K\}$$

$$K = \min \left( \left\lceil \frac{1.25}{\sigma_{P_n^{j=0}(\phi_0)}} \right\rceil, K_{\text{err}} \right). \quad (8.33)$$

We thus similarly bound  $K$  depending how well one has already converged to a value for  $\phi_0$  which constitutes some saving of resources. At large  $N$  we numerically find little difference between choosing  $k$  at random from  $\{1, \dots, K\}$  and cycling through  $k = 1, \dots, K$  in order. For this Bayesian estimator we draw  $\beta$  at random from a uniform distribution  $[0, 2\pi)$ . We find that the choice of  $\beta$  has no effect on the final estimation (as long as it is not chosen to be a single number) For the time-series estimator applied to single-round experiments, we choose to cycle over  $k = 1, \dots, K$  so that it obtains a complete estimate of  $g(k)$  as soon as possible, taking an equal number of experiments with final rotation  $\beta = 0$  and  $\beta = \pi/2$  at each  $k$ . Here again  $K \leq K_{\text{err}}$ , so that we choose the same number of experiments for each  $k \leq K$ . For the time-series estimator applied to multi-round experiments, we choose an equal number of rounds with  $\beta = 0$  and  $\beta = \pi/2$ , taking the total number of rounds equal to  $R = K$ .

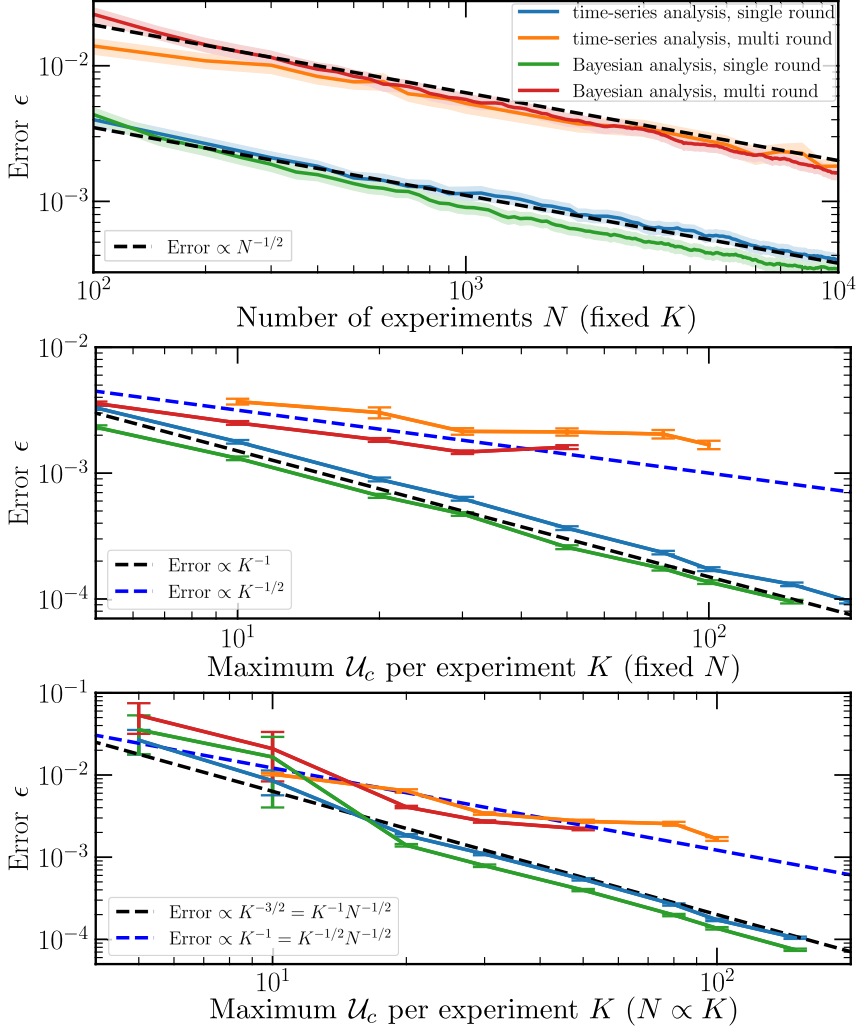
## 8.4. Results without experimental noise

We first focus on the performance of our estimators in the absence of experimental noise, to compare their relative performance and check the analytic predictions in Sec. 8.3.1. Although with a noiseless experiment our limit for  $K$  is technically infinite, we limit it to a make connection with the noisy results of the following section. Throughout this section we generate results directly by calculating the function  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$  and sampling from it. Note that  $P_{\mathbf{k},\beta}(\mathbf{m}|\phi, \mathbf{A})$  only depends on  $N_{\text{eig}}$  and not on the number of qubits in the system.

### 8.4.1. Single eigenvalues

To confirm that our estimators achieve the scaling bounds discussed previously, we first test them on the single eigenvalue scenario  $N_{\text{eig}} = 1$ . In

#### 8.4. Results without experimental noise



**Figure 8.3.:** Estimator performance for single eigenvalues with single and multi-round  $k = 1$  QPE schemes. Plots show scaling of the mean absolute error (Eq. (8.34)) with (top) the number of experiments (at fixed  $K = 50$ ), with (middle)  $K$  for a fixed total number of experiments ( $N = 10^6$ ), and (bottom) with  $K$  with a fixed number (100) of experiments per  $k = 1, \dots, K$  (i.e.  $N = 200K$ ). Data is averaged over 200-500 QPE simulations, with a new eigenvalue chosen for each simulation. Shaded regions (top) and error bars (middle, bottom) give 95% confidence intervals. Dashed lines show the scaling laws of Eq. (8.22) (fitted by eye). The top-right legend labeling the different estimation schemes is valid for all three plots.

## 8. Quantum phase estimation for noisy, small-scale experiments

Fig. 8.3, we plot the scaling of the average absolute error in an estimation  $\tilde{\phi}$  of a single eigenvalue  $\phi \in [-\pi, \pi)$ , defined so as to respect the  $2\pi$ -periodicity of the phase:

$$\epsilon := \langle \min(|\phi - \tilde{\phi}|, 2\pi - |\phi - \tilde{\phi}|) \rangle = \langle \left| \text{Arg} \left( e^{i(\phi - \tilde{\phi})} \right) \right| \rangle, \quad (8.34)$$

as a function of varying  $N$  and  $K$ . Here  $\langle \rangle$  represents an average over repeated QPE simulations, and the  $\text{Arg}$  function is defined using the range  $[-\pi, \pi)$  (otherwise the equality does not hold).

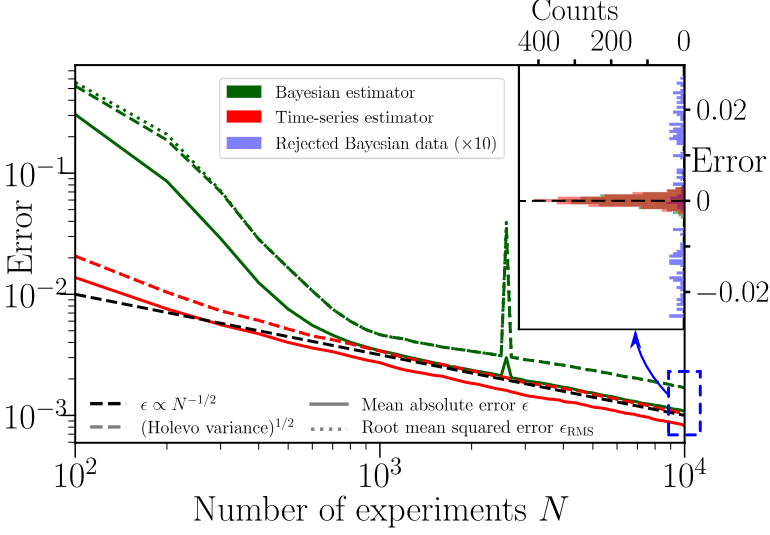
We see that both estimators achieve the previously-derived bounds in 8.3.1 (overlayed as dashed lines), and both estimators achieve almost identical convergence rates. The results for the Bayesian estimation match the scaling observed in Ref. [299]. Due to the worse scaling in  $K$ , the multi-round  $k = 1$  estimation significantly underperforms single-round phase estimation. This is a key observation of this chapter, showing that if the goal is to estimate a phase rather than to project onto an eigenstate, it is preferable to do single-round experiments.

### 8.4.2. Example behaviour with multiple eigenvalues

The performance of quantum phase estimation is dependent on both the estimation technique and the system being estimated. Before studying the system dependence, we first demonstrate that our estimators continue to perform at all in the presence of multiple eigenvalues. In Fig. 8.4, we demonstrate the convergence of both the Bayesian and time-series estimators in the estimation of a single eigenvalue  $\phi_0 = -0.5$  of a fixed unitary  $U$ , given a starting state  $|\Psi_0\rangle$  which is a linear combination of 10 eigenstates  $|\phi_j\rangle$ . We fix  $|\langle \phi_0 | \Psi_0 \rangle|^2 = 0.5$ , and draw other eigenvalues and amplitudes at random from  $[0, \pi]$  (making the minimum gap  $\phi_j - \phi_0$  equal to 0.5). We perform 2000 QPE simulations with  $K = 50$ , and calculate the mean absolute error  $\epsilon$  (Eq. (8.34), solid), Holevo variance  $\left| \langle e^{i\tilde{\phi}} \rangle \right|^{-2} - 1$  (dashed), and root mean squared error  $\epsilon_{\text{RMS}}$  (dotted), given by

$$\epsilon_{\text{RMS}}^2 := \langle \min(|\phi - \tilde{\phi}|, 2\pi - |\phi - \tilde{\phi}|)^2 \rangle = \langle \left| \text{Arg} \left( e^{i(\phi - \tilde{\phi})} \right) \right|^2 \rangle. \quad (8.35)$$

We observe that both estimators retain their expected  $\epsilon \propto N^{-1/2}$ , with one important exception. The Bayesian estimator occasionally (10% of simulations) estimates multiple eigenvalues near  $\phi_0$ . When this occurs, the estimations tend to repulse each other, making neither a good estimation



**Figure 8.4.:** Scaling of error for time-series (dark green) and Bayesian (red) estimators with the number of experiments performed for a single shot of a unitary with randomly drawn eigenphases (parameters given in text). Three error metrics are used as marked (described in text - note that the mean squared error and Holevo variance completely overlap for the time-series estimator). Data is averaged over 2000 simulations. The peak near  $N = 3000$  comes from deviation in a single simulation and is not of particular interest. With this exception, error bars are approximately equal to width of the lines used. (Inset) histogram of the estimated phases after  $N = 10^4$  experiments. Blue bars correspond to Bayesian estimates that were rejected (rejection method described in text). These have been magnified  $10\times$  to be made visible.

of the target. This is easily diagnosable without knowledge of the true value of  $\phi_0$  by inspecting the gap between estimated eigenvalues. While using this data to improve estimation is a clear target for future research, for now we have opted to reject simulations where such clustering occurs (in particular, we have rejected datapoints where  $\min(\bar{\phi}_0 - \bar{\phi}_j) < 0.05$ ). That this is required is entirely system-dependent: we find the physical Hamiltonians studied later in this text to not experience this effect. We attribute this difference to the distribution of the amplitudes  $A_j$  - physical Hamiltonians tend to have a few large  $A_j$ , whilst in this simulation the  $A_j$  were distributed uniformly.

In the inset to Fig. 8.4, we plot a histogram of the estimated eigenphases

## 8. Quantum phase estimation for noisy, small-scale experiments

after  $N = 10^4$  experiments. For the Bayesian estimator, we show both the selected (green) and rejected (blue) eigenphases. We see that regardless of whether rejection is used, the distribution appears symmetric about the target phase  $\phi_0$ . This suggests that in the absence of experimental noise, both estimators are unbiased. Proving this definitively for any class of systems is difficult, but we expect both estimators to be unbiased provided  $A_0 \gg 1/K$ . When  $A_0 \leq 1/K$ , one can easily construct systems for which no phase estimation can provide an unbiased estimation of  $\phi_0$  (following the arguments of Sec. 8.3). We further see that the scaling of the RMS error  $\epsilon_{\text{RMS}}$  and the Holevo variance match the behaviour of the mean absolute error  $\epsilon$ , implying that our results are not biased by the choice of estimator used.

### 8.4.3. Estimator scaling with two eigenvalues

The ability of QPE to resolve separate eigenvalues at small  $K$  can be tested in a simple scenario of two eigenvalues,  $\phi_0$  and  $\phi_1$ . The input to the QPE procedure is then entirely characterized by the overlap  $A_0$  with the target state  $|\phi_0\rangle$ , and the gap  $\delta = |\phi_0 - \phi_1|$ .

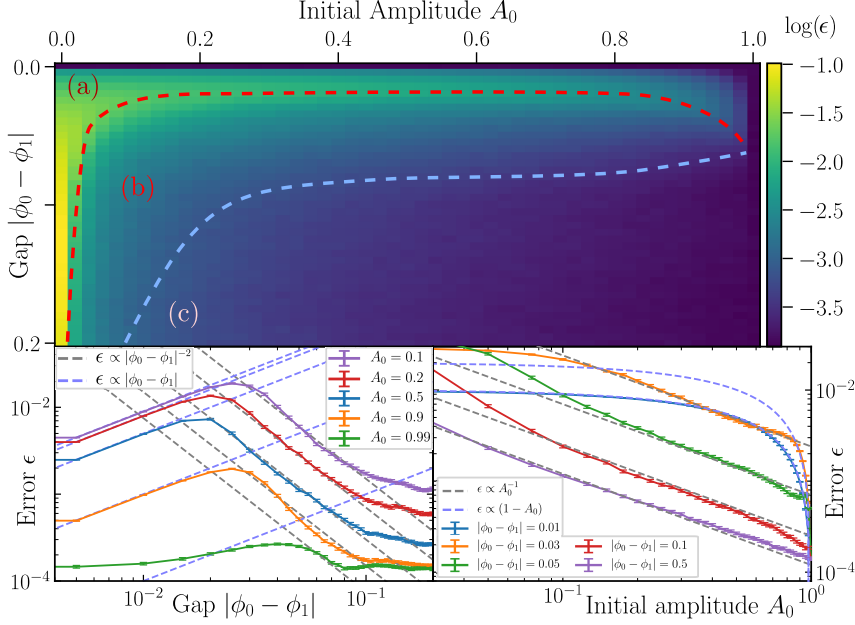
In Fig. 8.5, we study the performance of our time-series estimator in estimating  $\phi_0$  after  $N = 10^6$  experiments with  $K = 50$ , measured again by the mean error  $\epsilon$  (Eq. (8.34)). We show a two-dimensional plot (averaged over 500 simulations at each point  $A_0, \delta$ ) and log-log plots of one-dimensional vertical (lower left) and horizontal (lower right) cuts through this surface. Due to computational costs, we are unable to perform this analysis with the Bayesian estimator, or for the multi-round scenario. We expect the Bayesian estimator to have similar performance to the time-series estimator (given their close comparison in Sec. 8.4.1 and Sec. 8.4.2). We also expect the error in multi-round QPE to follow similar scaling laws in  $A_0$  and  $\delta$  as single-round QPE (i.e. multi-round QPE should be suboptimal only in its scaling in  $K$ ).

The ability of our estimator to estimate  $\phi_0$  in the presence of two eigenvalues can be split into three regions (marked as (a), (b), (c) on the surface plot). In region (a), we have performed insufficient sampling to resolve the eigenvalues  $\phi_0$  and  $\phi_1$ , and QPE instead estimates the weighted average phase  $A_0\phi_0 + A_1\phi_1$ . The error in the estimation of  $\phi_0$  then scales by how far it is from the average, and how well the average is resolved

$$\epsilon \propto (1 - A_0)\delta K^{-1}N^{-1/2}. \quad (8.36)$$

In region (b), we begin to separate  $\phi_0$ , from the unwanted frequency  $\phi_1$ ,





**Figure 8.5.:** Performance of the time-series estimator in the presence of two eigenvalues. (top) Surface plot of the error after  $N = 10^6$  experiments for  $K = 50$ , as a function of the overlap  $A_0$  with the target state  $|\phi_0\rangle$ , and the gap  $|\phi_0 - \phi_1|$ . Plot is divided by hand into three labeled regions where different scaling laws are observed. Each point is averaged over 500 QPE simulations. (bottom) log-log plots of vertical (bottom left) and horizontal (bottom right) cuts through the surface, at the labeled positions. Dashed lines in both plots are fits (by eye) to the observed scaling laws. Each point is averaged over 2000 QPE simulations, and error bars give 95% confidence intervals.

and our convergence halts,

$$\epsilon \propto A_0^{-1} \delta^{-2}. \quad (8.37)$$

In region (c), the gap is sufficiently well resolved and our estimation returns to scaling well with  $N$  and  $K$

$$\epsilon \propto A_0^{-1} K^{-1} N^{-1/2}. \quad (8.38)$$

The scaling laws in all three regions can be observed in the various cuts in the lower plots of Fig. 8.5. We note that the transition between the

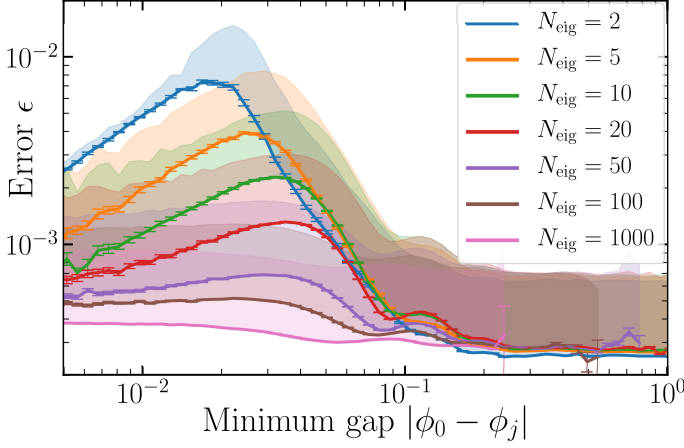
three regions is not sharp (boundaries estimated by hand), and is  $K$  and  $N$ -dependent.

#### 8.4.4. Many eigenvalues

To show that our observed scaling is applicable beyond the toy 2-eigenvalue system, we now shift to studying systems of random eigenvalues with  $N_{\text{eig}} > 1$ . In keeping with our insight from the previous section, in Fig. 8.6 we fix  $\phi_0 = 0$ , and study the error  $\epsilon$  as a function of the gap

$$\delta = \min_{j>1} (|\phi_j - \phi_0|). \quad (8.39)$$

We fix  $A_0 = 0.5$ , and draw the other parameters for the system from a uniform distribution:  $\phi_j \sim [\delta, \pi]$ ,  $A_j \sim [0, 0.5]$  (fixing  $\sum_{j=1}^{N_{\text{eig}}} A_j = 1 - A_0$ ). We plot both the average error  $\epsilon$  (line) and the upper 47.5% confidence interval  $[\epsilon, \epsilon + 2\sigma_\epsilon]$  (shaded region) for various choices of  $N_{\text{eig}}$ . We observe that increasing the number of spurious eigenvalues does not critically affect the error in estimation; indeed the error generally decreases as a function of the number of eigenvalues. This makes sense; at large  $N_{\text{eig}}$  the majority of eigenvalues sit in region (c) of Fig. 8.5, and we do not expect these to combine to distort the estimation. Then, the nearest eigenvalue  $\min_{j \neq 0} \phi_j$  has on average an overlap  $A_j \propto 1/N_{\text{eig}}$ , and its average contribution to the error in estimating  $\phi_0$  (inasmuch as this can be split into contributions) scales accordingly. We further note that the worst-case error remains that of two eigenvalues at the crossover between regions (a) and (b). In App.8.D we study the effect of confining the spurious eigenvalues to a region  $[\delta, \phi_{\text{max}}]$ . We observe that when most eigenvalues are confined to regions (a) and (b), the scaling laws observed in the previous section break down, however the worst-case behaviour remains that of a single spurious eigenvalue. This implies that sufficiently long  $K$  is not a requirement for QPE, even in the presence of large systems or small gaps  $\delta$ ; it can be substituted by sufficient repetition of experiments. However, we do require that the ground state is guaranteed to have sufficient overlap with the starting state -  $A_0 > 1/K$  (as argued in Sec. 8.3). As QPE performance scales better with  $K$  than it does with  $N$ , a quantum computer with coherence time  $2T$  is still preferable to two quantum computers with coherence time  $T$  (assuming no coherent link between the two).



**Figure 8.6.:** Performance of the time-series estimator in the presence of multiple eigenvalues. Error bars show 95% confidence intervals (data points binned from  $4 \times 10^6$  simulations). Shaded regions show upper  $2\sigma$  interval of data for each bin.

## 8.5. The effect of experimental noise

Experimental noise currently poses the largest impediment to useful computation on current quantum devices. As we suggested before, experimental noise limits  $K$  so that for  $K \gtrsim K_{\text{err}}$  the circuit is unlikely to produce reliable results. However, noise on quantum devices comes in various flavours, which can have different corrupting effects on the computation. Some of these corrupting effects (in particular, systematic errors) may be compensated for with good knowledge of the noise model. For example, if we knew that our system applied  $U = e^{-i\mathcal{H}(t+\epsilon)}$  instead of  $U = e^{-i\mathcal{H}t}$ , one could divide  $\tilde{\phi}_0$  by  $(t + \epsilon)/t$  to precisely cancel out this effect. In this study we have limited ourselves to studying and attempting to correct two types of noise: depolarizing noise, and circuit-level simulations of superconducting qubits. Given the different effects observed, extending our results to other noise channels is a clear direction for future research. In this section we do not study multi-round QPE, so each experiment consists of a single round. A clear advantage of the single-round method is that the only *relevant* effect of any noise in a single-round experiment is to change the outcome of the ancilla qubit, independent of the number of system qubits  $N_{\text{sys}}$ .

