TOWARD USEFUL APPLICATIONS OF QUANTUM PROCESSORS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Eliott Rosenberg

August 2023

TOWARD USEFUL APPLICATIONS OF QUANTUM PROCESSORS

Eliott Rosenberg, Ph.D.

Cornell University 2023

Recent years have seen tremendous progress toward building programmable quantum computers, with a variety of private companies and government agencies, as well as academic institutions, investing heavily in the emerging technology. Quantum computers promise to solve certain problems that are intractable for classical computers, but the utility of existing devices is severely restricted by noise. This thesis details several attempts to overcome this noise and make use of existing devices to address outstanding questions in physics. The work presented here demonstrates that, by employing error mitigation and avoidance techniques, existing devices can be used to study both ground state physics (Chapters 2, 3) and dynamics (Chapters 5, 6) of local spin systems. I also consider a non-local system of Majorana fermions (Chapter 4) but find its simulation to be beyond the capabilities of existing quantum processors at nontrivial system sizes. Whether the successful quantum simulations detailed here are extensible beyond the reach of classical techniques remains an open question, but the rapid pace of development suggests that useful quantum advantage may occur soon. Of the quantum simulations presented here, that of Chapter 5 is the most promising candidate for useful quantum advantage because it both addresses an outstanding question in the scientific community and extends beyond our classical simulations. However, more work is needed to determine whether it is beyond the capabilities of approximate classical methods as well. I expect that, in

the coming years, quantum processors will be able to perform quantum simulations that are increasingly useful for the scientific community and increasingly difficult to simulate classically.

BIOGRAPHICAL SKETCH

Eliott Rosenberg, recognizing the dangers of identity theft, prefers not to write his date and city of birth in a document that will be made publicly accessible. Suffice to say that he was born in the western United States in the mid-1990s. At the age of two, he made the arduous journey across the country in order to return to his ancestral homeland of Rochester, NY. Subsequently, at the age of three, he decided to become a vegetarian. Skipping over several formative years, he graduated from Penfield High School in 2013 and then from Brown University in 2017, obtaining a Bachelor of Science in Physics from the latter. Almost immediately thereafter, he matriculated at Cornell University in order to pursue theoretical high energy physics. He then changed his mind and pursued quantum computing instead.

This document is dedicated to my parents, who always encouraged my curiosity and provided me with endless opportunities to learn and explore.

ACKNOWLEDGEMENTS

I am deeply grateful to everyone who has made my time as a graduate student so enjoyable, only some of whom I can thank here.

First, I want to thank my advisor, Paul Ginsparg, whose course first sparked my interest in quantum computing. His decision to take me on as a student and the guidance he provided made this work possible. Importantly, he brought me into a collaboration with Peter McMahon and Tom Hartman, whom I would also like to thank, and then subsequently gave me the freedom to pursue collaborations outside of Cornell.

Peter McMahon was a second advisor to me. Without his knowledge of the field and guidance as to what problems are relevant and interesting, my work would not have had its direction. Further, the connections and partnerships that he forged gave me access to the hardware that made this work possible and introduced me to the people who shaped the latter part of my graduate career. His advice in navigating the field has proved invaluable.

Raphael Pooser and Titus Morris, both of Oak Ridge National Laboratory, gave me invaluable time and advice as we collaborated on a project to mitigate correlated readout errors.

Pedram Roushan and the whole Google Quantum AI team, who hosted me for close to a year, provided me not only with invaluable day-to-day guidance, but also with a sense of excitement, the feeling that I am participating in something important at the forefront of science and technology. I am incredibly grateful for the opportunity that they have given me.

Additionally, I would like to thank my first advisor, Maxim Perelstein, for being an excellent mentor as I pursued high energy phenomenology and for his flexibility and support in my decision to switch fields, as well as all of the Cornell professors from whom I have learned so much about quantum mechanics and physics more generally, including but not limited to Peter Lepage, Liam McAllister, Tom Hartman, Csaba Csaki, Yuval Grossman, Debanjan Chowdhury, Erich Mueller, Tomas Arias, and Farhan Rana.