### 8.5.1. Depolarizing noise

A very simple noise model is that of depolarizing noise, where the outcome of each experiment is either correct with some probability  $p$  or gives a completely random bit with probability  $1 - p$ . We expect this probability  $p$  to depend on the circuit time and thus the choice of  $k \geq 0$ , i.e.

$$p = p(k) = e^{-k/K_{\text{err}}}. \quad (8.40)$$

We can simulate this noise by directly applying it to the calculated probabilities  $P_{k,\beta}(m|\phi)$  for a single round

$$P_{k,\beta}(m|\phi) \rightarrow P_{k,\beta}(m|\phi)p(k) + \frac{1 - p(k)}{2}. \quad (8.41)$$

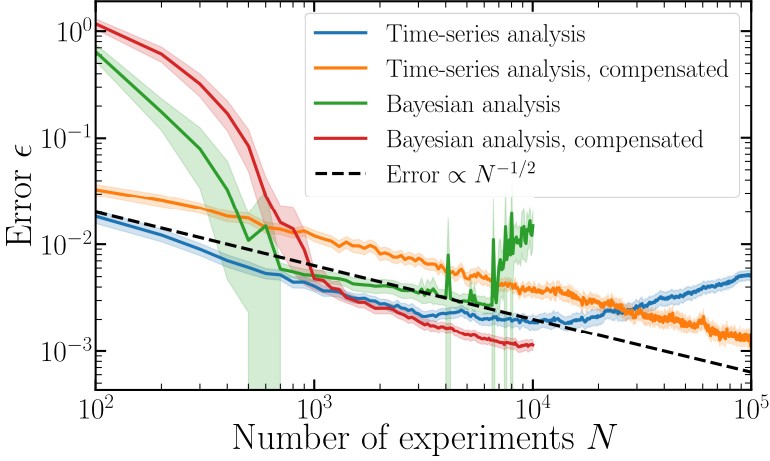
In Fig. 8.7, we plot the convergence of the time-series (blue) and Bayesian (green) estimators as used in the previous section as a function of the number of experiments, with fixed  $K = 50 = K_{\text{err}}/2$  fixed,  $A_0 = 0.5$ ,  $N_{\text{eig}} = 10$  and  $\delta = 0.5$ . We see that both estimators obey  $N^{-1/2}$  scaling for some portion of the experiment, however this convergence is unstable, and stops beyond some critical point.

Both the Bayesian and time-series estimator can be adapted rather easily to compensate for this depolarizing channel. To adapt the time-series analysis, we note that the effect of depolarizing noise is to send  $g(k) \rightarrow g(k)p(k)$  when  $k > 0$ , via Eq. (8.23) and Eq. (8.41). Our time-series analysis was previously performed over the range  $k = -K, \dots, K$  (getting  $g(-k) = g^*(k)$  for free), and over this range

$$g(k) \rightarrow g(k)p(|k|). \quad (8.42)$$

$g(k)$  is no longer a sum of exponential functions over our interval  $[-K, K]$ , as it is not differentiable at  $k = 0$ , which is the reason for the failure of our time-series analysis. However, over the interval  $[0, K]$  this is not an issue, and the time-series analysis may still be performed. If we construct a shift operator  $T$  using  $g(k)$  from  $k = 0, \dots, K$ , this operator will have eigenvalues  $e^{i\phi_j - 1/K_{\text{err}}}$ . This then implies that the translation operator  $T$  can be calculated using  $g(k)$  with  $k > 0$ , and the complex argument of the eigenvalues of  $T$  give the correct phases  $\phi_j$ . We see that this is indeed the case in Fig. 8.7 (orange line). Halving the range of  $g(k)$  that we use to estimate  $\phi_0$  decreases the estimator performance by a constant factor, but this can be compensated for by increasing  $N$ .

Adapting the Bayesian estimator requires simply that we use the correct



**Figure 8.7.:** Convergence of Bayesian and time-series estimators in the presence of depolarizing noise and multiple eigenvalues, both with and without noise compensation techniques (described in text). Fixed parameters for all plots are given in text. Shaded regions denote a 95% confidence interval (data estimated over 200 QPE simulations). The black dashed line shows the  $N^{-1/2}$  convergence expected in the absence of sampling noise. Data for the Bayesian estimator was not obtained beyond  $N = 10^4$  due to computational constraints.

conditional probability, Eq. (8.41). This in turn requires that we either have prior knowledge of the error rate  $K_{\text{err}}$ , or estimate it alongside the phases  $\phi_j$ . For simplicity, we opt to choose the former. In an experiment  $K_{\text{err}}$  can be estimated via standard QCVV techniques, and we do not observe significant changes in estimator performance when it is detuned. Our Fourier representation of the probability distribution of  $\phi_0$  can be easily adjusted to this change. The results obtained using this compensation are shown in Fig. 8.7: we observe that the data follows a  $N^{-1/2}$  scaling again.

### 8.5.2. Realistic circuit-level noise

Errors in real quantum computers occur at a circuit-level, where individual gates or qubits get corrupted via various error channels. To make connection to current experiments, we investigate our estimation performance on an error model of superconducting qubits. Full simulation details can be found in App.8.E. Our error model is primarily dominated by  $T_1$  and  $T_2$  decoherence, incoherent two-qubit flux noise, and dephasing

## 8. Quantum phase estimation for noisy, small-scale experiments

during single-qubit gates. We treat the decoherence time  $T_{\text{err}} = T_1 = T_2$  as a free scale parameter to adjust throughout our simulations, whilst keeping all other error parameters tied to this single scale parameter for simplicity. In order to apply circuit-level noise we must run quantum circuit simulations, for which we use the `quantumsim` density matrix simulator first introduced in [212]. We then choose to simulate estimating the ground state energy of four hydrogen atoms in varying rectangular geometries, with Hamiltonian  $\mathcal{H}$  taken in the STO-3G basis calculated via `psi4` [321], requiring  $N_{\text{sys}} = 8$  qubits. We make this estimation via a lowest-order Suzuki-Trotter approximation [322] to the time-evolution operator  $e^{-i\mathcal{H}t}$ . To prevent energy eigenvalues wrapping around the circle we fix  $t = 1/\sqrt{\text{Trace}[\mathcal{H}^\dagger\mathcal{H}]/(2^{N_{\text{sys}}})}$  \*. The resultant 9-qubit circuit is made using the `OpenFermion` package [298].

In lieu of any circuit optimizations (e.g. [264, 290]), the resulting circuit has a temporal length per unitary of  $T_U = 42 \mu\text{s}$  (with single- (two-) qubit gate times 20 ns (40 ns)). This makes the circuit unrealistic to operate at current decoherence times for superconducting circuits, and we focus on decoherence times 1 – 2 orders of magnitude above what is currently feasible, i.e.  $T_{\text{err}} = 5 - 50$  ms. However one may anticipate that the ratio  $T_U/T_{\text{err}}$  can be enlarged by circuit optimization or qubit improvement. Naturally, choosing a smaller system, less than 8 qubits, or using error mitigation techniques could also be useful.

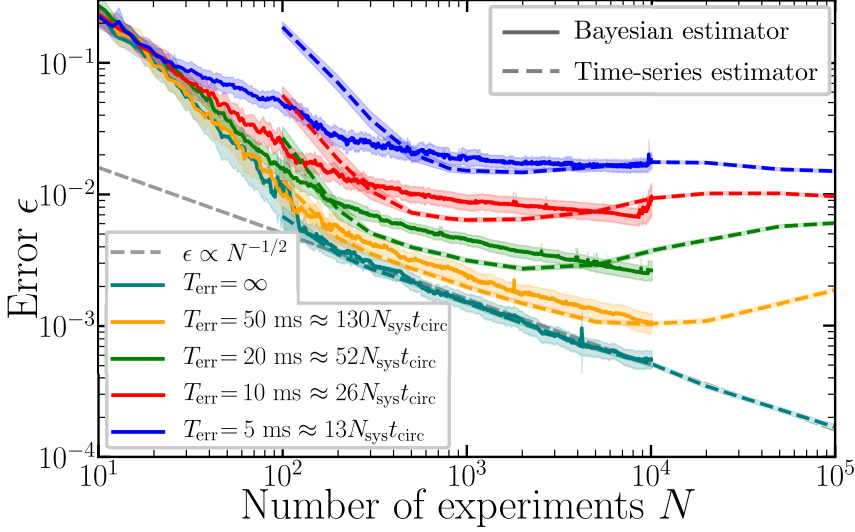
We observe realistic noise to have a somewhat different effect on both estimators than a depolarizing channel. Compared to the depolarizing noise, the noise may (1) be biased towards 0 or 1 and/or (2) its dependence on  $k$  may not have the form of Eq. (8.40).

In Fig. 8.8, we plot the performance of both estimators at four different noise levels (and a noiseless simulation to compare), in the absence of any attempts to compensate for the noise. Unlike for the depolarizing channel, where a  $N^{-1/2}$  convergence was observed for some time before the estimator became unstable, here we see both instabilities and a loss of the  $N^{-1/2}$  decay to begin with. Despite this, we note that reasonable convergence (to within 1 – 2%) is achieved, even at relatively low coherence times such as  $K_{\text{err}} = 10$ . Regardless, the lack of eventual convergence to zero error is worrying, and we now shift to investigating how well it can

---

\*This normalization is not good for large systems since it makes  $t$  exponentially small in system size. A scalable choice for normalization is to first determine upper and lower bounds on the eigenvalues of  $\mathcal{H}$  present in the starting state, assume that they occur in a some numerical window  $W$ . Given  $W$  (which is at most  $\text{poly}(N_{\text{sys}})$ ), one sets  $U = \exp(-i\pi\mathcal{H}/W)$ . The implementation of this  $U$  in Trotterized form with sufficient accuracy determines  $T_U$ .

be improved for either estimator.

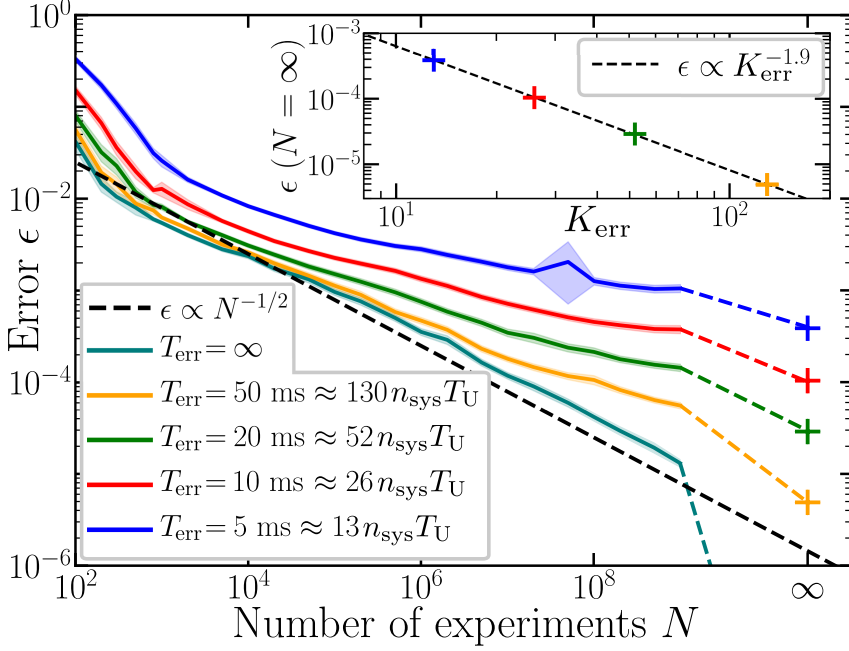


**Figure 8.8.:** Performance of Bayesian (solid) and time-series (dashed) estimators in the presence of realistic noise without any compensation techniques. Shaded regions denote 95% confidence intervals (averaged over 100 – 500 QPE simulations). The time-series analysis requires  $N > 2K$  experiments in order to produce an estimate, and so its performance is not plotted for  $N < 100$ .

Adjusting the time-series estimator to use only  $g(k)$  for positive  $k$  gives approximately 1 – 2 orders of magnitude improvement. In Fig. 8.9, we plot the estimator convergence with this method. We observe that the estimator is no longer unstable, but the  $N^{-1/2}$  convergence is never properly regained. We may study this convergence in greater detail for this estimator, as we may extract  $g(k)$  directly from our density-matrix simulations, and thus investigate the estimator performance in the absence of sampling noise (crosses on screen). We note that similar extrapolations in the absence of noise, or in the presence of depolarizing noise (when compensated) give an error rate of around  $10^{-10}$ , which we associate to fixed-point error in the solution to the least squares problem (this is also observed in the curve without noise in Fig. 8.9). Plotting this error as a function of  $K_{\text{err}}$  shows a power-law decay -  $\epsilon \propto K_{\text{err}}^{-\alpha} \propto T_{\text{err}}^{-\alpha}$  with  $\alpha = 1.9 \approx 2$ . We do not have a good understanding of the source of the obtained power law.

The same compensation techniques that restored the performance of the Bayesian estimator in the presence of depolarizing noise do not work

## 8. Quantum phase estimation for noisy, small-scale experiments



**Figure 8.9.:** Performance of time-series estimator with compensation techniques (described in text). Shaded regions denote 95% confidence intervals (averaged over 200 QPE simulations). Final crosses show the performance in the absence of any sampling noise (teal cross is at approximately  $10^{-10}$ ), i.e. in the limit  $N \rightarrow \infty$ ; dashed lines are present to demonstrate this limit. (inset) Plot of error without sampling noise as a function of the decoherence time  $T_{\text{err}}$ . Y-axis corresponds to y-axis on main plot (as color-coded).

nearly as well for realistic noise. Most likely this is due to the fact that the actual noise is not captured by a  $k$ -dependent depolarizing probability. In Fig. 8.10 we plot the results of using a Bayesian estimator when attempting to compensate for circuit-level noise by approximating it as a depolarizing channel with a decay rate (Eq. 8.40) of  $K_{\text{err}} = T_{\text{err}}/T_U N_{\text{sys}}$ . This can be compared with the results of Fig. 8.8 where this compensation is not attempted. We observe a factor 2 improvement at low  $T_{\text{err}}$ , however the  $N^{-1/2}$  scaling is not regained, and indeed the estimator performance appears to saturate at roughly this point. Furthermore, at  $T_{\text{err}} = 50 \text{ ms}$ , the compensation techniques do not improve the estimator, and indeed appear to make it more unstable.



	Time-series estimator	Bayesian estimator
Speed (scaling)	$O(K)$	$O(N^2)$
Speed (timing)	Processes large datasets in milliseconds	Takes hours to process $10^5$ experiments
Accuracy	$\epsilon \propto N^{-1/2} K^{-1} A_0^{-1} \delta^{-2}$ demonstrated.	$\epsilon \propto N^{-1/2} K^{-1}$ demonstrated $\epsilon \propto A_0^{-1} \delta^{-2}$ expected.
Number of eigenvalues estimated	100 – 200 with relative ease	Limited to 2 – 5
Improve accuracy via classical approximation	Not obvious	Can get speedup via choice of prior (not attempted in this chapter)
Account for error	Limited ability	Limited ability

**Table 8.1.:** Table comparing metrics of interest between the two studied estimators. All metrics are implementation-specific, and may be improvable.

To investigate this further, in Fig. 8.10 (inset) we plot a Bayes Factor analysis of the Bayesian estimators with and without compensation techniques. The Bayes Factor analysis is obtained by calculating the Bayes Factors

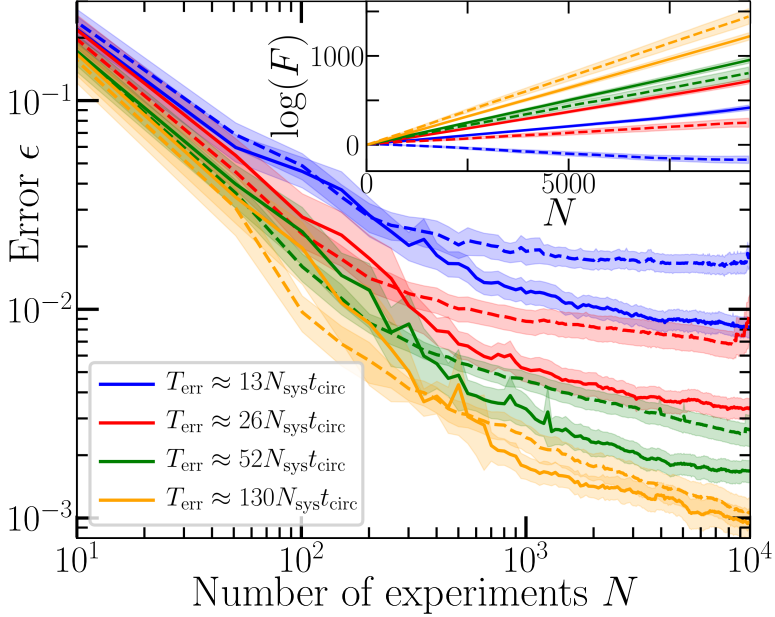
$$F = \prod_{\text{expt } n} \frac{P(m_n|M)}{P(m_n|M_0)}, \quad (8.43)$$

where  $M$  is the chosen Bayesian model (including the prior knowledge), and  $M_0$  is a reference model, and  $P(m|M)$  is the probability of observing measurement  $m$  given model  $M$ . As a reference model we take that of random noise -  $P(m|M_0) = 0.5$ . We observe that at large  $T_{\text{err}}$  the Bayes factor with compensation falls below that without, implying that the compensation techniques make the model worse. We also observe that at very small  $T_{\text{err}}$ , the estimator makes worse predictions than random noise ( $\log(F) < 0$ ). Despite our best efforts we have been unable to further improve the Bayesian estimator in noisy single-round QPE experiments.

## 8.6. Discussion

In this chapter, we have presented and studied the performance of two estimators for quantum phase estimation at low  $K$  for different experiment protocols, different systems (in particular those with one vs many eigenvalues), and under simplistic and realistic noise conditions. These findings are

## 8. Quantum phase estimation for noisy, small-scale experiments



**Figure 8.10.:** Performance of single-round Bayesian QPE with four sets of realistic noise using a compensation technique described in the text. Shaded regions are 95% confidence intervals over 200 – 500 QPE simulations. (inset) a Bayes factor analysis for the data below. Line color and style matches the legend of the main figure.

summarized in Table 8.1. From our numerical studies, we observe scaling laws for our time-series estimator; we find it first-order sensitive to the overlap  $A_0$  between starting state and ground state, second-order sensitive to the gap between the ground state and the nearest eigenstates, and second-order sensitive to the coherence time of the system. The Bayesian estimator appears to perform comparably to the time-series estimator in all circumstances, and thus should obey similar scaling laws.

We further observe that realistic noise has a worse effect on QPE than a depolarizing channel, for which the effects can largely be mitigated. We have numerically explored (but not reported) multi-round QPE in the presence of noise. Since each experiment has multiple outputs, it is harder to adapt the classical data analysis to the presence of noise and our results for realistic noise have not been convincing so far. Since the performance of multi-round noiseless QPE is already inferior to single-round noiseless

QPE, we do not advocate it as a near-term solution, although, for noiseless long circuits it does have the ability to project onto a single eigenstate, which single-round QPE certainly does not.

Despite our slightly pessimistic view of the effect of errors on the performance of QPE, we should note that the obtained error of  $10^{-3}$  at  $T_{\text{err}} \approx 13N_{\text{sys}}T_U$  or  $K_{\text{err}} = 13$  would be sufficient to achieve chemical accuracy in a small system. However, as the energy of a system scales with the number of particles, if we require a Hamiltonian's spectrum to fit in  $[-\pi, \pi)$ , we will need a higher resolution for QPE, making error rates of  $10^{-3}$  potentially too large. This could potentially be improved by improving the compensation techniques described in the text, applying error mitigation techniques to effectively increase  $T_{\text{err}}$ , or by using more well-informed prior distributions in the Bayesian estimator to improve accuracy. All of the above are obvious directions for future work in optimizing QPE for the NISQ era. Another possible direction is to investigate QPE performance in other error models than the two studied here. Following Ref. [295], we expect SPAM errors to be as innocuous as depolarizing noise. However, coherent errors can be particularly worrying as they imitate alterations to the unitary  $U$ . The time-series estimator is a clear candidate for such a study, due to its ease in processing a large number of experiments and its ability to be studied in the absence of sampling noise. We also expect that it is possible to combine the time-series estimator with the Heisenberg-limited scaling methods of Refs. [295, 316] so as to extend these optimal methods to the multiple-eigenvalue scenario with  $N_{\text{eig}} > 1$  eigenvalues, and that these methods could be extended to analog or ancilla-free QPE settings such as described in Ref. [295].

In this chapter we do not compare the performance of quantum phase estimation with purely classical methods. Let's assume that we have a classical efficient representation of the starting state  $\Psi$  and one can efficiently calculate  $\text{Tr} \mathcal{H}^k |\Psi\rangle\langle\Psi|$  for  $k = 1, \dots, K$  with  $K = O(1)$  (for fermionic Gaussian starting states and fermionic Hamiltonians this is possible as a single fermionic term in  $\mathcal{H}^k$  can be estimated as the Pfaffian of some matrix). Then, if there are at most  $K = O(1)$  eigenstates in this initial state, the time-series method would allow us to extract these eigenvalues efficiently. Thus in this setting and under these assumptions quantum phase estimation would not offer an exponential computational advantage.

## 8.A. Derivation of the identity in Eq. (8.25)

One first writes for  $0 \leq k \leq K/2$ :

$$\sum_j A_j \exp(ik\phi_j) = \sum_{\mathbf{m}, \mathbf{n}} \Pi_{i=1}^k [(-1)^{m_i} - i(-1)^{n_i}] \times \mathbb{P}(m_1, \dots, m_{K/2}, n_1, \dots, n_{K/2} | \phi, \mathbf{A}) \quad (8.44)$$

where  $\mathbb{P}(m_1, \dots, m_{K/2}, n_1, \dots, n_{K/2} | \phi, \mathbf{A})$  is the probability for a specific series of outcomes  $m_1, \dots, m_{K/2}$  for  $\beta = 0$  and  $n_1, \dots, n_{K/2}$  for  $\beta = \pi/2$ . To see that the above is true, note that it is quickly true for  $N_{\text{eig}} = 1$  by using Eq. (8.23) for  $g(1)$ . By linearity on the left and right hand side it then holds generally.