Finally, I would like to thank my friends who helped me along my grad school journey. I would like to thank members of my cohort at Cornell, including but not limited to Mijo, Ryan, Naomi, and Andres, whose presence in many classes that we took together I greatly appreciate. I would like to thank Ricky Oliver for his constant friendship. He and I have studied physics together since our freshman year at Brown, and much of what I know I attribute to our many hours of working through problem sets with him. Finally, I would like to thank my girlfriend, Nidia, for her love and support, including the hours that she spent listening to me practice my thesis defense and offering valuable feedback.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Ideal quantum computers

Ideal quantum computers differ from their classical counterparts in that, instead of performing logical operations on bits, they perform unitary operations and measurements on quantum states, which are written in terms of qubits. An ideal qubit is a two-state quantum system, whose Hilbert space $\mathcal{H}_q$ is spanned by the states $|0\rangle$ and $|1\rangle$. The Hilbert space of an n-qubit quantum computer is the tensor product $\mathcal{H}_q^{\otimes n}$.

The state of a single qubit can be readily visualized on (or in) the Bloch sphere, with pure states on the surface and mixed states in the interior. In particular, the state can be written

$$\rho = \frac{1}{2}\left(\mathbb{I} + \vec{a} \cdot \vec{\sigma}\right), \tag{1.1}$$

which has the nice property that

$$\langle \sigma_i \rangle = a_i, \tag{1.2}$$

seen easily by taking $\mathrm{Tr}[\rho\sigma_i]$. This implies, for example, that $|0\rangle$, the eigenstate of σ_z with eigenvalue $+1$, is at the north pole and $|1\rangle$, the eigenstate of σ_z with eigenvalue -1, is at the south pole. Similarly, $|+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)$, the eigenstate of σ_x with eigenstate $+1$ lies on the positive x-axis, with Cartesian coordinates $(1, 0, 0)$.

Further, using the fact that $\left(\vec{a} \cdot \vec{\sigma}\right)^2 = |a|^2\mathbb{I}$, and that $\mathrm{Tr}[\rho^2] \leq 1$, where it equals one if and only if ρ is pure, it is readily seen that $|a|\leq 1$ and is equal to one if and only if the state is pure. Thus, $\vec{a}$ describes the state of the qubit on or in a sphere.

Unitary operations on the Hilbert space act as rotations on $\vec{a}$. This stems from the fact that SU(2) is a double cover of SO(3). Explicitly, a rotation of angle θ about the axis $\hat{n}$, acting in the 3-dimensional real space in which $\vec{a}$ lives, can be represented as a unitary matrix acting on the 2-dimensional complex Hilbert space:

$$U(\theta, \hat{n}) = e^{-i(\hat{n}\cdot\vec{\sigma})\theta/2} = \cos\left(\frac{\theta}{2}\right)\mathbb{I} - i\sin\left(\frac{\theta}{2}\right)(\hat{n}\cdot\vec{\sigma})\,. \tag{1.3}$$

This allows us to easily visualize any single-qubit unitary as a rotation. There are a number of single qubit gates that are standard in quantum computing. These include the three Pauli matrices, which are often written as $X = \sigma_x$, $Y = \sigma_y$, and $Z = \sigma_z$. From Eq. 1.3, it is readily seen that, up to an unphysical global phase, these implement 180° rotations about their respective axes. For example the fact that $X|0\rangle = |1\rangle$ can be visualized as the north pole of the Bloch sphere, rotated 180° about the x-axis, equaling the south pole. Another important single-qubit gate is the Hadamard gate, $H = \frac{1}{\sqrt{2}}(X+Z)$, which implements a 180° rotation about the axis $\hat{n} = \frac{1}{\sqrt{2}}(\hat{x}+\hat{z})$.

Single-qubit gates alone are not sufficient for universal quantum computation because they cannot generate entanglement between qubits. This requires two-qubit gates. An example of a two-qubit gate is the controlled-not gate, or the CNOT gate, which implements an X on the target qubit if the control qubit is in the $|1\rangle$ state. If we number the qubits as $|q_1q_0\rangle$ and let qubit 0 be the control and qubit 1 be the target, then this has the matrix

$$\begin{aligned}
C_X &= \mathbb{I} \otimes |0\rangle\langle 0| + X \otimes |1\rangle\langle 1| \\
&= |00\rangle\langle 00| + |11\rangle\langle 01| + |10\rangle\langle 10| + |01\rangle\langle 11| \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}
\end{aligned} \tag{1.4}$$

where the states are ordered as $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$, like binary numbers.

Another important two-qubit gate is the controlled-Z gate, or the CZ gate. This is defined as

$$\begin{aligned}
C_Z &= \mathbb{I} \otimes |0\rangle\langle 0| + Z \otimes |1\rangle\langle 1| \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
&= |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes Z.
\end{aligned} \tag{1.5}$$

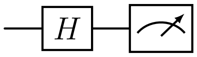
as is clear from the matrix representation, we can think of either qubit as the control qubit in this case. Note also that CX and CZ are related to each other by Hadamards on the target qubit since $HZH = X$ and $HXH = Z$. It turns out that single-qubit gates, plus either CX or CZ, are sufficient for universal quantum computation. Any n-qubit unitary matrix can be decomposed, to a specified precision, in terms of single qubit and CZ gates.

Another important class of two qubit gates are fSim gates, which are number-conserving gates described by the matrix

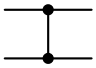
$$U_{\text{fSim}}(\theta, \phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & i\sin\theta & 0 \\ 0 & i\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & e^{i\phi} \end{pmatrix}, \tag{1.6}$$

which parameterizes all number-conserving two qubit gates up to single-qubit Z phases. For example, a Z gate is an fSim gate with $(\theta, \phi) = (0, \pi)$. The iSWAP gate is an fSim with $(\theta, \phi) = (\pi/2, 0)$. The SWAP gate, which swaps two qubits, is related to fSim$(\pi/2, \pi)$ by single-qubit phases.

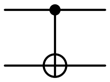
Sequences of single- and two-qubit gates are represented visually as quantum circuits, which are read from left to right. For example,



represents a single qubit to which a Hadamard is applied and which is then measured. Unless otherwise specified, qubits are assumed to begin in the $|0\rangle$ state, so the measurement outcomes of this circuit will be 50% “0” and 50% “1”. The symbol for a CZ gate is



and the symbol for a CX gate, with the top qubit as control and the bottom qubit as target, is



Ideal universal quantum computers would have a revolutionary impact on society. They would enable an exponential speedup in the prime factorization of large numbers [153], the difficulty of which is the basis of widely-used forms of encryption, such as RSA, a quadratic speedup in searching through a database [66], and the ability to simulate quantum dynamics, e.g. by implementing Trotterized time evolution of a quantum Hamiltonian or quantum phase estimation to measure energy eigenvalues and generate eigenstates [92].

Ideal quantum computers can be simulated by classical computers, but at an exponential cost in the number of qubits. This stems from the fact that the statevector describing n qubits consists of 2^n complex numbers. For example, a 36-qubit statevector, stored with single precision, takes 550 gigabytes, whereas a 37-qubit statevector takes 1100 gigabytes. The system requirements even to store the state quickly surpass the RAM available even in the most powerful supercomputers.

1.2 Real quantum computers

Various hardware implementations of qubits exist or are being pursued. Each has its own technical challenges. This section focuses on transmons, the hardware implementation chosen by both Google and IBM.

1.2.1 The spectrum of a transmon

A transmon consists of a Josephson junction and a capacitor, as illustrated in Figure 1.1.

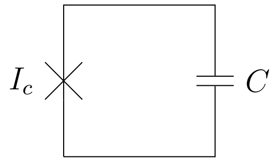


Figure 1.1: A transmon.

The Josephson junction is a superconducting circuit element that behaves like a nonlinear inductor. The current through it and voltage across it satisfy the Josephson equations:

$$\begin{aligned} I &= I_c \sin\varphi \\ V &= \bar{\Phi}_0 \dot{\varphi}, \end{aligned} \tag{1.7}$$

where φ is the Josephson phase (similar to the magnetic flux through an inductor), and $\bar{\Phi}_0 = \hbar/(2e)$. I_c is the critical current above which the junction behaves like a regular resistor. (For small currents, these are the same equations as for an inductor). (From now on, I will set $\hbar = 1$.) The energy stored in the Josephson junction is therefore

$$U = \int IV dt = -\bar{\Phi}_0 I_c \cos\varphi, \tag{1.8}$$

while the energy stored in the capacitor is

$$K = \frac{1}{2}CV^2 = \frac{1}{2}C(\Phi_0\dot{\varphi})^2 \tag{1.9}$$

Treating the Josephson phase as the canonical coordinate, we therefore have the Lagrangian

$$\mathcal{L} = K - U = \frac{1}{2}C(\bar{\Phi}_0\dot{\varphi})^2 + \bar{\Phi}_0 I_c \cos\varphi. \tag{1.10}$$