Since the order of the outcomes of the rounds does not matter, i.e.  $\mathbb{P}(m_1, \dots, m_{K/2}, n_1, \dots, n_{K/2} | \phi, \mathbf{A})$  only depends on the Hamming weights  $\mathbf{m} = |\mathbf{m}|$  and  $\mathbf{n} = |\mathbf{n}|$ , we can symmetrize the coefficient over permutations of the rounds and replace  $\mathbb{P}(m_1, \dots, m_{K/2}, n_1, \dots, n_{K/2} | \phi, \mathbf{A})$  by  $\mathbb{P}(\mathbf{m}, \mathbf{n} | \phi, \mathbf{A}) / ((\binom{K/2}{\mathbf{m}} \binom{K/2}{\mathbf{n}}))$ . This gives the following expression for  $\chi_k(m, n)$ :

$$\chi_k(\mathbf{m}, \mathbf{n}) = \frac{1}{((K/2)!)^2} \sum_{\pi_1 \in S_{K/2}, \pi_2 \in S_{K/2}} \prod_{i=1}^k ((-1)^{m_{\pi_1(i)}} - i(-1)^{n_{\pi_2(i)}}),$$

where  $m_i$  is the  $i$ th bit of a bitstring with Hamming weight  $\mathbf{m}$  (and similarly  $n_i$ ), and  $S_{K/2}$  is the symmetric group of permutations. We can expand this last expression as

$$\begin{aligned} \chi_k(\mathbf{m}, \mathbf{n}) &= \sum_{k=0}^l \binom{k}{l} (-i)^{k-l} \rho(l, \mathbf{m}) \rho(k-l, \mathbf{n}) \\ \rho(l, \mathbf{m}) &= \frac{1}{(K/2)!} \sum_{\pi} (-1)^{m_{\pi(1)}} \dots (-1)^{m_{\pi(l)}} \\ &= -1 + \frac{2}{(K/2)!} \sum_{\pi: m_{\pi(1)} \dots m_{\pi(l)} \text{ is even}} 1 \end{aligned}$$

The sum  $\sum_{\pi: m_{\pi(1)} \dots m_{\pi(l)} \text{ is even}}$  can be written as a sum over permutations such that  $m_{\pi(1)} \dots m_{\pi(l)}$  has Hamming weight  $2p$  with  $p = 0, 1, \dots, \lfloor l/2 \rfloor$ .

Then one counts the number of permutations of a  $K/2$ -bitstring of Hamming weight  $\mathbf{m}$  such that some segment of length  $l$  has Hamming weight  $2p$  which equals  $\binom{\mathbf{m}}{2p} \binom{K/2-\mathbf{m}}{l-2p} l! (K/2-l)!$ . All together this leads to  $\chi_k(\mathbf{m}, \mathbf{n})$  in Eq. (8.25). It is not clear whether one can simplify this equality or verify it directly using other combinatorial identities or (Chebyshev) polynomials.

## 8.B. Variance calculations for time-series estimator

For the case of estimating a single eigenvalue using single-round QPE with the time-series estimator, one can directly calculate the error in the estimation. In this situation, our matrices  $G_0$  and  $G_1$  are column vectors,

$$G_0^T = (g(-K), g(-K+1), \dots, g(K-1)), \quad (8.45)$$

$$G_1^T = (g(-K+1), g(-K+2), \dots, g(K)). \quad (8.46)$$

The least-squares solution for  $\mathfrak{T}$  is then

$$\mathfrak{T} = (G_0^\dagger G_0)^{-1} G_0^\dagger G_1 = \frac{\sum_{k=-K}^{K-1} g^*(k) g(k+1)}{\sum_{k=-K}^{K-1} g^*(k) g(k)}. \quad (8.47)$$

For a single frequency,  $g(k) = e^{ik\phi}$ , and immediately  $\mathfrak{T} = e^{i\phi}$ . However, we estimate the real and imaginary components of  $g(k)$  separately. Let us write in terms of our independent components

$$\mathfrak{T} = \mathfrak{T}_r + i\mathfrak{T}_i, \quad g(k) = g_k^0 + ig_k^1, \quad (8.48)$$

remembering that  $g_k^0 = g_{-k}^0$  and  $g_k^1 = -g_{-k}^1$  (i.e. the variables are correlated). Our target angle  $\phi = \tan^{-1} \mathfrak{T}_i / \mathfrak{T}_r$ , and so we can calculate

$$\begin{aligned} \text{Var}(\phi) &= \sum_{a,k} \left[ \frac{\partial \phi}{\partial g_k^a} \right]^2 \text{Var}[g_k^a] \\ &= \left[ \frac{1}{\mathfrak{T}_r^2 + \mathfrak{T}_i^2} \right]^2 \sum_{a,k} \left[ \mathfrak{T}_r \frac{\partial \mathfrak{T}_i}{\partial g_k^a} - \mathfrak{T}_i \frac{\partial \mathfrak{T}_r}{\partial g_k^a} \right]^2 \text{Var}[g_k^a]. \end{aligned} \quad (8.49)$$

## 8. Quantum phase estimation for noisy, small-scale experiments

Let us expand out our real and imaginary components of  $\mathfrak{T}$ :

$$\mathfrak{T}_r = \frac{\sum_{k=-K}^{K-1} (g_k^0 g_{k+1}^0 + g_k^1 g_{k+1}^1)}{\sum_{k=-K}^{K-1} (g_k^0)^2 + (g_k^1)^2}, \quad (8.50)$$

$$\mathfrak{T}_i = \frac{\sum_{k=-K}^{K-1} (g_k^0 g_{k+1}^1 - g_k^1 g_{k+1}^0)}{\sum_{k=-K}^{K-1} (g_k^0)^2 + (g_k^1)^2} \quad (8.51)$$

Then, we can calculate their derivatives as (recalling again that  $g_k^0 = g_{-k}^0$  and  $g_k^1 = g_{-k}^1$ )

$$\frac{\partial \mathfrak{T}_r}{\partial g_k^a} = \frac{2}{1 + \delta_{k,0}} \left[ \frac{(1 - \delta_{k,K}) g_{k+1}^a + g_{k-1}^a - 2 \mathfrak{T}_r g_k^a}{\sum_{k=-K}^{k+1} ((g_k^0)^2 + (g_k^1)^2)} \right] \quad (8.52)$$

$$\frac{\partial \mathfrak{T}_i}{\partial g_k^a} = \frac{2(-1)^a}{1 + \delta_{k,0}} \left[ \frac{(1 - \delta_{k,K}) g_{k+1}^{1-a} - g_{k-1}^{1-a} - 2 \mathfrak{T}_i g_k^a}{\sum_{k=-K}^{k+1} ((g_k^0)^2 + (g_k^1)^2)} \right]. \quad (8.53)$$

Substituting in for  $g_k^a$ , we find that everything precisely cancels when  $k \neq K!$

$$\frac{\partial \mathfrak{T}_r}{\partial g_k^0} = -\frac{\partial \mathfrak{T}_i}{\partial g_k^1} = -2\delta_{k,K} \frac{\cos((K+1)\phi)}{\sum_{k=-K}^{k+1} ((g_k^0)^2 + (g_k^1)^2)} \quad (8.54)$$

$$\frac{\partial \mathfrak{T}_i}{\partial g_k^0} = \frac{\partial \mathfrak{T}_r}{\partial g_k^1} = -2\delta_{k,K} \frac{\sin((K+1)\phi)}{\sum_{k=-K}^{k+1} ((g_k^0)^2 + (g_k^1)^2)}. \quad (8.55)$$

Our variance is then

$$\begin{aligned} \text{Var}(\phi) &= \left[ \frac{2}{(\mathfrak{T}_r^2 + \mathfrak{T}_i^2) \sum_{k=-K}^{k+1} ((g_k^0)^2 + (g_k^1)^2)} \right]^2 \times \\ &\quad \left\{ \text{Var}[g_K^0] (-\cos(\phi) \sin((K+1)\phi) + \sin(\phi) \cos((K+1)\phi))^2 \right. \\ &\quad \left. + \text{Var}[g_K^1] (\cos(\phi) \cos((K+1)\phi) + \sin(\phi) \sin((K+1)\phi))^2 \right\} \\ &= \left[ \frac{1}{K} \right]^2 \left\{ \text{Var}[g_K^0] \sin^2(K\phi) + \text{Var}[g_K^1] \cos^2(K\phi) \right\}. \end{aligned} \quad (8.56)$$

If  $g_K^a$  is estimated with  $N$  shots, we expect  $\text{Var}[g_K^0] = \frac{1}{N}$ , and

$$\text{Var}(\phi) \propto \frac{1}{K^2 N}. \quad (8.57)$$

As described in Sec. 8.3.1, for multi-round experiments we weight the least-squares inversion as per Eq. (8.27). This weighting adjusts the  $g_k^a$  values in Eqs. (8.52,8.53) so that  $\frac{\partial \phi}{\partial g_k^a}$  is no longer zero when  $k < K$ . The sum over  $k$  in Eq. (8.49) then lends an extra factor of  $K$  to the variance, reducing it to

$$\text{Var}(\phi) \propto \frac{1}{KN}. \quad (8.58)$$

## 8.C. Fourier representation for Bayesian updating

For simplicity, we first consider when the starting state is a simple eigenstate  $|\phi_j\rangle$ . After each multi-round experiment we would like to update the probability distribution  $P(\phi_j = \phi)$ , i.e.  $P_n(\phi) = \frac{P_{\mathbf{k},\beta}(\mathbf{m}|\phi)}{P(\mathbf{m})} P_{n-1}(\phi)$ . We will represent the  $2\pi$ -periodic probability distribution  $P_n(\phi)$  by a Fourier series with a small number of Fourier coefficients  $N_{\text{freq}}$  which are updated after each experiment, that is, we write

$$P(\phi) = p_0 + \sum_{j=1}^{N_{\text{freq}}-1} (p_{2j-1} \sin(j\phi) + p_{2j} \cos(j\phi)) \equiv \mathbf{p}. \quad (8.59)$$

We thus collect the coefficients as a  $N_{\text{freq}}$ -component vector  $\mathbf{p}$ . The Fourier representation has the advantage that integration is trivial i.e.  $\int_{-\pi}^{\pi} P(\phi) d\phi = 2\pi p_0$  so that the probability distribution is easily normalized. In addition, the current estimate  $\tilde{\phi}$  is easy to obtain:

$$\tilde{\phi} = \arg(\langle e^{i\phi} \rangle_P) = \arg(p_2 + ip_1). \quad (8.60)$$

Another observation is that the Holevo phase variance is easily obtained from this Fourier representation as

$$\text{Var}(P(\phi)) = \frac{1}{|\langle e^{i\phi} \rangle_P|^2} - 1 = \frac{1}{\pi^2(p_2^2 + p_1^2)} - 1. \quad (8.61)$$

Note that this is the Holevo phase variance of the posterior distribution of a single simulation instance. By comparison, in Fig. 8.4 we have calculated the same quantity over repeat simulations. However, in general we find the two to be equivalent.

The other advantage of the Fourier representation is that a single-round in an experiment is the application of a sparse matrix on  $\mathbf{p}$ . One has

### 8. Quantum phase estimation for noisy, small-scale experiments

$P(\phi) \rightarrow P_{k_r, \beta_r}(m_r|\phi)P(\phi) = \cos^2(k_r\phi/2 + \gamma/2)P(\phi)$ , where  $\gamma = \beta_r + m_r\pi$  which is equivalent to

$$\mathbf{p} \rightarrow \frac{1}{2}\mathbf{p} + \frac{1}{4}\cos(\gamma)M^0(k_r)\mathbf{p} + \frac{1}{4}\sin(\gamma)M^1(k_r)\mathbf{p}. \quad (8.62)$$

The coefficients of the update matrices  $M^{0,1}(k_r)$  can be simply calculated using the double angle formulae and employing

$$\begin{aligned} & \cos^2(k\phi/2 + \gamma/2) \cos(j\phi) \\ &= \frac{1}{2} \cos(j\phi) + \frac{1}{4} \cos(\gamma) (\cos((j+k)\phi) + \cos((j-k)\phi)) \\ &+ \frac{1}{4} \sin(\gamma) (\sin((j-k)\phi) - \sin((j+k)\phi)), \end{aligned} \quad (8.63)$$

and

$$\begin{aligned} & \cos^2(k\phi/2 + \gamma/2) \sin(j\phi) \\ &= \frac{1}{2} \sin(j\phi) + \frac{1}{4} \cos(\gamma) (\sin((j+k)\phi) + \sin((j-k)\phi)) \\ &+ \frac{1}{4} \sin(\gamma) (\cos((j+k)\phi) - \cos((j-k)\phi)). \end{aligned} \quad (8.64)$$

The matrices  $M^a(n)$  are then calculated from the above equations. When  $j > k$ , we have

$$\begin{aligned} [M^0(k)]_{2j+2k, 2j} &= 1, & [M^0(k)]_{2j-2k, 2j} &= 1, \\ [M^0(k)]_{2j+2k-1, 2j-1} &= 1, & [M^0(k)]_{2j-2k-1, 2j-1} &= 1, \\ [M^1(k)]_{2j+2k-1, 2j} &= -1, & [M^1(k)]_{2j-2k-1, 2j} &= 1, \\ [M^1(k)]_{2j+2k, 2j-1} &= 1, & [M^1(k)]_{2j-2k, 2j-1} &= -1, \end{aligned}$$

When  $j \leq k$ , we have to account for the sign change in  $\sin((j-k)\phi)$ :

$$\begin{aligned} [M^0(k)]_{j+2k, j} &= 1, & [M^0(k)]_{2k-2j, 2j} &= 1, \\ [M^0(k)]_{2k-2j-1, 2j-1} &= -1 \\ [M^0(k)]_{2k, 0} &= -2, & [M^0(k)]_{4k-1, 2k-1} &= 1 \\ [M^1(k)]_{2j+2k-1, 2j} &= -1, & [M^1(k)]_{2k-2j-1, 2j} &= -1, \\ [M^1(k)]_{2j+2k, 2j-1} &= 1, & [M^1(k)]_{2k-2j, 2j-1} &= -1, \\ [M^1(k)]_{2k-1, 0} &= 2, & [M^1(k)]_{4k-1, 2k} &= 1. \end{aligned}$$



For a multi-round experiment with  $R$  rounds, one thus applies such sparse matrices to the vector  $\mathbf{p}$   $R$  times. Note that each round with given  $k_r$  requires at most  $k_r$  more Fourier components, hence an experiment with at most  $K$  controlled- $U$  applications adds at most  $K$  Fourier components. Thus, when the total number of unitary rotations summed over all experiments  $K_{\text{tot}} = \sum_n \sum_r k_r > N_{\text{freq}}$ , our representation of the distribution is no longer accurate. When  $K_{\text{tot}} \leq N_{\text{freq}}$  on the other hand, it will be accurate.

### 8.C.1. Bayesian updating for multi-eigenvalue starting state

In this section we detail the method by which we store the distributions  $P_n^j(\phi_j)$  and  $P_n^{\text{red}}(\mathbf{A})$  of Eq. (8.31) and perform the Bayesian update of Eq. (8.30). We do so by representing the marginal probabilities  $P_n^j(\phi_j)$  by a Fourier series with a small number of Fourier coefficients which are updated after each experiment as shown in the previous section. We assume that there are most  $N_{\text{eig}}$  coefficients  $A_j > 0$  and thus  $N_{\text{eig}}$   $\phi_j$ .

From our independence assumption, individual updates of  $P^j(\phi_j)$  may be calculated by integrating out the other unknown variables in Eq. (8.30):

$$P_n^j(\phi_j) = \int \left( \prod_{l \neq j} d\phi_l P_{n-1}^l(\phi_l) \right) \int d\mathbf{A} P_{n-1}^{\text{red}}(\mathbf{A}) P_{\mathbf{k}, \beta}(\mathbf{m} | \phi, \mathbf{A}) P_{n-1}^j(\phi_j). \quad (8.65)$$

Expanding the conditional probability of Eq. (8.10) and rewriting leads to the form

$$P_n^j(\phi_j) = \frac{1}{P_{\mathbf{k}, \beta}(\mathbf{m})} \left( C + B_j \prod_r P_{k_r, \beta_r}(m_r | \phi_j) \right) P_{n-1}^j(\phi_j), \quad (8.66)$$

with

$$C = \sum_{k \neq j} B_k \int d\phi_k P_{n-1}^k(\phi_k) \prod_r P_{k_r, \beta_r}(m_r | \phi_k),$$

and  $B_j = \int d\mathbf{A} P_{n-1}^{\text{red}}(\mathbf{A}) A_j$ . Here we have used that  $\int d\phi_l P_{n-1}^l(\phi_l) = 1$ . One can concisely write  $B_j$  as the components of a vector  $\mathbf{B}$ . Computing Eq. (8.30) then involves creating an ‘update’ distribution for each  $\phi_j$ , calculating the integral of each distribution, and then forming the new distribution from a weighted sum from the ‘update’ distributions.

## 8. Quantum phase estimation for noisy, small-scale experiments

Calculating the distribution  $P_n^{\text{red}}(\mathbf{A})$  is complicated slightly by the restriction that  $\sum_j A_j = 1, A_j \geq 0$ , meaning that we cannot assume the distribution of individual  $A_j$  terms is uncorrelated. The marginal probability distribution equals

$$P_n^{\text{red}}(\mathbf{A}) = \frac{P_{n-1}^{\text{red}}(\mathbf{A})}{P_{\mathbf{k},\beta}(\mathbf{m})} \sum_j A_j \int d\phi_j P_{n-1}^j(\phi_j) \prod_r P_{k_r,\beta_r}(m_r|\phi_j). \quad (8.67)$$

or

$$P_n^{\text{red}}(\mathbf{A}) = \frac{P_{n-1}^{\text{red}}(\mathbf{A})}{P_{\mathbf{k},\beta}(\mathbf{m})} \mathbf{A} \cdot \mathbf{q}_{n-1}, \quad (8.68)$$

where the  $j$ th component  $(q_{n-1})_j$  is the integral

$$(q_{n-1})_j = \int d\phi_j P_{n-1}^j(\phi_j) \prod_r P_{k_r,\beta_r}(m_r|\phi_j). \quad (8.69)$$

As  $\mathbf{A}$  only enters our estimation through the vector  $\mathbf{B} = (B_0, \dots, B_{N_{\text{eig}}})$ , we only need approximate this value. Assuming we know the marginal probabilities  $P_n(\phi_j)$  for all experiments  $n = 1, \dots, N$ , we can estimate  $\mathbf{B}$  after all experiments by the maximum likelihood value  $\mathbf{A}^{(\text{max})}$ ,

$$\begin{aligned} \mathbf{A}_N^{(\text{max})} &= \underset{\mathbf{A}}{\text{argmax}} f(\mathbf{A}) \\ f(\mathbf{A}) &= \log \left( P_{\text{prior}}(\mathbf{A}) \prod_{n=1}^N \mathbf{A} \cdot \mathbf{q}_n \right) \\ &= \log(P_{\text{prior}}(\mathbf{A})) + \sum_{n=1}^N \log(\mathbf{A} \cdot \mathbf{q}_n). \end{aligned}$$

Evaluating this equation for up  $N = 1000$  experiments, taking  $N_{\text{freq}} = 10000$  frequency components of  $N_{\text{eig}} = 2$  eigenvalues takes less than a second on a laptop using a method such as sequential least-squares programming [323]. However, beyond this it becomes fairly computationally intensive. Thus, after  $N > 100$  experiments have been performed, we switch to a local optimization method. We determine the optimal  $\mathbf{B}_n$  after  $n$  experiments from its prior value  $\mathbf{B}_{n-1}$  via a single step of an approximate Newton's method, that is, we take

$$\mathbf{B}_n = \mathbf{B}_{n-1} - \Pi[\mathbf{H}^{-1}(f(\mathbf{B}_{n-1})) (\vec{\nabla} f)(\mathbf{B}_{n-1})].$$

where  $\vec{\nabla}f(\mathbf{A})$  is the first derivative of  $f$  at  $\mathbf{A}$  and  $\mathbf{H}$  is the Hessian matrix of  $f$ , i.e.  $H_{ij} = \partial_{A_i}\partial_{A_j}f(\mathbf{A})$ . Here  $\Pi[\mathbf{A}]$  is the projector onto the plane  $\sum_{j=0}^{N_{\text{eig}}} A_j = 1$  so that the update preserves the normalization. We have

$$\partial_{A_i}f(\mathbf{A}) = \frac{\partial_{A_i}P_{\text{prior}}(\mathbf{A})}{P_{\text{prior}}(\mathbf{A})} + \sum_{n=1}^N \frac{(q_n)_i}{\mathbf{A} \cdot \mathbf{q}_n}$$

We approximate the second term for each step as coming from only from the added term, i.e.

$$\vec{\nabla}f(\mathbf{B}_{n-1}) \approx \frac{\mathbf{q}_n}{\mathbf{B}_{n-1} \cdot \mathbf{q}_n}, \quad (8.70)$$

The Hessian equals

$$H_{ij}(f(\mathbf{A})) = - \sum_{n=1}^N \frac{(q_n)_i(q_n)_j}{(\mathbf{A} \cdot \mathbf{q}_n)^2}, \quad (8.71)$$

but we approximate this at the  $n$ th step

$$H_{ij}^{(n)}(f(\mathbf{B}_{n-1})) \approx H_{ij}^{(n-1)} - \frac{(q_n)_i(q_n)_j}{(\mathbf{B}_n \cdot \mathbf{q}_n)^2}. \quad (8.72)$$

This approximation allows  $H$  to be updated without summing over each experiment.