The canonical momentum is therefore

$$n = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = C\bar{\Phi}_0^2\dot{\varphi}, \tag{1.11}$$

which happens to be equal to $\bar{\Phi}_0 CV = Q/(2e)$, i.e. the charge on the capacitor in units of the Cooper pair charge, $2e$. It is called n because it counts the number of Cooper pairs. Now, we impose the canonical commutation relation:

$$[\varphi, n] = i \tag{1.12}$$

and write the Hamiltonian:

$$\begin{aligned} H &= \frac{n^2}{2C\bar{\Phi}_0^2} - \bar{\Phi}_0 I_c \cos\varphi \\ &= 4E_C n^2 - E_J \cos\varphi, \end{aligned} \tag{1.13}$$

where $E_C = e^2/(2C)$ is the charging energy and $E_J = \bar{\Phi}_0 I_c$. This is the Hamiltonian for a nonlinear oscillator with a cosine potential. The fact that the number of Cooper pairs on the capacitor is integer implies that φ is periodic. The energy eigenfunctions are exactly the Mathieu functions (see, e.g. Appendix B of [96]).

Although the exact eigenfunctions and eigenenergies are known, intuition can be gained by Taylor expanding the cosine. The lowest-order approximation is

$$H_0 = 4E_C n^2 + \frac{1}{2}E_J\varphi^2, \tag{1.14}$$

which is a harmonic oscillator Hamiltonian with frequency $\omega = \sqrt{8E_JE_C}$. Next, we consider the quartic correction:

$$H_1 = -\frac{1}{4!}E_J\varphi^4. \tag{1.15}$$

The first-order correction to the energy in the mth eigenstate is simply $\langle m|H_1|m\rangle$, which is easily computed in the oscillator's number basis using

$$\begin{aligned} \varphi &= \left(\frac{2E_C}{E_J}\right)^{1/4}\left(a + a^\dagger\right) \\ n &= -\frac{i}{2}\left(\frac{E_J}{2E_C}\right)^{1/4}\left(a - a^\dagger\right). \end{aligned} \tag{1.16}$$

The result is that

$$\begin{aligned} \omega_{10} &= \sqrt{8E_JE_C} - E_C \\ \omega_{21} &= \sqrt{8E_JE_C} - 2E_C \\ &\cdots \\ \omega_{m,m-1} &= \sqrt{8E_JE_C} - mE_C, \end{aligned} \tag{1.17}$$

that is, the spacing between consecutive energy levels decreases by the charging energy E_C. The quantity $\eta \equiv \omega_{21} - \omega_{10}$ is called the anharmonicity, which is $-E_C$ to leading order. The nonzero anharmonicity allows one to individually address the $0 \rightarrow 1$ transition and therefore isolate the lowest two energy levels as the qubit subspace. Care must be taken to avoid populating the higher energy levels, e.g. by minimizing the ω_{21} fourier component of any pulses used to control the qubits. Population of the higher states is known as leakage.

Note that there is a physical energy difference between the $|0\rangle$ and $|1\rangle$ states. Therefore, $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$ is not stationary; it evolves to $|\psi(t)\rangle = c_0|0\rangle +$

$c_1 e^{-i\omega_{10}t}|1\rangle$, i.e. it rotates about the z axis of the Bloch sphere at an angular frequency ω_{10}. This is accounted for by working in the interaction picture, i.e. in a rotating frame that also rotates about the z axis at frequency ω_{10}.

Depending on the type of two-qubit gate used, Z gates and Z-rotations more generally may be able to be implemented virtually. This means that, instead of applying any physical operations to the qubit, a Z gate simply updates the frame, which affects, for example, the lab-frame axes about which X and Y gates act. This strategy can be used, for example, if the native two-qubit gate is CZ, since single qubit Z phases commute with CZ, so the change of frame has no effect on the two-qubit gate. For more general fSim gates, it is not so trivial to apply the frame-change to the gate, so physical Z gates are used instead.

1.2.2 Frequency-tunable transmons

One feature of some transmon designs (such as Google's) but not others (such as IBM's) is the ability to tune the frequency of the qubit (i.e. ω_{10}) during the execution of a quantum circuit. This is useful for applying physical Z gates. For example, to apply a Z rotation by an angle θ, simply detune the qubit's frequency by an amount $\Delta\omega$ such that $\tau\Delta\omega = \theta$, where τ is the duration of the gate. Tuning the frequencies of individual qubits is also useful for avoiding any unwanted resonances. The lifetime of a qubit (i.e. the time that it remains in the $|1\rangle$ state before decaying to the $|0\rangle$ state) turns out to be highly frequency-dependent in a complicated and time-dependent way that is still not fully understood (see e.g. [103, 179]). Being able to tune the qubit's frequency helps optimize its lifetime.

The mechanism used to control a qubit's frequency is relatively simple in principle and is described, e.g., in [33]. Consider two Josephson junctions in parallel, as depicted in Figure 1.2. Two Josephson junctions in parallel are known as a superconducting quantum interference device, or SQUID.

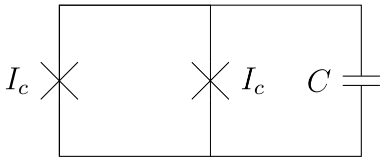


Figure 1.2: A tunable frequency transmon.

Current will flow through the two Josephson junctions to exactly cancel any external magnetic flux through the loop. In particular, if the phases of the two junctions are φ_1 and φ_2, and Φ_{ext} is the external magnetic flux, then

$$\varphi_2 - \varphi_1 = \Phi_{\text{ext}}/\bar{\Phi}_0. \tag{1.18}$$

Meanwhile, assuming for simplicity that I_c is the same for both junctions, the energy stored in the junctions is

$$\begin{aligned} U &= -E_J\left(\cos\varphi_1 + \cos\varphi_2\right) \\ &= -2E_J \cos\left(\frac{\varphi_1+\varphi_2}{2}\right)\cos\left(\frac{\varphi_1-\varphi_2}{2}\right) \\ &= -2E_J\cos\left(\Phi_{\text{ext}}/\bar{\Phi}_0\right)\cos\varphi, \end{aligned} \tag{1.19}$$

where $\varphi = \frac{\varphi_1+\varphi_2}{2}$. Note that because the circuit elements are in parallel, $\dot{\varphi}_1 = \dot{\varphi}_2 = \dot{\varphi} = V$, so φ is conjugate to n as before. Thus, we have the same Hamiltonian as before, except with $E_J \to 2E_J \cos\left(\Phi_{\text{ext}}/\bar{\Phi}_0\right)$. Therefore, to leading order in perturbation theory, the $0 \to 1$ transition frequency is

$$\omega_{10} = \sqrt{8E_{J,\text{max}} E_C \cos\left(\Phi_{\text{ext}}/\bar{\Phi}_0\right)} - E_C. \tag{1.20}$$

Where, in this case, $E_{J,\text{max}} = 2E_J$. This shows that, by controlling a D.C. external magnetic flux through the SQUID, one can tune the qubit frequency.

1.2.3 Single-qubit gates

So far, we have seen how to perform Z rotations, either virtually through a change of frame, or physically by detuning the qubit frequency. This subsection shows how to perform single-qubit gates. The discussion here follows that of [33]. Consider a qubit that is capacitively coupled to a voltage source, as shown in Figure 1.3.

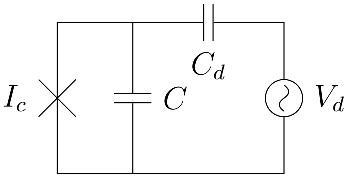


Figure 1.3: Circuit for driving a quibt.

Written in terms of $\dot{\varphi}$, and setting $\bar{\Phi}_0 = 1$ for convenience, the energy in the capacitors is now

$$K = \frac{1}{2} C \dot{\varphi}^2 + \frac{1}{2} C_d \left(V_d(t) - \dot{\varphi}\right)^2. \tag{1.21}$$

The conjugate momentum to φ is therefore

$$\begin{aligned}\tilde{n} &= \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} \\ &= C\dot{\varphi} - C_d(V_d(t) - \dot{\varphi}) \\ &= C_\Sigma \dot{\varphi} - C_d V_d(t),\end{aligned} \tag{1.22}$$

where $C_\Sigma = C + C_d$. This leads to the Hamiltonian

$$H = \frac{\tilde{n}^2}{2C_\Sigma} + \frac{C_d}{C_\Sigma} V_d(t)\tilde{n} - E_J \cos\varphi, \tag{1.23}$$

after dropping a term proportional to V_d^2 because it does not contain any quantum operators. If the drive frequency is on resonance with the qubit frequency, it can be used to excite $|0\rangle \leftrightarrow |1\rangle$ transitions.