With the above implemented, we observe that our estimator can process data from  $N = 10,000$  experiments to estimate  $N_{\text{eig}} = 2$  eigenvalues with  $N = 20,000$  Fourier components within approximately two minutes on a laptop. Unfortunately, this method scales as  $N^2$ , as the number of frequencies required for accurate estimation grows as the total number of unitaries applied.

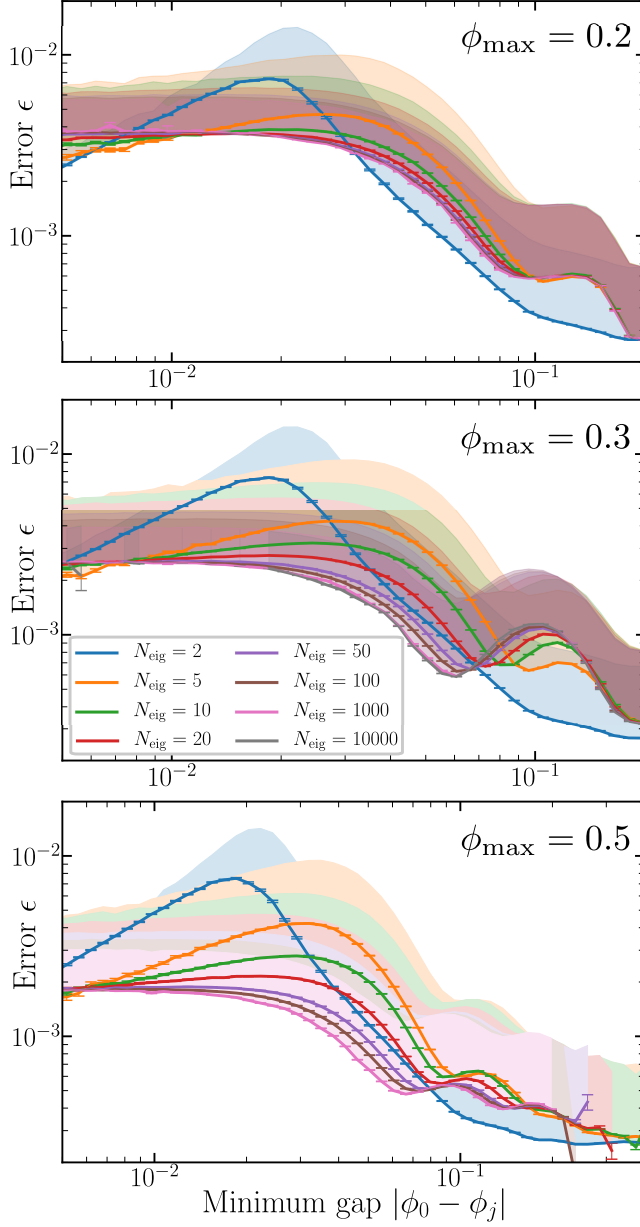
As the mean, variance and integration calculations only require the first few frequencies of the distribution, it may be possible to reduce this cost by finding approximation techniques for higher frequency components.

### 8.D. Convergence of the (noiseless) time-series analysis in case of multiple eigenvalues.

In this section we present an expansion of Fig. 8.6, namely Fig. 8.11, by drawing the spurious eigenvalues  $\phi_j$  from a range closer to the target eigenvalue  $\phi_0$ . This negates the drop in estimation error observed in Fig. 8.6 that was caused by the majority of eigenvalues lying in region (c) of Fig. 8.5. We observe that for certain gaps  $\delta$ , multiple eigenvalues confined to a thin region  $[\delta, \phi_{\max}]$  can have a worse effect on our ability to estimate  $\phi_0$  than that of a single eigenvalue at  $\delta$ . However, this loss in accuracy does not get critically worse with the addition of more eigenvalues. Neither is it worse than the worst-possible estimation with two eigenvalues.

### 8.E. Details of realistic simulation

In this Appendix we give details of the method for the realistic noisy circuit simulation of Sec. 8.5.2. Our density-matrix simulator is fairly limited in terms of qubit number, and so we opt to simulate  $H_4$  in the STO-3G basis. This molecule has 8 spin orbitals and thus requires 9 qubits for the QPE simulation (with the additional qubit being the ancilla). We choose 10 rectangular molecular geometries for the  $H_4$  system, parametrized by a horizontal distance  $d_x$  and a vertical distance  $d_y$  (i.e. the four H atoms are in the positions  $(\pm d_x/2, \pm d_y/2, 0)$ ). We calculate the Hartree-Fock and full-CI solutions to the ground state using the psi4 package [321] with the openfermion interface [298]. This allows to calculate the true ground state energy  $E_0$  for each geometry, and the overlap  $A_0$  between the ground state and the Hartree-Fock state, which we choose as our starting state  $|\Psi\rangle$ . Due to symmetry and particle number conservation,  $|\Psi\rangle$  has non-zero overlap with only 8 eigenstates of the full-CI solution, separated from the ground state by a minimum gap  $\delta$ . (When  $d_x = d_y$ , the true ground state of  $H_4$  is actually orthogonal to the Hartree-Fock state, and so we do not include any such geometries in our calculation.) The full error in our calculation of the energy (at a fixed geometry) is then a combination of three separate contributions: basis set error (i.e. from the choice of orbitals), Trotter error, and the estimator error studied in this chapter (which includes error from experimental noise). The Trotter error  $\epsilon_{\text{Trotter}}$  is reasonably large due to our use of only the first-order Suzuki-Trotter



**Figure 8.11.:** Variations of Fig. 8.6, but with eigenstates  $\phi_j$  drawn from a range  $[0, \phi_{\max}]$  as labeled. Error bars are 95% confidence intervals for each point, shaded regions denote top 2σ interval (i.e. region containing the top 2.5% – 50% of the population).

## 8. Quantum phase estimation for noisy, small-scale experiments

approximation  $U = \prod_i e^{-iH_i t} \approx e^{-i\mathcal{H}t}$ . Higher-order Suzuki-Trotter expansions require longer quantum circuits, which in turn increase the estimator error from experimental noise. Balancing these two competing sources of error is key to obtaining accurate calculations and a clear target for future study. In Tab. 8.2, we list some parameters of interest for each studied geometry. We normalize the gap and the Trotter error by the Frobenius norm  $\|\mathcal{H}\|_F = \sqrt{\text{Trace}[\mathcal{H}^\dagger \mathcal{H}]/2^{N_{\text{sys}}}}$ , as we chose an evolution time  $t = 1/\|\mathcal{H}\|_F$ , making this the relevant scale for comparison with scaling laws and errors calculated in the text.

$d_x$ [Å]	$d_y$ [Å]	$E_0$	$A_0$	$\delta/\ \mathcal{H}\ _F$	$\epsilon_{\text{Trotter}}/\ \mathcal{H}\ _F$
0.4	0.5	-0.26	0.98	0.09	$3.7 \times 10^{-4}$
0.6	0.7	-1.46	0.94	0.17	$3.1 \times 10^{-3}$
0.8	0.9	-1.84	0.88	0.24	0.016
1.0	1.1	-1.96	0.80	0.23	0.017
1.2	1.3	-1.98	0.71	0.18	0.013
1.6	1.7	-1.94	0.55	0.09	$6.0 \times 10^{-3}$
0.2	1.8	0.32	0.996	0.67	$2.0 \times 10^{-4}$
0.4	1.6	-1.80	0.993	1.14	$2.6 \times 10^{-3}$
0.6	1.4	-2.15	0.98	1.27	0.014
0.8	1.2	-2.09	0.96	0.73	0.021

**Table 8.2.:** Parameters of the  $H_4$  geometries used in the text. Terms are described in App.8.E.  $\|\mathcal{H}\|_F = \sqrt{\text{Trace}[\mathcal{H}^\dagger \mathcal{H}]/2^{N_{\text{sys}}}}$ .

### 8.E.1. Error model and error parameters

Throughout this chapter we simulate circuits using an error model of superconducting qubits first introduced in Ref. [212]. This captures a range of different error channels with parameters either observed in experimental data or estimated via theory calculations. All error channels used are listed in Tab. 8.3, and we will now describe them in further detail.

Transmon qubits are dominated primarily by decoherence, which is captured via  $T_1$  and  $T_2$  channels [23]. Typical  $T_1$  and  $T_2$  times in state-of-the-art devices are approximately  $10 - 100 \mu\text{s}$ . As other error parameters are derived from experimental results on a device with  $T_1 = T_2 \approx 30 \mu\text{s}$ , we take these as a base set of parameters [182, 183]. Single-qubit gates in transmon qubits incur slight additional dephasing due to inaccuracies or

Parameter	Symbol	Standard Value	Scaling
Qubit relaxation time	$T_1$	30 $\mu$ s	$\lambda$
Qubit dephasing time	$T_2$	30 $\mu$ s	$\lambda$
Single-qubit gate time	$T_{\text{sq}}$	20 ns	1
Two-qubit gate time	$T_{2\text{q}}$	40 ns	1
In-axis rotation error	$p_{\text{axis}}$	$10^{-4}$	$\lambda^{-1}$
In-plane rotation error	$p_{\text{plane}}$	$5 \times 10^{-4}$	$\lambda^{-1}$
Incoherent flux noise	$A$	$(1\mu\Phi_0)^2$	$\lambda^{-1}$
Measurement time	$T_{\text{meas}}$	300 ns	1
Depletion time	$T_{\text{dep}}$	300 ns	1
Readout infidelity	$\epsilon_{\text{RO}}$	$5 \times 10^{-3}$	$\lambda^{-1}$
Measurement induced decay	$p_{\text{d,i}}, p_{\text{d,f}}$	0.005, 0.0015	$\lambda^{-1}$

**Table 8.3.:** Standard parameters of error models used in density matrix simulation. Table adapted from Ref. [212] with all parameters taken from therein (with the exception of the  $1/f$  flux noise, which is made incoherent as described in text).

fluctuations in microwave pulses. We assume such dephasing is Markovian, in which case it corresponds to a shrinking of the Bloch sphere along the axis of rotation by a value  $1 - p_{\text{axis}}$ , and into the perpendicular plane by a value  $1 - p_{\text{plane}}$ . We take typical values for these parameters as  $p_{\text{axis}} = 10^{-4}$ ,  $p_{\text{plane}} = 5 \cdot 10^{-4}$  [212].

Two-qubit gates in transmon qubits incur dephasing due to  $1/f$  flux noise. Assuming that the phase in an ideal C-Phase gate  $G = \text{diag}(1, 1, 1, e^{i\phi})$  is controlled by adjusting the time of application, this suggests a model for the applied gate which is

$$G(\delta_{\text{flux}}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{i\delta_{\text{flux}}\phi} & 0 \\ 0 & 0 & 0 & e^{i(1+\delta_{\text{flux}}/2)\phi} \end{pmatrix}, \quad (8.73)$$

where  $\delta_{\text{flux}}$  is drawn from a normal distribution around 0 with standard deviation  $\sigma_{\text{flux}}$ . One can estimate  $\sigma_{\text{flux}} \approx 0.01$  rad for a typical gate length of 40 ns [212]. The noise is in general non-Markovian, as  $\delta_{\text{flux}}$  fluctuates on longer timescale than a single gate. However, to make the simulation

## 8. Quantum phase estimation for noisy, small-scale experiments

tractable, we approximate it as Markovian. The Pauli transfer matrix of this averaged channel [324] reads

$$\Lambda[G] = \int d\delta_{\text{flux}} P(\delta_{\text{flux}}) \Lambda[G(\delta_{\text{flux}})], \quad (8.74)$$

where the Pauli transfer matrix of a channel  $G$  is given by  $\Lambda[G]_{i,j} = \text{Tr}[\sigma_i G \sigma_j]$ .

During qubit readout, we assume that the qubit is completely dephased and projected into the computational basis. We then allow for a  $T_{\text{meas}} = 300$  ns period of excitation and de-excitation (including that from  $T_1$ -decay), during which the qubit state is copied onto a classical bit. This copying is also assumed to be imperfect, with a probability  $\epsilon_{\text{RO}}$  of returning the wrong result. The qubit then has an additional  $T_{\text{dep}} = 300$  ns waiting period before it may participate in gates again (to allow resonator depletion [182]), over which additional excitation and de-excitation may occur. Though simple, this description is an accurate model of experimental results. Typically experiments do not observe measurement-induced excitation to the  $|1\rangle$  state, but do observe measurement-induced decay [212]. Typical values of such decay are 0.005 prior to the copy procedure, and 0.015 after.

Though reasonably accurate, this error model does fail to capture some details of real experimental systems. In particular, we do not include leakage to the  $|2\rangle$  state, which is a dominant source of two-qubit gate error. Furthermore, we have not included cross-talk between qubits.

To study the effect of changing noise levels while staying as true as possible to our physically-motivated model, we scale our noise parameters by a dimensionless parameter  $\lambda$  such that the contribution from each error channel to the simulation remains constant. In Tab. 8.3 we show the power of  $\lambda$  that each error term is multiplied by during this scaling. We report  $T_{\text{err}} := T_1 = T_2$  in the main text instead of  $\lambda$  to make connection to parameters regularly reported in experimental works.



# Bibliography

- [1] K. von 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).
- [2] C. L. Kane and E. J. Mele,  *$Z_2$  Topological Order and the Quantum Spin Hall Effect*, Phys. Rev. Lett. **95**, 146802 (2005).
- [3] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, *A topological Dirac insulator in a quantum spin Hall phase*, Nature **452**, 970-974 (2008).
- [4] S. Murakami, *Phase transition between the quantum spin Hall and insulator phases in 3D: emergence of a topological gapless phase*, New J. Phys. **9**, 356 (2007).
- [5] S-Y. Xu, I. Belopolski, N. Alidoust, M. Neupane, G. Bian, C. Zhang, R. Sankar, G. Chang, Z. Yuan, C-C. Lee, S-M. Huang, H. Zheng, J. Ma, D. S. Sanchez, B. Wang, A. Bansil, F. Chou, P. P. Shibayev, H. Lin, S. Jia, and M. Z. Hasan, *Discovery of a Weyl fermion semimetal and topological Fermi arcs*, Science **349** (6248), 613-617 (2015).
- [6] F. D. M. Haldane, *Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the "Parity Anomaly"*, Phys. Rev. Lett. **61**, 2015 (1988).
- [7] M. Konig, S. Wiedmann, C. Brune, A. Roth, H. Buhmann, L. W. Molenkamp, X-L. Qi, and S-C. Zhang, *Quantum Spin Hall Insulator State in HgTe Quantum Wells*, Science **318** (5851), 766-770 (2007).
- [8] G. Moore, and N. Read, *Nonabelions in the fractional quantum hall effect*, Nuc. Phys. B **360** (2-3), 362-396 (1991).
- [9] S. Aaronson, and A. Arkhipov, *The computational complexity of linear optics*, Proc. 43rd Ann. ACM Symp. Theo. Comp. 333-342 (2011).

## Bibliography

- [10] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, *Characterizing quantum supremacy in near-term devices*, Nat. Phys. **14**, 595-600 (2018).
- [11] K. Temme, S. Bravyi, and J. M. Gambetta, *Error Mitigation for Short-Depth Quantum Circuits* Phys. Rev. Lett. **119**, 180509 (2017).
- [12] Y. Li, and S. C. Benjamin, *Efficient Variational Quantum Simulator Incorporating Active Error Minimization*, Phys. Rev. X **7**, 021050 (2017).
- [13] S. Endo, S. C. Benjamin, and Y. Li, *Practical Quantum Error Mitigation for Near-Future Applications*, Phys. Rev. X **8**, 031027 (2018).
- [14] S. McArdle, X. Yuan, and S. C. Benjamin, *Error mitigated digital quantum simulation*, ArXiv:1807.02467 (2018).
- [15] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien, *Low-cost error mitigation by symmetry verification*, Phys. Rev. A **98**, 062339 (2018).
- [16] D. Loss and D. P. DiVincenzo, *Quantum computation with quantum dots*, Phys. Rev. A **57**, 120 (1998).
- [17] R. P. Feynman, *Simulating Physics with Computers*, Int. J. Theo. Phys. **21**, 6-7 (1982).
- [18] D. P. DiVincenzo, *The Physical Implementation of Quantum Computation*, Fortschr. Phys. **48**, 771-783 (2000).
- [19] F. Bloch, *Nuclear Induction*, Phys. Rev. **70**, 460 (1946).
- [20] W. G. Unruh, *Maintaining coherence in quantum computers*, Phys. Rev. A **51** (2), 992-997 (1995).
- [21] R. Landauer, *Is quantum mechanics useful?* Proc. Roy. Soc. A **353** (1703), 367-376 (1995).
- [22] A. Y. Kitaev, *Quantum measurements and the Abelian Stabilizer Problem*, ArXiv:quant-ph/9511026 (1995).
- [23] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press (2000).

- [24] C. Kittel, *Introduction to Solid State Physics*, 8th edition, Wiley (2004).
- [25] N. W. Ashcroft, and N. D. Mermin, *Solid State Physics*, 33rd edition, Holt, Rinehart and Winston (1976).
- [26] F. Bloch, *Über die Quantenmechanik der Elektronen in Kristallgittern*, Z. Phys. **52** (7-8), 555-600 (1929).
- [27] X. G. Wen, *Topological Orders in Rigid States*, Int. J. Mod. Phys. B **4**, 239 (1990).
- [28] P. W. Anderson, *Absence of Diffusion in Certain Random Lattices*, Phys. Rev. **109**, 1492 (1958)
- [29] N. F. Mott and R. Peierls, *Discussion of the paper by de Boer and Verwey*, Proc. Roy. Soc. **49** (4S), 72 (1937).
- [30] L. D. Landau, *On the Theory of Phase Transitions*, Zh. Eksp. Teor. Fiz. **7**, 19-37 (1937)
- [31] A. Y. Kitaev, *Unpaired Majorana fermions in quantum wires*, Phys.-Usp. **44**, 131 (2001).
- [32] P.-G. de Gennes, *Superconductivity of Metals and Alloys*, 2nd edition, Advanced Book Program, Perseus Books (1999).
- [33] N. N. Bogoljubov, V. V. Tolmachov, D. V. Širkov, *A New Method in the Theory of Superconductivity*, Fort. Phys. **6**, 11-12 (1958).
- [34] M. Leijnse, and K. Flensberg, *Introduction to topological superconductivity and Majorana fermions*, Semi. Sci. Tech. **27**, 12 (2012).
- [35] V. Mourik, K. Zuo, S. M. Frolov, S. R. Plissard, E. P. A. M. Bakkers, and L. P. Kouwenhoven, *Signatures of Majorana Fermions in Hybrid Superconductor-Semiconductor Nanowire Devices*, Science **336** (6804), 1003-1007 (2012).
- [36] 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).
- [37] Y. Oreg, G. Refael, and F. von Oppen, *Helical Liquids and Majorana Bound States in Quantum Wires*, Phys. Rev. Lett. **105**, 177002 (2010).

- [38] R. M. Lutchyn, E. P. A. M. Bakkers, L. P. Kouwenhoven, P. Krogstrup, C. M. Marcus and Y. Oreg, *Majorana zero modes in superconductor-semiconductor heterostructures*, Nat. Rev. Mat. **3**, 52-68 (2018).
- [39] S. Nadj-Perge, I. K. Drozdov, J. Li, H. Chen, S. Jeon, J. Seo, A. MacDonald, B. A. Bernevig, A. Yazdani, *Observation of Majorana fermions in ferromagnetic atomic chains on a superconductor*, Science **346** (6209), 602-607 (2014).
- [40] , M. T. Deng, S. Vaitiekėnas, E. B. Hansen, J. Danon, M. Leijnse, K. Flensberg, J. NygøArd, P. Krogstrup, and C. M. Marcus, *Majorana bound state in a coupled quantum-dot hybrid-nanowire system*, Science **354** (6319), 1557-1562 (2016).
- [41] D. I. Pikulin, J. P. Dahlhaus, M. Wimmer, H. Schomerus, and C. W. J. Beenakker, *A zero-voltage conductance peak from weak antilocalization in a Majorana nanowire*, New J. Phys. **14**, 125011 (2012).
- [42] G. Kells, D. Meidan, and P. W. Brouwer, *Near-zero-energy end states in topologically trivial spin-orbit coupled superconducting nanowires with a smooth confinement*, Phys. Rev. B **86**, 100503(R) (2012).
- [43] T. Hyart, B. van Heck, I. C. Fulga, M. Burrello, A. R. Akhmerov, and C. W. J. Beenakker, *Flux-controlled quantum computation with Majorana fermions*, Phys. Rev. B **88**, 035121 (2013).
- [44] M. Wimmer, A. R. Akhmerov, J. P. Dahlhaus, and C. W. J. Beenakker, *Quantum point contact as a probe of a topological superconductor*, New J. Phys. **13**, 053106 (2011).
- [45] J. D. Sau, D. J. Clarke, and S. Tewari, *Controlling non-Abelian statistics of Majorana fermions in semiconductor nanowires*, Phys. Rev. B **84**, 094505 (2011).
- [46] 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).
- [47] S. B. Bravyi and A. Y Kitaev, *Fermionic Quantum Computation*, Ann. Phys. **298** (1), 210-266 (2002).
- [48] S. Bravyi and A. Kitaev, *Universal quantum computation with ideal Clifford gates and noisy ancillas*, Phys. Rev. A **71**, 022316 (2005).

- [49] B. Bradlyn, J. Cano, Z. Wang, M. G. Vergniory, C. Felser, R. J. Cava, and B. A. Bernevig, *Beyond Dirac and Weyl fermions: Unconventional quasiparticles in conventional crystals*, Science **353** (6299), aaf5037 (2016).
- [50] H. Weyl, *Elektron und Gravitation. I*, Z. Phys. **56**, 330 (1929).
- [51] P. Huhtala and G. E. Volovik, *Fermionic microstates within the Painlevé-Gullstrand black hole* JETP **94**, 853 (2002).
- [52] T. Kawarabayashi, Y. Hatsugai, T. Morimoto, and H. Aoki, *Generalization of chiral symmetry for tilted Dirac cones*, Int. J. Mod. Phys. Conf. Series **11**, 145 (2012).
- [53] M. Trescher, B. Sbierski, P. W. Brouwer, and E. J. Bergholtz, *Quantum transport in Dirac materials: Signatures of tilted and anisotropic Dirac and Weyl cones*, Phys. Rev. B **91**, 115135 (2015).
- [54] Y. Sun, S.-C. Wu, M. N. Ali, C. Felser, and B. Yan, *Prediction of Weyl semimetal in orthorhombic  $\text{MoTe}_2$* , Phys. Rev. B **92**, 161107(R) (2015).
- [55] Y. Xu, F. Zhang, and C. Zhang, *Structured Weyl Points in Spin-Orbit Coupled Fermionic Superfluids*, Phys. Rev. Lett. **115**, 265304 (2015).
- [56] A. A. Soluyanov, D. Gresch, Z. Wang, Q.-S. Wu, M. Troyer, X. Dai, and B. A. Bernevig, *Type-II Weyl semimetals*, Nature **527**, 495 (2015).
- [57] G. Autés, D. Gresch, A. A. Soluyanov, M. Troyer, and O. V. Yazyev, *Robust Type-II Weyl Semimetal Phase in Transition Metal Diphosphides  $\text{XP}_2$  ( $X=\text{Mo}, \text{W}$ )*, Phys. Rev. Lett. **117**, 066402 (2016).
- [58] K. Koepnick, D. Kasinathan, D. V. Efremov, S. Khim, S. Borisenko, B. Büchner, and J. van den Brink,  *$\text{TaIrTe}_4$ : A ternary type-II Weyl semimetal*, Phys. Rev. B Rap. Com. **93**, 201101 (2016).
- [59] L. Muechler, A. Alexandradinata, T. Neupert, and R. Car, *Topological Nonsymmorphic Metals from Band Inversion*, Phys. Rev. X **6**, 041069 (2016).
- [60] S.-Y. Xu, N. Alidoust, G. Chang, H. Lu, B. Singh, I. Belopolski, D. S. Sanchez, X. Zhang, G. Bian, H. Zheng, M.-A. Hsuanu, Y. Bian, S.-M. Huang, C.-H. Hsu, T.-R. Chang, H.-T. Jeng, A. Bansil, T. Neupert, V. N. Strocov, H. Lin, S. Jia, and M. Z. Hasan, *Discovery of Lorentz-violating type II Weyl fermions in  $\text{LaAlGe}$* , Sci. Adv. **3** (6), e1603266 (2017).

## Bibliography

- [61] M. V. Berry, *Quantal phase factors accompanying adiabatic changes*, Proc. Roy. Soc. A **392** (1802), 45-57 (1984).
- [62] A. A. Zyuzin and A. A. Burkov, *Topological response in Weyl semimetals and the chiral anomaly* Phys. Rev. B **86**, 115133 (2012).
- [63] A. A. Burkov, *Chiral anomaly and transport in Weyl metals*, J. Phys. Condens. Matter **27**, 113201 (2015).
- [64] H. B. Nielsen and M. Ninomiya, *The Adler-Bell-Jackiw anomaly and Weyl fermions in a crystal*, Phys. Lett. B **130**, 389 (1983).
- [65] M. M. Vazifeh and M. Franz, *Electromagnetic Response of Weyl Semimetals*, Phys. Rev. Lett. **111**, 027201 (2013).
- [66] B. M. Terhal, *Quantum error correction for quantum memories*. Rev. Mod. Phys. **87**, 307-346 (2015).
- [67] P. W. Shor, *Scheme for reducing decoherence in quantum computer memory*, Phys. Rev. A **52**, R2493 (1995).
- [68] D. Gottesman, *Stabilizer Codes and Quantum Error Correction*. Ph.D. thesis, Caltech (1997).
- [69] J. F. Ziegler, H. W. Curtis, H. P. Muhlfeld, C. J. Montrose, B. Chin, M. Nicewicz, C. A. Russell, W. Y. Wang, L. B. Freeman, P. Hosier, L. E. LaFave, J. L. Walsh, J. M. Orro, G. J. Unger, J. M. Ross, T. J. O’Gorman, B. Messina, T. D. Sullivan, A. J. Sykes, H. Yourke, T. A. Enger, V. Tolat, T. S. Scott, A. H. Taber, R. J. Sussman, W. A. Klein, and C. W. Wahauss, *IBM experiments in soft fails in computer electronics (1978-1994)*, IBM J. Res. Dev. **40**, 1 (1996).
- [70] J. Park, *The concept of transition in quantum mechanics*, Found. Phys. **1** (1), 23-33 (1970).
- [71] W. Wootters, and W. Zurek, *A single quantum cannot be cloned*, Nature **299**, 802-803 (1982).
- [72] D. Dieks, *Communication by EPR devices*, Phys. Lett. A **92** (6), 271-272 (1982).
- [73] A. R. Calderbank, and P. W. Shor, *Good quantum error-correcting codes exist*, Phys. Rev. A **54**, 1098 (1996).

- [74] J-P. Tillich, and G. Zemor, *Quantum LDPC codes with positive rate and minimum distance proportional to  $n^{\frac{1}{2}}$* , IEEE Int. Symp. Inf. Theo., 799-803 (2009).
- [75] H. Bombin, and M. A. Martin-Delgado, *Topological Quantum Distillation*, Phys. Rev. Lett. **97**, 180501 (2006).
- [76] D. Aharonov, and M. Ben-Or, *Fault Tolerant Quantum Computation with Constant Error*, ArXiv:quant-ph/9611025 (1996).
- [77] R. Versluis, S. Poletto, N. Khammassi, B. Tarasinski, N. Haider, D. J. Michalak, A. Bruno, K. Bertels, and L. DiCarlo, *Scalable Quantum Circuit and Control for a Superconducting Surface Code*, Phys. Rev. App. **8**, 034021 (2017).
- [78] S. Bravyi, and R. Konig, *Classification of Topologically Protected Gates for Local Stabilizer Codes*, Phys. Rev. Lett. **110**, 170503 (2013).
- [79] H. Bombin, and M. A. Martin-Delgado, *Optimal resources for topological two-dimensional stabilizer codes: Comparative study*, Phys. Rev. A **76**, 012305 (2007).
- [80] H. Bombin, and M. A. Martin-Delgado, *Exact topological quantum order in  $D=3$  and beyond: Branyons and brane-net condensates*, Phys. Rev. B **75**, 075103 (2007).
- [81] M-H. Hsieh, and F. Le Gall, *NP-hardness of decoding quantum error-correction codes*, Phys. Rev. A **83**, 052331 (2011).
- [82] S. Bravyi, M. Suchara, and A. Vargo, *Efficient algorithms for maximum likelihood decoding in the surface code*, Phys. Rev. A **90**, 032326 (2014).
- [83] P. Baireuther, M. D. Caio, B. Criger, C. W. J. Beenakker, and T. E. O'Brien, *Neural network decoder for topological color codes with circuit level noise*, New J. Phys. **21**, 013003 (2019).
- [84] P. W. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, SIAM J. Comp. **26** (5), 1484-1509 (1997).
- [85] A. W. Harrow, A. Hassidim, and S. Lloyd, *Quantum Algorithm for Linear Systems of Equations*, Phys. Rev. Lett. **103**, 150502 (2009).

- [86] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, *The theory of variational hybrid quantum-classical algorithms*
- [87] 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**, 242-246 (2017).
- [88] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, *Charge-insensitive qubit design derived from the Cooper pair box*, Phys. Rev. A **76**, 042319 (2007).
- [89] E. T. Jaynes, and F. W. Cummings, *Comparison of quantum and semiclassical radiation theories with application to the beam maser*, Proc. IEEE. **51** (1), 89-109 (1963).
- [90] J. Ghosh, A. G. Fowler, J. M. Martinis, and M. R. Geller, *Understanding the effects of leakage in superconducting quantum-error-detection circuits*, Phys. Rev. A **88**, 062329 (2013).
- [91] A. G. Fowler, *Coping with qubit leakage in topological codes*, Phys. Rev. A **88**, 042308 (2013).
- [92] C. Herring, *Accidental Degeneracy in the Energy Bands of Crystals*, Phys. Rev. **52**, 365 (1937).
- [93] G. E. Volovik, *The Universe in a Helium Droplet*, Oxford University Press (2003).
- [94] M. O. Goerbig, J.-N. Fuchs, G. Montambaux, F. Piechon, *Tilted anisotropic Dirac cones in quinoid-type graphene and  $\alpha$ -(BEDT-TTF) $_2$ I $_3$* , Phys. Rev. B **78**, 045415 (2008).
- [95] L. Huang, T. M. McCormick, M. Ochi, Z. Zhao, M. Suzuki, R. Arita, Y. Wu, D. Mou, H. Cao, J. Yan, N. Trivedi, and A. Kaminski, *Spectroscopic evidence for a type II Weyl semimetallic state in MoTe $_2$* , Nat. Mater. **15**, 1155-1160 (2016).
- [96] S.-Y. Xu, N. Alidoust, G. Chang, H. Lu, B. Singh, I. Belopolski, D. Sanchez, X. Zhang, G. Bian, H. Zheng, M.-A. Husanu, Y. Bian, S.-M. Huang, C.-H. Hsu, T.-R. Chang, H.-T. Jeng, A. Bansil, V. N. Strocov, H. Lin, S. Jia, and M. Z. Hasan, *Discovery of Lorentz-violating type II Weyl fermions in LaAlGe*, Sci. Adv. **3** (6), e1603266 (2017).



- [97] K. Deng, G. Wan, P. Deng, K. Zhang, S. Ding, E. Wang, M. Yan, H. Huang, H. Zhang, Z. Xu, J. Denlinger, A. Fedorov, H. Yang, W. Duan, H. Yao, Y. Wu, S. Fan, H. Zhang, X. Chen, and S. Zhou, *Experimental observation of topological Fermi arcs in type-II Weyl semimetal  $\text{MoTe}_2$* , Nat. Phys. **12**, 1105-1110 (2016).
- [98] J. Jiang, Z. K. Liu, Y. Sun, H. F. Yang, R. Rajamathi, Y. P. Qi, L. X. Yang, C. Chen, H. Peng, C.-C. Hwang, S. Z. Sun, S.-K. Mo, I. Vobornik, J. Fujii, S. S. P. Parkin, C. Felser, B. H. Yan, and Y. L. Chen, *Signature of type-II Weyl semimetal phase in  $\text{MoTe}_2$* , Nat. Comm. **8**, 13973 (2017).
- [99] A. Liang, J. Huang, S. Nie, Y. Ding, Q. Gao, C. Hu, S. He, Y. Zhang, C. Wang, B. Shen, J. Liu, P. Ai, L. Yu, X. Sun, W. Zhao, S. Lv, D. Liu, C. Li, Y. Zhang, Y. Hu, Y. Xu, L. Zhao, G. Liu, Z. Mao, X. Jia, F. Zhang, S. Zhang, F. Yang, Z. Wang, Q. Peng, H. Weng, X. Dai, Z. Fang, Z. Xu, C. Chen, and X. J. Zhou, *Electronic Evidence for Type II Weyl Semimetal State in  $\text{MoTe}_2$* , ArXiv:1604.01706 (2016).
- [100] N. Xu, Z. J. Wang, A. P. Weber, A. Magrez, P. Bugnon, H. Berger, C. E. Matt, J. Z. Ma, B. B. Fu, B. Q. Lv, N. C. Plumb, M. Radovic, E. Pomjakushina, K. Conder, T. Qian, J. H. Dil, J. Mesot, H. Ding, and M. Shi, *Discovery of Weyl semimetal state violating Lorentz invariance in  $\text{MoTe}_2$* , ArXiv:1604.02116 (2016).
- [101] C. Wang, Y. Zhang, J. Huang, S. Nie, G. Liu, A. Liang, Y. Zhang, B. Shen, J. Liu, C. Hu, Y. Ding, D. Liu, Y. Hu, S. He, L. Zhao, L. Yu, J. Hu, J. Wei, Z. Mao, Y. Shi, X. Jia, F. Zhang, S. Zhang, F. Yang, Z. Wang, Q. Peng, H. Weng, X. Dai, Z. Fang, Z. Xu, C. Chen, and X. J. Zhou, *Observation of Fermi arc and its connection with bulk states in the candidate type-II Weyl semimetal  $\text{WTe}_2$* , Phys. Rev. B **94**, 241119(R) (2016).
- [102] N. Harrison and S. E. Sebastian, *Dirac nodal pockets in the antiferromagnetic parent phase of  $\text{FeAs}$  superconductors*, Phys. Rev. B **80**, 224512 (2009).
- [103] M.-C. Chang and M.-F. Yang, *Chiral magnetic effect in the absence of Weyl node*, Phys. Rev. B **92**, 205201 (2015).
- [104] A. A. Zyuzin and R. P. Tiwari, *Intrinsic Anomalous Hall Effect in Type-II Weyl Semimetals*, JETP **103** (11), 717-722 (2016).

## Bibliography

- [105] Z. Yu, Y. Yao, and S. A. Yang, *Predicted Unusual Magnetoresponse in Type-II Weyl Semimetals* Phys. Rev. Lett. **117**, 077202 (2016).
- [106] C. M. Wang, H.-Z. Lu, and S.-Q. Shen, *Anomalous Phase Shift of Quantum Oscillations in 3D Topological Semimetals*, Phys. Rev. Lett. **117**, 077201 (2016).
- [107] L. Onsager, *Interpretation of the de Haas-van Alphen effect*, Phil. Mag. **43**, 1006 (1952).
- [108] D. Shoenberg, *Magnetic Oscillations in Metals*, Cambridge University Press (1984).
- [109] C. W. J. Beenakker, *Colloquium: Andreev reflection and Klein tunneling in graphene*, Rev. Mod. Phys. **80**, 1337 (2008).
- [110] R. Nandkishore and L. Levitov, *Common-path interference and oscillatory Zener tunneling in bilayer graphene p-n junctions*, PNAS **108**, 14021 (2011).
- [111] L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3, Elsevier, Oxford (1977).
- [112] E. I. Blount, *Bloch Electrons in a Magnetic Field*, Phys. Rev. **126**, 1636 (1962).
- [113] S. Keppeler and R. Winkler, *Anomalous Magneto-oscillations and Spin Precession*, Phys. Rev. Lett. **88**, 046401 (2002).
- [114] C. W. Groth, M. Wimmer, A. R. Akhmerov, and X. Waintal, *Kwant: a software package for quantum transport*, New J. Phys. **16**, 063065 (2014).
- [115] Z. Zhu, X. Lin, J. Liu, B. Fauqué, Q. Tao, C. Yang, Y. Shi, and K. Behnia, *Quantum Oscillations, Thermoelectric Coefficients, and the Fermi Surface of Semimetallic WTe<sub>2</sub>*, Phys. Rev. Lett. **114**, 176601 (2015).
- [116] B. A. Bernevig, *It's been a Weyl coming*, Nat. Phys. **11**, 698 (2015).
- [117] D. Ciudad, *Weyl fermions: Massless yet real*, Nat. Mat. **14**, 863 (2015).
- [118] S. Jia, S.-Y. Xu, and M. Z. Hasan, *Weyl semimetals, Fermi arcs and chiral anomalies*, Nat. Mat. **15**, 1140 (2016).

- [119] A. A. Burkov, *Topological semimetals*, Nat. Mat. **15**, 1145 (2016).
- [120] S. Rao, *Weyl semi-metals: A short review*, J. Indian Inst. Science **96**, 145 (2016).
- [121] D. Kharzeev, *The Chiral Magnetic Effect and anomaly-induced transport*, Progr. Part. Nucl. Phys. **75**, 133 (2014).
- [122] Heon-Jung Kim, Ki-Seok Kim, J.-F. Wang, M. Sasaki, N. Satoh, A. Ohnishi, M. Kitaura, M. Yang, and L. Li, *Dirac versus Weyl fermions in topological insulators: Adler-Bell-Jackiw anomaly in transport phenomena*, Phys. Rev. Lett. **111**, 246603 (2013).
- [123] J. Xiong, S. K. Kushwaha, T. Liang, J. W. Krizan, M. Hirschberger, W. Wang, R. J. Cava, and N. P. Ong, *Evidence for the chiral anomaly in the Dirac semimetal Na<sub>3</sub>Bi*, Science **350**, 413 (2015).
- [124] Xiaochun Huang, Lingxiao Zhao, Yujia Long, Peipei Wang, Dong Chen, Zhanhai Yang, Hui Liang, Mianqi Xue, Hongming Weng, Zhong Fang, Xi Dai, and Genfu Chen, *Observation of the chiral-anomaly-induced negative magnetoresistance in 3d Weyl semimetal TaAs*, Phys. Rev. X **5**, 031023 (2015).
- [125] Cai-Zhen Li, Li-Xian Wang, Haiwen Liu, Jian Wang, Zhi-Min Liao, and Da-Peng Yu, *Giant negative magnetoresistance induced by the chiral anomaly in individual Cd<sub>3</sub>As<sub>2</sub> nanowires*, Nat. Comm. **6**, 10137 (2015).
- [126] Q. Li, D. E. Kharzeev, C. Zhang, Y. Huang, I. Pletikosić, A. V. Fedorov, R. D. Zhong, J. A. Schneeloch, G. D. Gu, and T. Valla, *Chiral magnetic effect in ZrTe<sub>5</sub>*, Nat. Phys. **12**, 550 (2016).
- [127] J. Zhou, H. Jiang, Q. Niu, and J. Shi, *Topological invariants of metals and the related physical effects*, Chinese Phys. Lett. **30**, 027101 (2013).
- [128] Y. Chen, S. Wu, and A. A. Burkov, *Axion response in Weyl semimetals*, Phys. Rev. B **88**, 125105 (2013).
- [129] P. Goswami and S. Tewari, *Axionic field theory of (3+1)-dimensional Weyl semimetals*, Phys. Rev. B **88**, 245107 (2013).
- [130] G. Basar, D. E. Kharzeev, and H.-U. Yee, *Triangle anomaly in Weyl semimetals*, Phys. Rev. B **89**, 035142 (2014).

## Bibliography

- [131] Ming-Che Chang and Min-Fong Yang, *Chiral magnetic effect in a two-band lattice model of Weyl semimetal*, Phys. Rev. B **91**, 115203 (2015).
- [132] J. Ma and D. A. Pesin, *Chiral magnetic effect and natural optical activity in metals with or without Weyl points*, Phys. Rev. B **92**, 235205 (2015).
- [133] Y. Alavirad and J. D. Sau, *Role of boundary conditions, topology, and disorder in the chiral magnetic effect in Weyl semimetals*, Phys. Rev. B **94**, 115160 (2016).
- [134] S. Zhong, J. E. Moore, and I. Souza, *Gyrotropic magnetic effect and the magnetic moment on the Fermi surface*, Phys. Rev. Lett. **116**, 077201 (2016).
- [135] P. Baireuther, J. A. Hutasoit, J. Tworzydło, and C. W. J. Beenakker, *Scattering theory of the chiral magnetic effect in a Weyl semimetal: interplay of bulk Weyl cones and surface Fermi arcs*, New J. Phys. **18**, 045009 (2016).
- [136] M. A. Zubkov, *Absence of equilibrium chiral magnetic effect*, Phys. Rev. D **93**, 105036 (2016).
- [137] L. S. Levitov, Yu. V. Nazarov, and G. M. Eliashberg, *Magnetostatics of superconductors without an inversion center*, JETP Lett. **41**, 445 (1985).
- [138] Yu. V. Nazarov, *Instability due to magnetically induced currents*, Sov. Phys. JETP **64**, 193 (1986).
- [139] V. P. Mineev, *Magnetostatics and optics of noncentrosymmetric metals*, Phys. Rev. B **88**, 134514 (2013).
- [140] P. Baireuther, J. Tworzydło, M. Breitzkreiz, Ī Adagideli, and C. W. J. Beenakker, *Weyl-Majorana solenoid*, New J. Phys. **19**, 025006 (2017).
- [141] Kai-Yu Yang, Yuan-Ming Lu, and Ying Ran, *Quantum Hall effects in a Weyl semimetal: Possible application in pyrochlore iridates*, Phys. Rev. B **84**, 075129 (2011).
- [142] T. Meng and L. Balents, *Weyl superconductors*, Phys. Rev. B **86**, 054504 (2012).

- [143] G. Bednik, A. A. Zyuzin, and A. A. Burkov, *Superconductivity in Weyl metals*, Phys. Rev. B **92**, 035153 (2015).
- [144] G. Y. Cho, J. H. Bardarson, Y.-M. Lu, and J. E. Moore, *Superconductivity of doped Weyl semimetals: Finite-momentum pairing and electronic analog of the  $^3\text{He-A}$  phase*, Phys. Rev. B **86**, 214514 (2012).
- [145] V. Shivamoggi and M. J. Gilbert, *Weyl phases in point-group symmetric superconductors*, Phys. Rev. B **88**, 134504 (2013).
- [146] Huazhou Wei, Sung-Po Chao, and Vivek Aji, *Odd-parity superconductivity in Weyl semimetals*, Phys. Rev. B **89**, 014506 (2014).
- [147] P. Hosur, X. Dai, Z. Fang, and X.-L. Qi, *Time-reversal-invariant topological superconductivity in doped Weyl semimetals*, Phys. Rev. B **90**, 045130 (2014).
- [148] S. Kobayashi and M. Sato, *Topological superconductivity in Dirac semimetals*, Phys. Rev. Lett. **115**, 187001 (2015).
- [149] Tao Zhou, Yi Gao, and Z. D. Wang, *Superconductivity in doped inversion-symmetric Weyl semimetals*, Phys. Rev. B **93**, 094517 (2016).
- [150] Rui Wang, Lei Hao, Baigeng Wang, and C. S. Ting, *Quantum anomalies in superconducting Weyl metals*, Phys. Rev. B **93**, 184511 (2016).
- [151] T. Hashimoto, S. Kobayashi, Y. Tanaka, and M. Sato, *Superconductivity in doped Dirac semimetals*, Phys. Rev. B **94**, 014510 (2016).
- [152] M. Alidoust, K. Halterman, and A. A. Zyuzin, *Superconductivity in type-II Weyl semimetals*, Phys. Rev. B **95**, 155124 (2017).
- [153] Z. Faraei and S. A. Jafari, *Superconducting proximity in three dimensional Dirac materials: odd-frequency, psueudo-scalar, pseudo-vector and tensor-valued superconducting orders*, Phys. Rev. B **96**, 134516 (2017).
- [154] X. Wan, A. M. Turner, A. Vishwanath, and S. Y. Savrasov, *Topological semimetal and Fermi-arc surface states in the electronic structure of pyrochlore iridates*, Phys. Rev. B **83**, 205101 (2011).
- [155] F. D. M. Haldane, *Attachment of surface Fermi Arcs to the bulk Fermi surface: Fermi-level plumbing in topological metals*, ArXiv:1401.0529.

## Bibliography

- [156] Y. Li and F. D. M. Haldane, *Topological nodal Cooper pairing in doped Weyl metals*, Phys. Rev. Lett. **120**, 067003 (2018).
- [157] D. I. Pikulin, A. Chen, and M. Franz, *Chiral anomaly from strain-induced gauge fields in Dirac and Weyl semimetals*, Phys. Rev. X **6**, 041021 (2016).
- [158] H. Sumiyoshi and S. Fujimoto, *Torsional chiral magnetic effect in a Weyl semimetal with a topological defect*, Phys. Rev. Lett. **116**, 166601 (2016).
- [159] M. Tinkham, *Introduction to Superconductivity*, Dover Publications, (2004).
- [160] A. Boyarsky, J. Fröhlich, and O. Ruchayskiy, *Self-consistent evolution of magnetic fields and chiral asymmetry in the early universe*, Phys. Rev. Lett. **108**, 031301 (2012).
- [161] P. Goswami and B. Roy, *Effective field theory, chiral anomaly and vortex zero modes for odd parity topological superconducting state of three dimensional Dirac materials*, ArXiv:1211.4023.
- [162] V. A. Miransky and I. A. Shovkovy, *Quantum field theory in a magnetic field: From quantum chromodynamics to graphene and Dirac semimetals*, Phys. Rep. **576**, 1 (2015).
- [163] M. Stone and P. Lopes, *Effective action and electromagnetic response of topological superconductors and Majorana-mass Weyl fermions*, Phys. Rev. B **93**, 174501 (2016).
- [164] P. J. J. O'Malley, R. Babbush, I.D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Lov, H. Neven, A. Aspuru-Guzik and J. M. Martinis, *Scalable quantum simulation of molecular energies*. Phys. Rev. X **6**, 031007 (2016).
- [165] R. Barends, L. Lamata, J. Kelly, L. Garcá-Álvarez, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, I.-C. Hoi, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, J. Wenner, E. Solano and J. M. Martinis, *Digital quantum simulation of fermionic models with a superconducting circuit*. Nat. Comm. **6**, 7654 (2015).

- [166] N. K. Langford, R. Sagastizabal, M. Kounalakis, C. Dickel, A. Bruno, F. Luthi, D. J. Thoen, A. Endo and L. DiCarlo, *Experimentally simulating the dynamics of quantum light and matter at ultrastrong coupling*, Nat. Comm. **8**, 1715 (2017).
- [167] J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Yu Chen, Z. Chen, B. Chiaro, A. Dunsworth, I.-C. Hoi, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and J. M. Martinis, *State preservation by repetitive error detection in a superconducting quantum circuit*, Nature **519**, 66-69 (2015).
- [168] D. Ristè, S. Poletto, M.-Z. Huang, A. Bruno, V. Vesterinen, O.-P. Saira, and L. DiCarlo, *Detecting bit-flip errors in a logical qubit using stabilizer measurements*. Nat. Comm. **6**, 6983 (2015).
- [169] A. D. Córcoles, E. Magesan, S. J. Srinivasan, A. W. Cross, M. Steffen, J. M. Gambetta, and J. M. Chow, *Demonstration of a quantum error detection code using a square lattice of four superconducting qubits*, Nat. Comm. **6**, 6979 (2015).
- [170] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, *Extending the lifetime of a quantum bit with error correction in superconducting circuits*, Nature **536**, 441 (2016).
- [171] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, *Characterizing Quantum Supremacy in Near-Term Devices*, Nat. Phys. **14**, 595-600 (2018).
- [172] J. M. Martinis, *Qubit metrology for building a fault-tolerant quantum computer*. npj Quant. Inf. **1**, 15005 (2015).
- [173] A. G. Fowler, M. Mariantoni, J. M. Martinis and A. N. Cleland, *Surface codes: Towards practical large-scale quantum computation*, Phys. Rev. A **86**, 032324 (2012).
- [174] A. J. Landahl, J. T. Anderson, and P. R. Rice, *Fault-tolerant quantum computing with color codes*, ArXiv:1108.5738 (2011).
- [175] T. J. Yoder, and I. H. Kim, *The surface code with a twist*, Quantum **1**, 2 (2017).

- [176] Y. Tomita and K. M. Svore, *Low-distance surface codes under realistic quantum noise*, Phys. Rev. A **90**, 062320 (2014).
- [177] A. G. Fowler, A. M. Stephens and P. Groszkowski, *High-threshold universal quantum computation on the surface code*. Phys. Rev. A **80**, 052312 (2009).
- [178] B. Heim, K. M. Svore and M. B. Hastings, *Optimal circuit-level decoding for surface codes*, ArXiv:1609.06373 (2016).
- [179] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin and R. J. Schoelkopf, *Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation*. Phys. Rev. A **69**, 062320 (2004).
- [180] C. Horsman, A. G. Fowler, S. Devitt, and R. V. Meter, *Surface code quantum computing by lattice surgery*. New J. Phys. **14**, 123011 (2012).
- [181] R. Versluis, S. Poletto, N. Khammassi, B. Tarasinski, N. Haider, D. J. Michalak, A. Bruno, K. Bertels, and L. DiCarlo, *Scalable quantum circuit and control for a superconducting surface code*, Phys. Rev. App. **8**, 034021 (2017).
- [182] C. C. Bultink, M. A. Rol, T. E. O'Brien, X. Fu, B. C. S. Dikken, C. Dickel, R. F. L. Vermeulen, J. C. de Sterke, A. Bruno, R. N. Schouten, and L. DiCarlo, *Active resonator reset in the nonlinear dispersive regime of circuit QED*. Phys. Rev. Appl. **6**, 034008 (2016).
- [183] M. A. Rol, C. C. Bultink, T. E. O'Brien, S. R. de Jong, L. S. Theis, X. Fu, F. Luthi, R. F. L. Vermeulen, J. C. de Sterke, A. Bruno, D. Deurloo, R. N. Schouten, F. K. Wilhelm, and L. DiCarlo, *Restless tuneup of high-fidelity qubit gates*. Phys. Rev. Applied **7**, 041001 (2017).
- [184] S. Asaad, C. Dickel, N. K. Langford, S. Poletto, A. Bruno, M. A. Rol, D. Deurloo, and L. DiCarlo, *Independent, extensible control of same-frequency superconducting qubits by selective broadcasting*. npj Quant. Inf. **2**, 16029 (2016).
- [185] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potočník, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, and A. Wallraff, *Realizing Rapid, High-Fidelity, Single-Shot Dispersive Readout of Superconducting Qubits*. Phys. Rev. App. **7**, 054020 (2017).



- [186] A. Frisk Kockum, L. Tornberg, and G. Johansson, *Undoing measurement-induced dephasing in circuit QED*. Phys. Rev. A **85**, 052318 (2012).
- [187] H. Paik, A. Mezzacapo, M. Sandberg, D. T. McClure, B. Abdo, A. D. Córcoles, O. Dial, D. F. Bogorin, B. L. T. Plourde, M. Steffen, A. W. Cross, J. M. Gambetta, and J. M. Chow Experimental demonstration of a resonator-induced phase gate in a multiqubit circuit-QED system. Phys. Rev. Lett. **117**, 250502 (2016).
- [188] S. Gustavsson, F. Yan, G. Catelani, J. Bylander, A. Kamal, J. Birenbaum, D. Hover, D. Rosenberg, G. Samach, A. P. Sears, S. J. Weber, J. L. Yoder, J. Clarke, A. J. Kerman, F. Yoshihara, Y. Nakamura, T. P. Orlando, W. D. Oliver, *Suppressing relaxation in superconducting qubits by quasiparticle pumping*. Science **354**, 1573-1577 (2016).
- [189] A. G. Fowler, D. Sank, J. Kelly, R. Barends, and J. M. Martinis, *Scalable extraction of error models from the output of error detection circuits*, ArXiv:1405.1454 (2014).
- [190] N. Delfosse and J. P. Tillich, *A decoding algorithm for css codes using the x/z correlations*, 2014 IEEE Int. Sym. on Inf. Th., 1071–1075 (2014).
- [191] A. G. Fowler, *Optimal complexity correction of correlated errors in the surface code*. ArXiv:1310.0863 (2013).
- [192] R. Blume-Kohout, J. K. Gamble, E. Nielsen, J. Mizrahl, J. D. Sterk, P. Maunz, *Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit*, ArXiv:1310.4492 (2013).
- [193] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland and John M. Martinis *Superconducting quantum circuits at the surface code threshold for fault tolerance*. Nature **508**, 500 (2014).
- [194] Z. Chen, Y. Wang, T. Li, L. Tian, Y. Qiu, K. Inomata, F. Yoshihara, S. Han, F. Nori, J. S. Tsai and J. Q. You, *Multi-photon sideband transitions in an ultrastrongly-coupled circuit quantum electrodynamics system*, Phys. Rev. A **96**, 012325 (2017).

- [195] A. G. Fowler, *Coping with qubit leakage in topological codes*. Phys. Rev. A **88**, 042308 (2013).
- [196] A. P. Sears, A. Petrenko, G. Catelani, L. Sun, Hanhee Paik, G. Kirchmair, L. Frunzio, L. I. Glazman, S. M. Girvin, and R. J. Schoelkopf, *Photon shot noise dephasing in the strong-dispersive limit of circuit QED*, Phys. Rev. B **86**, 180504 (2012).
- [197] D. Ristè, C. C. Bultink, M. J. Tiggelman, R. N. Schouten, K. W. Lehnert, and L. DiCarlo, *Millisecond charge-parity fluctuations and induced decoherence in a superconducting transmon qubit*, Nat. Comm. **4**, 1913 (2013).
- [198] R. Blume-Kohout, J. K. Gamble, E. Nielsen, K. Rudinger, J. Mizrahi, K. Fortier, and P. Maunz *Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography*, Nat. Comm. **8**, 14485 (2017).
- [199] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf *Demonstration of two-qubit algorithms with a superconducting quantum processor*, Nature **460**, 240 (2009).
- [200] F. Yan, S. Gustavsson, A. Kamal, J. Birenbaum, A.P. Sears, D. Hover, D. Rosenberg, G. Samach, T.J. Gudmundsen, J.L. Yoder, T.P. Orlando, J. Clarke, A.J. Kerman, W.D. Oliver, *The flux qubit revisited to enhance coherence and reproducibility*, Nat. Comm. **7**, 12964 (2016).
- [201] C. M. Quintana, Y. Chen, D. Sank, A. G. Petukhov, T. C. White, D. Kafri, B. Chiaro, A. Megrant, R. Barends, B. Campbell, Z. Chen, A. Dunsworth, A. G. Fowler, R. Graff, E. Jeffrey, J. Kelly, E. Lucero, J. Y. Mutus, M. Neeley, C. Neill, P. J. J. O'Malley, P. Roushan, A. Shabani, V. N. Smelyanskiy, A. Vainsencher, J. Wenner, H. Neven, J. M. Martinis, *Observation of classical-quantum crossover of  $1/f$  flux noise and its paramagnetic temperature dependence*, Phys. Rev. Lett. **118**, 057702 (2017).
- [202] D. Ristè, J. G. van Leeuwen, H.-S. Ku, K. W. Lehnert, and L. DiCarlo, *Initialization by Measurement of a Superconducting Quantum Bit Circuit*, Phys. Rev. Lett. **109**, 050507 (2012).
- [203] M. A. Castellanos-Beltran, K. D. Irwin, G. C. Hilton, L. R. Vale, and K. W. Lehnert, *Amplification and squeezing of quantum noise with a tunable Josephson metamaterial*, Nat. Phys. **4**, 929-931 (2008).

- [204] A. G. Fowler, *Time-optimal quantum computation*, ArXiv:1210.4626 (2012).
- [205] D. A. Lidar and T. A. Brun, editors, *Quantum Error Correction* (Cambridge University Press, 2013).
- [206] A. Steane, *Multiple-particle interference and quantum error Correction*, Proc. Roy. Soc. Lond. A **452**, 2551 (1996).
- [207] S. B. Bravyi and A. Yu. Kitaev, *Quantum codes on a lattice with boundary*, ArXiv:quant-ph/9811052 (1998).
- [208] R. Raussendorf, J. Harrington, and K. Goyal, *Topological fault-tolerance in cluster state quantum computation*, New. J. Phys. **9**, 199 (2007).
- [209] J. Edmonds, *Paths, trees, and flowers*, Canad. J. Math. **17**, 449 (1965).
- [210] A. G. Fowler, *Minimum weight perfect matching of fault-tolerant topological quantum error correction in average  $O(1)$  parallel time*, Quantum Inf. Comp. **15**, 0145 (2015).
- [211] P. Baireuther, T. E. O'Brien, B. Tarasinski, and C. W. J. Beenakker, *Machine-learning-assisted correction of correlated qubit errors in a topological code*, Quantum **2**, 48 (2018).
- [212] T. E. O'Brien, B. Tarasinski, and L. DiCarlo, *Density-matrix simulation of small surface codes under current and projected experimental noise*, npj Quant. Inf. **3**, 39 (2017).
- [213] A. G. Fowler, A. C. Whiteside, A. L. McInnes, and A. Rabbani, *Topological code Autotune*, Phys. Rev. X **2**, 041003 (2012).
- [214] J. Combes, C. Ferrie, C. Cesare, M. Tiersch, G. J. Milburn, H. J. Briegel, and C. M. Caves, *In-situ characterization of quantum devices with error correction*, ArXiv:1405.5656 (2014).
- [215] Y. Fujiwara, *Instantaneous quantum channel estimation during quantum information processing*, ArXiv:1405.6267 (2014).
- [216] D. Orsucci, M. Tiersch, and H. J. Briegel, *Estimation of coherent error sources from stabilizer measurements*, Phys. Rev. A **93**, 042303 (2016).

## Bibliography

- [217] Ming-Xia Huo and Ying Li, *Learning time-dependent noise to reduce logical errors: Real time error rate estimation in quantum error correction*, New J. Phys. **19**, 123032 (2017).
- [218] S. J. Devitt, K. Nemoto, and W. J. Munro, *Quantum error correction for beginners*, Rep. Prog. Phys. **76**, 076001 (2013).
- [219] D. Gottesman, *An introduction to quantum error correction and fault-tolerant quantum computation*, Proc. Sympos. Appl. Math. **68**, 13 (2010).
- [220] A. Yu. Kitaev, *Fault-tolerant quantum computation by anyons*, Ann. Phys. **303**, 2 (2003).
- [221] E. T. Campbell, B. M. Terhal, and C. Vuillot, *The steep road towards robust and universal quantum computation*, Nature **549**, 172-179 (2017).
- [222] H. Bombin and M. A. Martin-Delgado, *Topological computation without braiding*, Phys. Rev. Lett. **98**, 160502 (2007).
- [223] H. Bombin, G. Duclos-Cianci, and D. Poulin, *Universal topological phase of two-dimensional stabilizer codes*, New J. Phys. **14**, 073048 (2012).
- [224] A. Kubica, B. Yoshida, and F. Pastawski, *Unfolding the color code*, New J. Phys. **17**, 083026 (2015).
- [225] B. Criger and B. Terhal, *Noise Thresholds for the  $[4, 2, 2]$ -concatenated Toric Code*, Quant. Inf. Comput. **16**, 1261 (2016).
- [226] R. S. Andrist, H. G. Katzgraber, H. Bombin, and M. A. Martin-Delgado, *Tricolored lattice gauge theory with randomness: Fault tolerance in topological color codes*, New J. Phys. **13**, 083006 (2011).
- [227] H. Bombin, *Gauge color codes: optimal transversal gates and gauge fixing in topological stabilizer codes*, New J. Phys. **17**, 083002 (2015).
- [228] A. Kubica and M. E. Beverland, *Universal transversal gates with color codes: A simplified approach*, Phys. Rev. A **91**, 032330 (2015).
- [229] D. Gottesman and I. L. Chuang, *Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations*, Nature **402**, 390 (1999).

- [230] D. Litinski, M. S. Kesselring, J. Eisert, and F. von Oppen, *Combining topological hardware and topological software: color code quantum computing with topological superconductor networks*, Phys. Rev. X **7**, 031048 (2017).
- [231] D. Litinski and F. von Oppen, *Braiding by Majorana tracking and long-range CNOT gates with color codes*, Phys. Rev. B **96**, 205413 (2017).
- [232] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, *Topological quantum memory*, J. Math. Phys. **43**, 4452 (2002).
- [233] D. S. Wang, A. G. Fowler, C. D. Hill, and L. C. L. Hollenberg, *Graphical algorithms and threshold error rates for the 2d colour code*, Quantum Inf. Comput. **10**, 780 (2010).
- [234] G. Duclos-Cianci and D. Poulin, *Fast decoders for topological quantum codes*, Phys. Rev. Lett. **104**, 050504 (2010).
- [235] P. Sarvepalli and R. Raussendorf, *Efficient decoding of topological color codes*, Phys. Rev. A **85**, 022317 (2012).
- [236] N. Delfosse, *Decoding color codes by projection onto surface codes*, Phys. Rev. A **89**, 012317 (2014).
- [237] A. M. Stephens, *Efficient fault-tolerant decoding of topological color codes*, ArXiv:1402.3037 (2014).
- [238] P. W. Shor, *Fault-tolerant quantum computation*, Proceedings of 37th Conference on Foundations of Computer Science, Burlington, VT, USA, pp. 56–65 (1996).
- [239] A. M. Steane, *Active stabilization, quantum computation, and quantum state synthesis*, Phys. Rev. Lett. **78**, 2252 (1997).
- [240] E. Knill, *Scalable quantum computing in the presence of large detected-error rates*, Phys. Rev. A **71**, 042322 (2005).
- [241] R. Chao and B. W. Reichardt, *Quantum error correction with only two extra qubits*, Phys. Rev. Lett. **121**, 050502 (2018).
- [242] R. Chao and B. W. Reichardt, *Fault-tolerant quantum computation with few qubits*, npj Quantum Information **4**, 42 (2018).
- [243] C. Chamberland and M. E. Beverland, *Flag fault-tolerant error correction with arbitrary distance codes*, Quantum **2**, 53 (2018).

## Bibliography

- [244] M. Gutiérrez, M. Müller and A. Bermudez, *Transversality and lattice surgery: exploring realistic routes towards coupled logical qubits with trapped-ion quantum processors*, ArXiv:1801.07035 (2018).
- [245] T. Tansuwannont, C. Chamberland, and D. Leung, *Flag fault-tolerant error correction for cyclic CSS codes*, ArXiv:1803.09758 (2018).
- [246] G. Torlai and R. G. Melko, *Neural decoder for topological codes*, Phys. Rev. Lett. **119**, 030501 (2017).
- [247] S. Varsamopoulos, B. Criger, and K. Bertels, *Decoding small surface codes with feedforward neural networks*, Quantum Sci. Technol. **3**, 015004 (2018).
- [248] S. Krastanov and L. Jiang, *Deep neural network probabilistic decoder for stabilizer codes*, Sci. Rep. **7**, 11003 (2017).
- [249] N. P. Breuckmann and X. Ni, *Scalable neural network decoders for higher dimensional quantum codes*, Quantum **2**, 68 (2018).
- [250] A. Davaasuren, Y. Suzuki, K. Fujii, and M. Koashi, *General framework for constructing fast and near-optimal machine-learning-based decoder of the topological stabilizer codes*, ArXiv:1801.04377 (2018).
- [251] C. Chamberland and P. Ronagh, *Deep neural decoders for near term fault-tolerant experiments*, Quantum Sci. Technol. **3**, 044002 (2018).
- [252] D. Poulin and Y. Chung, *On the iterative decoding of sparse quantum codes*, Quantum Inf. Comput. **8**, 987 (2008).
- [253] N. Maskara, A. Kubica, and T. Jochym-O'Connor, *Advantages of versatile neural-network decoding for topological codes*, ArXiv:1802.08680 (2018).
- [254] K. M. Svore, B. M. Terhal, and D. P. DiVincenzo, *Local fault-tolerant quantum computation*, Phys. Rev. A **72**, 022317 (2005).
- [255] R. Rojas, *Neural Networks - A Systematic Introduction*, (Springer, Berlin, Heidelberg, 1996).
- [256] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, 2016).
- [257] H. T. Siegelmann and E. D. Sontag, *Turing computability with neural nets*, Appl. Math. Lett. **4**, 77 (1991).

- [258] S. Hochreiter and J. Schmidhuber, *Long short-term memory*, Neural Computation **9**, 1735 (1997).
- [259] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, *TensorFlow: Large-scale machine learning on heterogeneous distributed systems*, ArXiv:1603.04467 (2016).
- [260] F. A. Gers, J. Schmidhuber, and F. Cummins, *Learning to forget: Continual prediction with LSTM*, Neural Computation **12**, 2451 (2000).
- [261] W. Zaremba, I. Sutskever, and O. Vinyals, *Recurrent neural network regularization*, ArXiv:1409.2329 (2014).
- [262] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, 3rd Int. Conf. for Learn. Rep. (2018).
- [263] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, *Elucidating reaction mechanisms on quantum computers*, PNAS, **114** (29), 7555-7560 (2017).
- [264] M. B. Hastings, D. Wecker, B. Bauer, M. Troyer, *Improving Quantum Algorithms for Quantum Chemistry*, Quant. Inf. Comp. **15**, 1 (2015).
- [265] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K-L. Chan, and R. Babbush, *Physical Review Letters Quantum Simulation of Electronic Structure with Linear Depth and Connectivity* Phys. Rev. Lett. **120**, 110501 (2018).
- [266] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven and G. K-L. Chan, *Low-Depth Quantum Simulation of Materials*, Phys. Rev. X **8**, 011044 (2018).
- [267] J. Alicea, *New directions in the pursuit of Majorana fermions in solid state systems*, New J. Phys. **75** (7), 076501 (2012).
- [268] C. W. J. Beenakker, *Search for Majorana Fermions in Superconductors*, Ann. Rev. Cond. Mat. Phys. **4**, 113-136 (2013).

## Bibliography

- [269] M. Leijnse and K. Flensberg, *Introduction to topological superconductivity and Majorana fermions*, Semicond. Sci. Tech. **27** (12), 124003 (2012).
- [270] S. Das Sarma, M. Freedman, and C. Nayak, *Majorana zero modes and topological quantum computation*, npj Quant. Inf. **1**, 15001 (2015).
- [271] P-L. Dallaire-Demers, and F. K. Wilhelm, *Method to efficiently simulate the thermodynamic properties of the Fermi-Hubbard model on a quantum computer*, Phys. Rev. A **93**, 032303 (2016).
- [272] Y. Li, *Fault-tolerant fermionic quantum computation based on color code*, Phys. Rev. A **98**, 012336 (2018).
- [273] T. Hyart, B. van Heck, I. C. Fulga, M. Burrello, A. R. Akhmerov, and C. W. J. Beenakker, *Flux-controlled quantum computation with Majorana fermions*, Phys. Rev. B **88**, 035121 (2013).
- [274] T. W. Larsen, K. D. Petersson, F. Kuemmeth, T. S. Jespersen, P. Krogstrup, J. Nyg ard, and C. M. Marcus, *Semiconductor-Nanowire-Based Superconducting Qubit*, Phys. Rev. Lett. **115**, 127001 (2015).
- [275] G. de Lange, B. van Heck, A. Bruno, D. J. van Woerkom, A. Geresdi, S. R. Plissard, E. P. A. M. Bakkers, A. R. Akhmerov, and L. Di-Carlo, *Realization of Microwave Quantum Circuits Using Hybrid Superconducting-Semiconducting Nanowire Josephson Elements*, Phys. Rev. Lett. **115**, 127002 (2015).
- [276] L. Fu, *Electron Teleportation via Majorana Bound States in a Mesoscopic Superconductor*, Phys. Rev. Lett. **104**, 056402 (2010).
- [277] B. van Heck, F. Hassler, A. R. Akhmerov, and C. W. J. Beenakker, *Coulomb stability of the  $4\pi$ -periodic Josephson effect of Majorana fermions*, Phys. Rev. B **84**, 180502(R) (2011).
- [278] J. D. Sau, D. J. Clarke, and S. Tewari, *Controlling non-Abelian statistics of Majorana fermions in semiconductor nanowires*, Phys. Rev. B **84**, 094505 (2011).
- [279] 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).



- [280] C. Knapp, M. Zaletel, D. E. Liu, M. Cheng, P. Bonderson, and C. Nayak, *The Nature and Correction of Diabatic Errors in Anyon Braiding*, Phys. Rev. X **6**, 041003 (2016).
- [281] A. Peruzzo, J. McClean, P. Shadbolt, M-H. Yung, X-Q. Zhou, P. J. Love, A. Aspuru-Guzik, J. L. O’Brien, *A variational eigenvalue solver on a photonic quantum processor*, Nat. Comm. **5**, 4213 (2014).
- [282] H. Tasaki, *The Hubbard model - an introduction and selected rigorous results*, J. Phys. Cond. Mat. **10**, 4353 (1998).
- [283] Z. Jiang, K. J. Sung, K. Kechedzhi, V. N. Smelyanskiy, and S. Boixo, *Quantum Algorithms to Simulate Many-Body Physics of Correlated Fermions*, Phys. Rev. App. **9**, 044036 (2018).
- [284] V. Havlíček, M. Troyer, and J. D. Whitfield, *Operator locality in the quantum simulation of fermionic models*, Phys. Rev. A **95**, 032332 (2017).
- [285] J. D. Whitfield, V. Havlíček, and M. Troyer, *Local spin operators for fermion simulations*, Phys. Rev. A **94**, 030301(R) (2016).
- [286] G. Zhu, Y. Subaşı, J. D. Whitfield, and M. Hafezi, *Hardware-efficient fermionic simulation with a cavity-QED system*, npj Quant. Inf. **4**, 16 (2018).
- [287] S. Vijay, and L. Fu, *Quantum Error Correction for Complex and Majorana Fermion Qubits*, ArXiv:1703.00459 (2017).
- [288] 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).
- [289] D. W. Berry, and A. M. Childs, *Black-box hamiltonian simulation and unitary implementation*, Quant. Inf. Comp. **12**, 29-62 (2012).
- [290] D. W. Berry, M. Kieferová, A. Scherer, Y. R. Sanders, G. H. Low, N. Wiebe, C. Gidney, and R. Babbush, *Improved techniques for preparing eigenstates of fermionic Hamiltonians*, npj Quant. Inf. **4**, 22 (2018).
- [291] D. Poulin, A. Kitaev, D. S. Steiger, M. B. Hastings, and M. Troyer, *Quantum Algorithm for Spectral Measurement with a Lower Gate Count*, Phys. Rev. Lett. **121**, 010501 (2018).

- [292] P. Wocjan and S. Zhang. *Several natural BQP-Complete problems*, ArXiv:quant-ph/0606179 (2006).
- [293] J. Preskill. Quantum Computing in the NISQ era and beyond, *Quantum* **2**, 79, (2018).
- [294] R. B. Griffiths and C.-S. Niu. Semiclassical Fourier transform for quantum computation, *Phys. Rev. Lett.* **76**, 3228–3231 (1996).
- [295] S. Kimmel, G. H. Low, and T. J. Yoder. Robust calibration of a universal single-qubit gate set via robust phase estimation, *Phys. Rev. A* **92** (6), 062315 (2015).
- [296] E. Knill, G. Ortiz, and R. D. Somma. Optimal quantum measurements of expectation values of observables, *Phys. Rev. A* **75**, 012328, (2007).
- [297] K. M. Svore, M. B. Hastings, and M. Freedman. Faster phase estimation. *Quant. Inf. Comp.* **14** (3-4), 306–328 (2013).
- [298] J. R. McClean, I. D. Kivlichan, K. J. Sung, D. S. Steiger, Y. Cao, C. Dai, E. Schuyler Fried, C. Gidney, B. Gimby, P. Gokhale, T. Häner, T. Hardikar, V. Havlíček, C. Huang, J. Izaac, Z. Jiang, X. Liu, M. Neeley, T. E. O’Brien, I. Ozfidan, M. D. Radin, J. Romero, N. Rubin, N. P. D. Sawaya, K. Setia, S. Sim, M. Steudtner, Q. Sun, W. Sun, F. Zhang, and R. Babbush, OpenFermion: The Electronic Structure Package for Quantum Computers. ArXiv:1710.07629 (2017).
- [299] N. Wiebe and C. Granade. Efficient Bayesian phase estimation. *Phys. Rev. Lett.* **117**, 010503, (2016).
- [300] D. Potts and M. Tasche. Parameter estimation for nonincreasing exponential sums by Prony-like methods. *Linear Algebra and its Applications*, **439** (4), 1024–1039 (2013).
- [301] E. Onorati, A. H. Werner, and J. Eisert. Randomized benchmarking for individual quantum gates. ArXiv:1811.11775 (2018).
- [302] G. Y. Xiang, B. L. Higgins, D. W. Berry, H. M. Wiseman, and G. J. Pryde, Entanglement-enhanced measurement of a completely unknown optical phase. *Nat. Photon.* **5** (1), 43–47, (2011).
- [303] S. Daryanoosh, S. Slussarenko, D. W. Berry, H. M. Wiseman, and G. J. Pryde, *Experimental optical phase measurement approaching the exact Heisenberg limit*. *Nat. Comm.* **9**(1), 4606, (2018).

- [304] A. A. Berni, T. Gehring, B. M. Nielsen, V. Handchen, M. G. A. Paris, and U. L. Andersen, *Ab-initio quantum-enhanced optical phase estimation using real-time feedback control*, Nat. Phot., **9**, 577-581, (2015).
- [305] X.-Q. Zhou, P. Kalasuwan, T. C. Ralph, and J. L. O'Brien. *Calculating unknown eigenvalues with a quantum algorithm*. Nat. Phot. **7**, 223-228 (2013).
- [306] S. Paesani, A. A. Gentile, R. Santagati, J. Wang, N. Wiebe, D. P. Tew, J. L. O'Brien, and M. G. Thompson, *Experimental Bayesian quantum phase estimation on a silicon photonic chip*. Phys. Rev. Lett. **118**, 100503, (2017).
- [307] R. Santagati, J. Wang, A. A. Gentile, S. Paesani, N. Wiebe, J. R. McClean, S. Morley-Short, P. J. Shadbolt, D. Bonneau, J. W. Silverstone, D. P. Tew, X. Zhou, J. L. O'Brien, and M. G. Thompson, *Witnessing eigenstates for quantum simulation of Hamiltonian spectra*. Sci. Adv. **4**, 1, 2018.
- [308] D. S. Abrams and S. Lloyd. *Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors*. Phys. Rev. Lett. **83**, 5162-5165, (1999).
- [309] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, *Simulation of electronic structure Hamiltonians using quantum computers*. Mol. Phys. **109**, 735-750 (2011).
- [310] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, *Simulating Hamiltonian dynamics with a truncated Taylor series*. Phys. Rev. Lett. **114**, 090502, (2015).
- [311] G. H. Low and I. L. Chuang. *Optimal Hamiltonian simulation by quantum signal processing*. Phys. Rev. Lett. **118**, 010501 (2017).
- [312] A. M. Childs, *On the relationship between continuous- and discrete-time quantum walk*. Comm. Math. Phys. **294**, 581-603, (2010).
- [313] A. Y. Kitaev, A. H. Shen, and M. N. Vyal'yi. *Classical and quantum computation*. Graduate studies in mathematics. American Mathematical Society, Providence, Rhode Island, (2002).
- [314] S. Bravyi, D. Gosset, R. Koenig, and K. Temme, *Approximation algorithms for quantum many-body problems*, ArXiv:1808.01734 (2018).

## Bibliography

- [315] Jay Gambetta, Alexandre Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, *Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting*. Phys. Rev. A, **74**, 042318, (2006).
- [316] B. L. Higgins, D. W. Berry, S. D. Bartlett, M. W. Mitchell, H. M. Wiseman, and G. J. Pryde, *Demonstrating Heisenberg-limited unambiguous phase estimation without adaptive measurements*. New J. Phys. **11** (7), 073023 (2009).
- [317] B. M. Terhal and D. Weigand, *Encoding a qubit into a cavity mode in circuit QED using phase estimation*, Phys. Rev. A. **93**, 012315, (2016).
- [318] N. Wiebe, Chr. Granade, and D. G. Cory. *Quantum bootstrapping via compressed quantum Hamiltonian learning*. New Journal of Physics, **17** (2), 022005 (2015).
- [319] R. Babbush and N. Ding. *Quantum phase estimation of multiple eigenvalues*, U.S. Patent Application WO2017116446A1 (2017).
- [320] N. Wiebe, C. Granade, C. Ferrie, and D. Cory. *Hamiltonian learning and certification using quantum resources*. Phys. Rev. Lett. **112**, 190501 (2014).
- [321] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince III, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer III, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, Psi4 1.1: An open-source electronic structure program emphasizing automation, advanced libraries, and interoperability. Journal of Chemical Theory and Computation, **13** (7), 3185-3197 (2017).
- [322] M. Suzuki. *Generalized Trotter formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems*. Comm. Math. Phys. **51**, 183-190 (1976).
- [323] D. Kraft. *Algorithm 733: TOMP-fortran modules for optimal control calculations*, ACM Trans. Math. Soft. **20** (3), 262-281 (1994).
- [324] J. M. Chow, J. M. Gambetta, A. D. Córcoles, S. T. Merkel, J. A. Smolin, C. Rigetti, S. Poletto, G. A. Keefe, M. B. Rothwell, J. R.

Rozen, M. B. Ketchen, and M. Steffen, *Universal Quantum Gate Set Approaching Fault-Tolerant Thresholds with Superconducting Qubits* Phys. Rev. Lett. **109**, 060501 (2012).



# Acknowledgments

No PhD is produced in a vacuum, and if there is anything particularly notable about the work that appears in these pages, it would be the quality and quantity of collaborators, colleagues, and mentors that I have had the privilege to interact with over the last few years. It gives me immense pleasure to give some (albeit small) acknowledgment to the academic contributions these people have made to this work.

Firstly and foremostly, I would like to thank Carlo for your guidance and inspiration over these past few years. I have never met a person with more insight into physics and mathematics, nor with more dedication to their work and to their students.

I would like to thank Leo for showing me physics in the real world, and for teaching me how to write in my own native language. You taught me the care, rigor, and methodicity that good science demands.

I would like to thank Inanc for sharing your insight into physics and mathematics. You have an uncanny ability to see solutions from far away, even when they require novel mathematical methods and heavy machinery to get there.

I would like to thank Barbara for taking me under your wing for nearly a whole year, and leading me through the wilderness of quantum phase estimation. You are a brilliant physicist and fantastic mentor, and I look forward to working with you more in the future.

I would like to thank Anton for making my PhD experience more interesting, and for teaching me about models for Majorana quantum computing. I hope that our next few publications together are as easy and quick to write as the last.

Brian, I first met you as a PhD student showing me how to do calculations of Landau levels in Weyl semimetals, then as a postdoc teaching me about decoders and simulation in quantum error correction, and then re-inventing Prony's method in front of my eyes for the purposes of phase estimation, and pushing me towards hidden Markov models for leakage detection. You always seem to know the answers to the problems that I can't solve, and you are probably the person that taught me the most during my PhD.

Paul (et al), thank you for teaching me about Weyl semimetals and neural networks, and for your advice on programming, academic life, and

## *Acknowledgments*

legal matters.

Slava, thank you so much for all of the coding and cluster assistance throughout my time in Leiden, even when you moved on to work in Delft.

To my other collaborators on various Weyl projects - Mathias, Nicandro, Maxim, and Jakub, thank you for correcting my mistakes and filling in the details that I missed (including the notorious cancellation of Maslov index terms).

To my, dare I say it, group - Xavi, Bruno, Stefano, and Alice. Thank you for your assistance with programming and mathematics, and your insight into physics and chemistry. I have no idea why such talented people would take me on as their advisor, but at least now I can save you some embarrassment by getting a PhD.

To my other collaborators in Leiden on various projects; Cameron, Marcello, Yaroslav, Stephen, and Stéphanie. It has been an immense pleasure working with you, thank you for putting up with me, and for very insightful and intriguing discussions.

To Adriaan, Niels, and Ramiro, who I have had the pleasure of working incredibly closely with on actual experiments (but not too close, for the sake of the cables). Thank you for your patience with this poor theorist, and your willingness to share interesting problems.

To Luuk, Franco and Detlef, for your collaboration and mentorship as we begin this exciting new project on quantum chemistry. Thank you for taking the time to explain relatively simple ideas to, and to find connections with, a person from outside of your field.

I would also like to thank the other Beenakker group members from my time in Leiden — Bernard, Denis, Benjamin, Jimmy, Shuo, Marco, Nikolay, Mark, Yevheniia, Aurélien, Gal, and Michal — and two professors in our extended Lorentz family — Vadim and Vedran — for insight and discussion throughout my PhD (especially between the hours of 11:00 and 11:30am).

I would also like to thank additional collaborators in Delft — Xiang, Florian, Ale, Jacob, Chris Dickel, Boris, Malay, Ben, Piotr and Nandini — for experimental assistance (i.e. actually doing the experiments), theoretical support (i.e. actually doing the calculations), and fruitful discussions. In addition to this, I'd like to thank Nathan, Marios, Konstantin, Thijs, Morten, Stefano, Filip, Francesco, Michael, Lieven, Stephan, Xiaotong, Chris Chamberland, Victoria and Jonas, for insightful discussions on various projects throughout the last four years.

I would also like to give a special mention to the fantastic support staff in Leiden, and in particular to Fran, Leonardo, Marianne, and Manon, without whom there would be either no institute or no working computers



within it.

The above acknowledgments are solely academic, and do not touch on the vast amount of personal support that I've received over the last few years (and prior to my time in Leiden). However, I must not go without thanking my family; Mum, Dad and Tess, and my amazing partner Louise, for all the support that you have given me. I hope that I've made you proud.



# Samenvatting

Dit proefschrift behandelt onderwerpen uit twee onderzoeksgebieden. De eerste twee hoofdstukken hebben betrekking op het bestuderen en voorspellen van elektronische eigenschappen van een exotische klasse van nieuwe materialen — zogenaamde Weyl-semimetalen. De resterende vijf hoofdstukken bestuderen enkele van de vereiste onderdelen voor de besturing en werking van een toekomstige quantumcomputer — de hardware laag, de software laag, en de laag van foutcorrectie die daartussen nodig is. Het thema dat deze delen verbindt, is topologie; het idee dat de eigenschappen van een systeem afhangen van macroscopische in plaats van microscopische kenmerken. In quantuminformatie leidt dit tot de contra-intuïtieve voorspelling dat er macroscopische vrijheidsgraden kunnen worden ontworpen die bestand zijn tegen ruis. Dit is contra-intuïtief aangezien quantumverschijnselen in principe sneller vervallen in grotere systemen, dan in kleinere systemen. Om deze reden is de quantumwereld tot de 20e eeuw onopgemerkt gebleven.

In het eerste deel onderzoeken we topologische effecten in de vaste stof fysica. In de vaste stof zorgt topologie voor het bestaan van systemen met bijzondere deeltjes, die niet aan de gebruikelijk Schrödingervergelijking voldoen: deeltjes in een Weyl-semimetaal voldoen aan de Weyl-vergelijking, en deeltjes in supergeleidende nanodraden voldoen aan de Majorana-vergelijking. Deze vergelijkingen komen uit de hoge-energiefysica en het was onverwacht dat ze in de lage-energie fysica van de vaste stof zijn opgedoken. De topologie van de elektronische bandstructuur zorgt ervoor dat deze deeltjes stabiel zijn ook als er imperfecties in het materiaal aanwezig zijn.

Een belangrijk kenmerk dat voortkomt uit de elektronische structuur van een metaal is hoe de eigenschappen ervan variëren onder invloed van een magnetisch veld. Dit heeft toepassingen in de elektronische en computerindustrie, maar biedt ook een middel om de elektronische structuur zelf te bestuderen. Een Leidse ontdekking is de periodieke variatie van materiaaleigenschappen (specifieke warmte, elektrische weerstand) als functie van het magneetveld, effecten die naar de ontdekkers genoemd zijn (het de Haas-van Alphen effect en het Shubnikov-de Haas effect — prof. W.J. de Haas was een opvolger van Kamerlingh Onnes). In hoofdstuk 2

laten we zien hoe een speciaal soort Weyl-semimetalen (type-II genaamd) door deze metingen bestudeerd kan worden. Het effect dat optreedt is dat elektronen kunnen tunnelen tussen verschillende toegestane orbitalen op een manier die analoog is aan het Klein-tunnelingeffect uit de hoge-energiefysica. De topologie van dit effect vindt haar oorsprong in het feit dat twee lijnen die rond een Möbius-band lopen elkaar op minstens één punt moeten kruisen.

Een van de meest verrassende voorspellingen in Weyl-semimetalen staat bekend onder de naam chiraal magnetisch effect, wat verwijst naar het optreden van een elektrische stroom parallel aan een aangelegd magnetisch veld. Dit is onverwacht omdat geladen deeltjes, volgens de Lorentzkracht, loodrecht op een magnetisch veld worden versneld. Men kan inderdaad bewijzen dat de ijsymmetrie van het elektromagnetisme een evenwichtsstroom evenwijdig aan een aangelegd magnetisch veld verbiedt. De ijsymmetrie is echter gebroken in een supergeleider, waardoor het chiraal magnetische effect in evenwicht tot de mogelijkheden behoort. In hoofdstuk 3 demonstreren we het optreden van een superstroom in evenwicht langs het magnetische veld, als manifestatie van het chiraal magnetische effect. De stroom van geladen deeltjes wordt gecompenseerd door een tegengestelde stroom van neutrale deeltjes, zodat er netto geen warmtestroom optreedt in evenwicht — hetgeen door thermodynamische principes verboden wordt. De richting van de elektrische stroom is afhankelijk van de vraag of de Weyl-deeltjes een linkshandige of rechtshandige chiraliteit hebben.

In deel twee van dit proefschrift verleggen we onze focus naar topologische effecten in quantumcomputers. Allereerst onderzoeken we quantumfoutcorrectie, waarbij topologie wordt gebruikt om de informatie op een meer robuuste wijze in een quantumcomputer op te slaan. Hoewel algemeen wordt aangenomen dat fouttolerante quantumberekeningen uiteindelijk mogelijk zullen zijn, ontbraken nauwkeurige numerieke voorspellingen van de mate waarin fouten met de huidige hardware gecorrigeerd kunnen worden. In hoofdstuk 4 verhelpen we dit door een simulatie uit te voeren van een prototype van een code voor quantumfoutcorrectie, bekend als de “surface code”, op supergeleidende quantum hardware met 17 qubits (“Surface-17”). Als onderdeel hiervan beschrijven en ontwikkelen we een expliciet schema voor het meten van de prestaties van de Surface-17-chip. Met de simulatie die in dit werk is ontwikkeld, kunnen we de experimentele prestaties nauwkeurig modelleren als functie van het soort verstoringen wat op kan treden. Deze mogelijkheid is sindsdien essentieel gebleken voor zowel theoretische als experimentele studies van de quantum hardware.

Zoals vermeld in het vorige hoofdstuk, is quantumfoutcorrectie gebaseerd

op een methode om lokale fouten te diagnosticeren en te corrigeren (“decoderen”), voordat ze zich ophopen en de quantuminformatie onherstelbaar beschadigen. Het optimaliseren van deze decoding vereist gedetailleerde kennis van de onderliggende foutpercentages. In hoofdstuk 5 onderzoeken we hoe iemand deze kennis precies kan bepalen binnen het kader van één van de meest populaire decodeeralgoritmen, gebaseerd op een probleem uit de grafentheorie. We testen en demonstreren deze decoder zowel voor het schatten van onderliggende foutpercentages in een statisch systeem, als voor het bijhouden van veranderingen in de tijd in een dynamisch systeem. We vinden dat de decoder de tijdsafhankelijke variaties van de foutpercentages in een supergeleidende quantumchip tot op een tijdschaal onder de milliseconde kan bijhouden.

De voorgaande twee hoofdstukken hadden betrekking op foutcorrectie voor de surface code, wat een populaire topologische code is voor quantumfoutcorrectie. Een andere topologische code, de “color code”, heeft voordelen boven de surface code, maar onderzoek aan deze alternatieve code wordt gehinderd door het gebrek aan goede decoders. In hoofdstuk 6 gebruiken we de techniek van machinaal leren door middel van een neurale netwerk om een decoder voor de color code te ontwikkelen. We demonstreren in een simulatie dat ons effectief bescherming biedt tegen veelvuldig voorkomende fouten in de quantumberekening. Voor zover aan ons bekend is dit de best presterende decoder voor de color code tot nu toe. Een bijkomend voordeel van het neurale netwerk, is dat het bruikbaar is voor andere topologische codes, voor de methode van machinaal leren maakt het immers niet uit welke de code is. Dit biedt interessante mogelijkheden voor vervolgonderzoek.

In de laatste twee hoofdstukken onderzoeken we toekomstige implementaties van quantumcomputers en de algoritmen die daarop kunnen worden uitgevoerd. Majoranadeeltjes zijn veelbelovend voor quantumberekeningen en worden om die reden intensief onderzocht als mogelijke qubits. Vanwege hun topologische bescherming verwacht men lagere foutpercentages dan met de gangbare qubits (die ook met foutencorrectie nog heel beperkte mogelijkheden hebben). Wij hebben onderzocht of Majoranadeeltjes nuttig ingezet kunnen worden voor berekeningen uit de quantumchemie. Dat is niet vanzelfsprekend, omdat de deeltjes in een molecuul elektronen (of meer in het algemeen fermionen) zijn en geen Majoranadeeltjes. In hoofdstuk 7 stellen we een schema voor waarin men paren van Majoranadeeltjes kan combineren om er fermionen van te maken. We laten zien hoe bekende quantumoperaties voor qubits kunnen worden herschreven in de taal van fermionen, zonder eerst de gebruikelijke omweg te maken via spinvariabelen.

Om quantumchemieberekeningen op een quantumcomputer uit te voeren met een versnelling ten opzichte van klassieke computers, moet men bestaande algoritmen aanpassen aan de mogelijkheden en onmogelijkheden van quantumcomputers. Het belangrijkste algoritme heet quantumfaseschatting. Zoals de naam aangeeft, biedt dit algoritme een manier aan om de fase te bepalen die wordt opgebouwd in de tijd wanneer een quantumtoestand via de Schrödingervergelijking evolueert. Bekende algoritmes voor quantumfaseschatting vereisen een nauwkeurige bereiding van de golffunctie, wat veel extra qubits kost. In het laatste hoofdstuk van dit proefschrift laten we zien hoe men de efficiëntie van het algoritme aanzienlijk kan verbeteren door gebruik te maken van de methode van Prony om fase-informatie uit een quantumcomputer te extraheren. Deze methode heeft voordelen boven de gangbare Fourieranalyse, omdat het minder qubits nodig heeft om toch een fijne resolutie te behalen. We bestuderen de prestatie van deze methode met behulp van de simulatietechniek uit hoofdstuk 4.

# Summary

This thesis is split between different topics over two fields. The first two chapters are concerned with studying and predicting electronic properties of an exotic class of new materials — Weyl semimetals. The remaining five chapters study some of the required parts for the control and operation of a future quantum computer — the architecture and software, and the layer of error correction needed in between. A theme that connects these parts is topology; the idea that one may find behaviour in a system that depends on macroscopic rather than microscopic features. In quantum information this leads to the rather counterintuitive prediction that these macroscopic degrees of freedom may be designed to be resilient against noise. This is counterintuitive as quantum phenomena traditionally decohere faster in larger systems, not smaller, allowing them to remain mostly unnoticed until the 20th century. In condensed matter, topology allows for the existence of systems with exotic effective particles: electronic excitations in a Weyl semimetal with momentum near specific Weyl points obey the Weyl equation, and low-energy electronic excitations in superconducting nanowires obey the Majorana equation. Engineering such systems without defect or disorder is a very difficult task, so we cannot expect to realise such quasiparticles via fine-tuning alone. Luckily, topological protection allows us to circumvent these concerns, and even construct systems (such as effective single Weyl cones) that are fundamentally prohibited in nature.

In part one of this thesis, we investigate novel features of Weyl semimetals. A key feature arising from the electronic structure of a metal (or any conductor) is how its properties vary under the application of a magnetic field. This has applications in the electronics and computing industry, but also gives a means to study the electronic structure itself. The celebrated de Haas-van Alphen and Shubnikov-de Haas effects (discovered in Leiden), and their related counterparts, have been used with great success to probe metals for the last fifty years. These effects appear as fluctuations of various material properties (e.g. magnetic susceptibility, specific heat, and electrical resistivity), periodic in one over the strength of an applied magnetic field. In chapter 2, we show how a signature of a broad class of Weyl semimetals (type-II Weyl semimetals) appears in measurements of this type. This signature manifests by electrons tunneling between different

## Summary

allowed orbitals in a manner forbidden by a semi-classical description, in a momentum-space analogue to the celebrated Klein tunneling effect. Interestingly, we find that this effect has a topological protection separate to that of a Weyl semimetal, that emerges from the inability to prevent two lines running around a Mobius band from crossing at at least one point.

One of the most stunning predictions in Weyl semimetals is the chiral magnetic effect, whereby electrons are accelerated parallel to an applied magnetic field. This is in contradiction to the standard behaviour of charged particles which, following the Lorentz rule, are accelerated perpendicular to an applied magnetic field. Indeed, one may prove that the gauge symmetry of the electron forbids an equilibrium current parallel to an applied magnetic field. Thus, the chiral magnetic effect is not observable in a regular metal in equilibrium. However, the gauge symmetry of the electron is broken in a superconductor, re-opening the possibility of such an observation, which we demonstrate and study in chapter 3. This manifests as a charge current balanced by a flow of zero-charge particles in the opposite direction to balance the heat current (which must remain net-zero). The flow of these charged particles depends on the chirality of an effective single Weyl cone in the bulk of the system, which may be switched by an applied supercurrent, thus providing a ‘chirality switch’ for this effect.

In part two of this thesis we switch our focus to quantum error correction, where topology is used to enhance control over quantum degrees of freedom, rather than ensuring their existence. Although it is widely-accepted that fault tolerant quantum computing is eventually possible, accurate numerical predictions of the performance of current efforts to demonstrate scalable full quantum error correction have been lacking, in part due to the complexity of the system. In chapter 4 we rectify this, by performing a full-density matrix simulation of a small quantum error correcting code prototype, known as Surface-17, on superconducting quantum hardware. As part of this, we explicitly detail and develop a scheme for measuring the experimental performance of the Surface-17 chip, and provide and benchmark a decoder to analyze the information from repeated parity checks performed during this experiment. The density-matrix simulation toolbox developed during this work allows us to accurately model experimental performance as noise parameters are varied, which has proved essential in both theory and experimental studies since then.

As mentioned in the previous chapter, quantum error correction requires classical processing of repeated syndrome measurements to diagnose and correct local errors before they accumulate to kill the macroscopically-



spread logical quantum information. Optimizing this decoding requires detailed knowledge of the rates at which these local errors occur in order to determine the most-likely set of underlying events from the error syndrome. In chapter 5 we investigate how one may determine this knowledge exactly within the framework of one of the most popular decoding algorithms, based on the minimum-weight perfect matching problem in graph theory. We test and demonstrate this adaptive decoder both in estimating underlying error rates in a static system, and in keeping up with noise fluctuations on the time-scale of a single quantum error correcting experiment. We find that the decoder may keep up with what would be sub-millisecond fluctuations on a superconducting transmon chip, and demonstrate the trade-off between lagging behind said fluctuations and trying too hard to keep up (resulting in insufficient convergence).

The previous two chapters dealt with error correction for the surface code, which is a popular choice for quantum error correction due to the excellent performance of minimum-weight perfect matching decoders. Other topological quantum error correcting codes, such as the color code, have even lower qubit counts and easier access to logical operations than the surface code, however research in these has been somewhat stymied due to the lack of good decoders. In chapter 6, we extend a class of neural network quantum error correction decoders to work on the color code. Importantly, this requires both processing of traditional parity-check measurements, and additional flag measurements, which are required to diagnose a small class of particularly dangerous ‘hook’ errors that appear in the color code. We demonstrate numerical evidence that our scheme protects against these hook errors, despite using fewer flag measurements than expected in the rigorous theoretical proofs of their performance. This allows us significant savings in measurement overhead, giving us (to the best of our knowledge) the best-performing color code decoder yet. Furthermore, the presented scheme is usable for any quantum error correcting code, opening up a range of interesting possibilities for code design.

In the final part of this thesis, we look towards future implementations of quantum computers and the algorithms to run thereon. Majorana modes, or Majorana bound states, have emerged as an exciting prospect for quantum computing. Their topological protection raises hopes of lower error rates than current state-of-the-art devices (which is potentially the largest barrier to future useful quantum computers, even with quantum error correction). However, most work studying Majorana modes as a platform for quantum computing has paid little attention to their non-abelian quasiparticle nature, save for the purposes of constructing

## Summary

a universal quantum gate set. This is of importance when looking to perform future quantum chemistry calculations on a quantum computer: these calculations typically study electronic behaviour, the simulation of which on a bosonic qubit system incurs a significant overhead. In chapter 7, we propose a scheme where one might circumvent this overhead on a Majorana architecture by directly combining pairs of Majoranas to make fermions. We demonstrate how one may decompose traditional quantum operations into the language of fermions (albeit in a non-fault tolerant setting), and show a large advantage over traditional quantum computing implementations in simulating lattice fermionic systems.

In order to perform quantum chemistry calculations, or any calculations, on a quantum computer (with a quantum speedup over classical computers), one must tailor new algorithms to quantum devices. The work-horse behind most (if not all) algorithms with a provable quantum speedup is quantum phase estimation. As suggested in the name, this algorithm provides a means to determine the relative phase accumulated by an otherwise stationary state (also known as an eigenstate) as it evolves following the Schrödinger equation (simulated on a quantum register). Traditionally, this has either required accumulating this phase onto multiple ancillary qubits, or precise preparation of the eigenstate itself, both of which come with significant overhead. In the last chapter of this thesis, we demonstrate how one may separate the phase information from a superposition or mixture of eigenstates by single rounds of the quantum phase estimation protocol with a single ancilla qubit. This work pioneers the use of Prony's method to extract phase information from a quantum computer, which is preferable to more traditional Fourier-type methods as it allows for finer resolution of a small number of phases with a low classical computation cost. We finally study the performance of this method, and that of Bayesian post-processing, in the presence of nearby eigenenergies (which make distinction of a single frequency difficult), and two prototypical noise models (including the accurate density-matrix simulations presented in Chapter 4).

# Curriculum Vitæ

I was born on the 18th of November 1990 in Otahuhu, Auckland, New Zealand. In 1996, I emigrated with my parents to Wollongong, Australia, where I remained for the rest of my childhood.

In 2009, I began attending the University of Wollongong, graduating in 2012 with a Bachelor of Science (Chemistry and Physics) and a Bachelor of Mathematics (Advanced). During this time I undertook research projects in biochemistry (under Prof. Nicholas Dixon), physical chemistry (under Prof. Adam Trevitt), algebraic topology (under Prof. David Pask), and theoretical condensed matter physics (under Dr. Anthony Wright and Prof. Chao Zhang). In 2013, I shifted to the University of Queensland, Australia, to study for an BSc(Honours)(Physics) for a year, taking on a project studying Majorana bound states (under Dr. Wright) in strongly correlated systems, and graduating first in this class. In 2014, I was accepted into the Perimeter Institute's PSI Master's program, in Waterloo, Canada. As part of this 9-month intensive program I wrote my Master's thesis on numerical investigations of many-body localization under the tutelage of Prof. Guifre Vidal, Prof. Dmitry Abanin, and Dr. Zlatko Papić.

In 2015, I was accepted for a PhD position in the group of Prof. Carlo Beenakker at Leiden University on a project with QuTech in Delft funded by a Synergy grant from the European Research Council. I began in August 2015, studying transport properties of Weyl semimetals, but towards the end of that year I took on an additional project; to provide theoretical assistance to the IARPA project of Prof. Leonardo DiCarlo (Delft University of Technology) on quantum error correction in superconducting transmon qubits. Collaboration on this quantum error correction project continues actively to this day, and both this and the Weyl project play a large role in this thesis. In 2017, with Prof. DiCarlo, Prof. Lieven Vandersypen (Delft University of Technology), and industry partners, I applied for and successfully obtained an NWA Start-Impuls grant from the Dutch National Science Foundation, to hire a postdoctoral fellow for three years to study quantum chemistry on a quantum computer. This project was the seed for a collaboration with Shell Research that will occupy me for the next several years.



# List of publications

- T. E. O'Brien, B. Senjean, R. Sagastizabal, X. Bonet-Monroig, A. Dutkiewicz, F. Buda, L. DiCarlo, and L. Visscher. *Calculating energy derivatives for quantum chemistry on a quantum computer*. ArXiv:1905.03742 (2019).
- R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. C. Bultink, X. Fu, C. H. Price, V. P. Ostroukh, N. Muthusubramanian, A. Bruno, M. Beekman, N. Haider, T. E. O'Brien, and L. DiCarlo. *Error mitigation by symmetry verification on a variational quantum eigensolver*. ArXiv:1902.11258 (2019).
- P. Baireuther, M. D. Caio, B. Criger, C. W. J. Beenakker, and T. E. O'Brien. *Neural network decoder for topological color codes with circuit level noise*. New Journal of Physics **21** (1), 013003 (2019). [Chapter 6]
- T. E. O'Brien, B. Tarasinski, and B. M. Terhal. *Quantum phase estimation of multiple eigenvalues for small-scale (noisy) experiments*. New Journal of Physics **21**, 023022 (2019). [Chapter 8]
- X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien. *Low-cost error mitigation by symmetry verification*. Physical Review A **98** (6), 062339 (2018).
- S. T. Spitz, B. Tarasinski, C. W. J. Beenakker, and T. E. O'Brien. *Adaptive weight estimator for quantum error correction in a time-dependent environment*. Advanced Quantum Technologies **1** (1), 1870015 (2018). [Chapter 5]
- T. E. O'Brien, P. Rožek, and A. R. Akhmerov. *Majorana-based fermionic quantum computation*. Physical Review Letters **120** (22), 220504 (2018). [Chapter 7]
- N. Bovenzi, M. Bretkreiz, T. E. O'Brien, J. Tworzydło, and C. W. J. Beenakker. *Twisted Fermi surface of a thin-film Weyl semimetal*. New Journal of Physics **20** (2), 023023 (2018).

- P. Baireuther, T.E. O'Brien, B. Tarasinski, and C.W.J. Beenakker. *Machine-learning-assisted correction of correlated qubit errors in a topological code*. Quantum **2**, 48 (2018).
- J.R. McClean, I.D. Kivlichan, K.J. Sung, D.S. Steiger, Y. Cao, C. Dai, E.S. Fried, C. Gidney, B. Gimby, P. Gokhale, T. Häner, T. Hardikar, V. Havlíček, C. Huang, J. Izaac, Z. Jiang, X. Liu, M. Neeley, T.E. O'Brien, I. Ozfidan, M.D. Radin, J. Romero, N. Rubin, N.P.D. Sawaya, K. Setia, S. Sim, M. Steudtner, Q. Sun, W. Sun, F. Zhang, and R. Babbush. *OpenFermion: the electronic structure package for quantum computers*. ArXiv:1710.07629 (2017).
- T.E. O'Brien, B. Tarasinski, and L. DiCarlo. *Density-matrix simulation of small surface codes under current and projected experimental noise*. npj Quantum Information **3**, 27 (2017). [Chapter 4]
- N. Bovenzi, M. Bretkreiz, P. Baireuther, T.E. O'Brien, J. Tworzydło, Ī. Adagideli, and C.W.J. Beenakker. *Chirality blockade of Andreev reflection in a magnetic Weyl semimetal*. Physical Review B **96** (3), 035437 (2017).
- T.E. O'Brien, C.W.J. Beenakker, and Ī. Adagideli. *Superconductivity provides access to the chiral magnetic effect of an unpaired Weyl cone*. Physical Review Letters **118** (20), 207701 (2017). [Chapter 3]
- M.A. Rol, C.C. Bultink, T.E. O'Brien, S.R. de Jong, L.S. Theis, X. Fu, F. Luthi, R.F.L. Vermeulen, J.C. de Sterke, A. Bruno, D. Deurloo, R.N. Schouten, F.K. Wilhelm, and L. DiCarlo. *Restless tuneup of high-fidelity qubit gates*. Physical Review Applied **7** (4), 041001 (2017).
- T.E. O'Brien, D.A. Abanin, G. Vidal, and Z. Papić. *Explicit construction of local conserved operators in disordered many-body systems*. Physical Review B **94** (14), 144208 (2016).
- C.C. Bultink, M.A. Rol, T.E. O'Brien, X. Fu, B.C.S. Dikken, C. Dickel, R.F.L. Vermeulen, J.C. de Sterke, A. Bruno, R.N. Schouten, and L. DiCarlo. *Active resonator reset in the nonlinear dispersive regime of circuit QED*. Physical Review Applied **6** (3), 034008 (2016).
- T.E. O'Brien, M. Diez, and C.W.J. Beenakker. *Magnetic breakdown and Klein tunneling in a type-II Weyl semimetal*. Physical Review Letters **116** (23), 236401 (2016). [Chapter 2].

- T.E. O'Brien, and A.R. Wright. *A many-body interpretation of Majorana bound states, and conditions for their localisation.* ArXiv:1508.06638 (2015).
- T.E. O'Brien, A.R. Wright, and M. Veldhorst. *Many-particle Majorana bound states: derivation and signatures in superconducting double quantum dots.* Physica Status Solidi (b) **252** (8), 1731-1742 (2015).
- T.E. O'Brien, C. Zhang, and A.R. Wright. *Universal geometric classification of armchair honeycomb nanoribbons by their properties in a staggered sublattice potential.* Applied Physics Letters **103** (17), 171608 (2013).
- T.E. O'Brien, and C. Zhang. *Optical conductance of a two-dimensional semiconductor in the presence of Rashba spin-orbit coupling and a periodic potential.* Modern Physics Letters B **26** (26), 1250174 (2012).
- A.R. Wright, T.E. O'Brien, D. Beaven, and C. Zhang. *Gapless insulator and a band gap scaling law in semihydrogenated graphene.* Applied Physics Letters **97** (4), 043104 (2010).

# Stellingen

behorende bij het proefschrift

*Applications of topology to Weyl semimetals and quantum computing*

1. Weyl points in a type-II Weyl semimetal that are too far from the Fermi level to exhibit Klein tunnelling will not play a role in the properties of the material (Chapter 2).
2. Superconductivity provides a loophole in the no-go theorem that relies on gauge invariance to prohibit the equilibrium chiral magnetic effect (Chapter 3).
3. Quantum error correction cannot optimally account for noise fluctuations on a scale faster than approximately 1000 stabilizer measurement cycles (Chapter 5).
4. The advantages of fermionic quantum computing schemes for digital quantum simulation will outweigh the increase in error brought on by a lack of local fermion parity conservation (Chapter 7).
5. The flag-fault tolerance scheme of Chamberland and Beverland needs only a single flag qubit per ancilla.  
C. Chamberland and M. E. Beverland, *Quantum* **2**, 53 (2018).
6. The ‘leakage paralysis’ effect of Ghosh *et al.* will not occur in a fully quantum mechanical code.  
J. Ghosh, A. G. Fowler, J. M. Martinis, and M. R. Geller, *Phys. Rev. A* **88**, 062329 (2013).
7. Error mitigation strategies such as those considered in McArdle *et al.* and Bonet-Monroig *et al.* will be used by fault-tolerant quantum computers.  
S. McArdle, X. Yuan, and S. C. Benjamin, *Phys. Rev. Lett.* **122**, 180501 (2019).  
X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O’Brien, *Phys. Rev. A* **98**, 062339 (2019).
8. Estimation of energy derivatives, along the lines of arXiv:1905.03742, will be as essential for the success of digital quantum simulation as the estimation of the eigenenergies themselves.



9. Freedom of movement is a human right.
10. Quantum computers are only at the same stage as classical computers were in the early 1930s.
11. Evil gains power by being an emergent phenomenon.
12. An undergraduate degree teaches you how to find answers. A PhD teaches you how to find questions.

Thomas Eugene O'Brien  
Leiden, 20 juni 2019