To gain some intuition for the effect of the driving term, let's truncate to the qubit subspace. Without the driving term, the qubit Hamiltonian is

$$H_0 = -\frac{\omega}{2}\sigma_z, \tag{1.24}$$

where ω is the qubit frequency. The driving term is

$$\begin{aligned}H_1 &= \frac{C_d}{C_\Sigma} V_d(t)\tilde{n} \\ &= -i\frac{C_d V_d(t)}{2C_\Sigma}\left(\frac{E_J}{2E_C}\right)^{1/4}\left(a - a^\dagger\right) \\ &\to \frac{C_d V_d(t)}{2C_\Sigma}\left(\frac{E_J}{2E_C}\right)^{1/4}\sigma_y \\ &= \Omega f(t)\sigma_y\end{aligned} \tag{1.25}$$

where now $E_C = e^2/(2C_\Sigma)$ and $V_d(t) = V_0 f(t)$. $\Omega = \frac{C_d V_0}{2C_\Sigma}\left(\frac{E_J}{2E_C}\right)^{1/4}$. Next, we write

H_1 in the interaction picture:

$$\begin{aligned} H_{\text{int}} &= \Omega f(t) e^{i\omega\sigma_z/2}\sigma_y e^{-i\omega\sigma_z/2} \\ &= \Omega f(t)\left(\cos(\omega t)\sigma_y + \sin(\omega t)\sigma_x\right). \end{aligned} \tag{1.26}$$

Next, suppose that $f(t) = \sin(\omega t + \phi)$. Expanding the sines and cosines in terms of $e^{i\omega t}$ and $e^{-i\omega t}$ and discarding the terms that rotate at a frequency of 2ω (the rotating wave approximation), we arrive at

$$H_{\text{int}} \approx \frac{\Omega}{2}\left(\sigma_x \cos\phi + \sigma_y \sin\phi\right). \tag{1.27}$$

That is, the interaction causes the qubit to rotate on the Bloch sphere at an angular frequency of Ω, known as the Rabi frequency, about an axis determined by the phase of the driving signal: $\hat{n} = \cos\phi\hat{x} + \sin\phi\hat{y}$. For example, by setting $\phi = 0$ and applying the pulse for a time $\tau = \pi/\Omega$, we can implement an X gate. Similarly, setting $\phi = \pi/2$, we can implement a Y gate. Now, we see that we can implement arbitrary single qubit gates. Care must be taken in shaping the pulse $f(t)$ to avoid exciting transitions to higher states of the transmon. For futher details, see Ref. [121].

1.2.4 Readout

Qubits are measured using a technique known as dispersive readout, which is detailed, for example, in Refs. [33, 148]. A qubit is capacitively coupled to a linear LC oscillator, called the readout resonator, as depicted in Figure 1.4.

Following the same procedures to obtain the Hamiltonian for this circuit, the capacitive coupling leads to an interaction term between the qubit and the oscillator

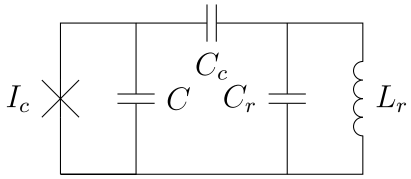


Figure 1.4: A qubit coupled to a linear resonator, the setup used for dispersive readout.

of the form

$$H_g = g(a^\dagger - a)(\sigma^- - \sigma^+), \tag{1.28}$$

where a is the oscillator's lowering operator and σ_- is the qubit's lowering operator. g is a function of the capacitances and the qubit and oscillator frequencies, ω_q and ω_r, given by Eq. 2.93 in Ref. [33]. This interaction term modifies the joint eigenstates of the qubit-resonator system. In particular, the energy required to excite the oscillator depends on the state of the qubit. The dispersive shift, χ, is defined as

$$\begin{aligned} 2\chi &= \omega_r^{|1\rangle} - \omega_r^{|0\rangle} \\ &= \left(E_{|1,1\rangle} - E_{|1,0\rangle}\right) - \left(E_{|0,1\rangle} - E_{|0,0\rangle}\right), \end{aligned} \tag{1.29}$$

where $E_{|n_q,n_r\rangle}$ is the eigenenergy of the eigenstate connected to the state, which, in the absence of H_g, has n_q excitations in the transmon and n_r excitations in the readout resonator. If one makes the rotating wave approximation to eliminate the terms that do not conserve the total number of excitations and then computes the eigenenergies using second-order perturbation theory, the dispersive shift is found to be

$$2\chi = \frac{-2g^2\eta}{\Delta^2(1-\eta/\Delta)}, \tag{1.30}$$

where $\eta = \omega_{21} - \omega_{10}$ is the qubit's anharmonicity, and $\Delta = \omega_q - \omega_r$ is the detuning of the qubit away from the resonator's frequency.

The readout resonator is connected to a transmission line via another capacitor with capacitance C_κ, which determines the coupling, κ between the readout resonator and the external electronics. Microwave signals are sent through the transmission line at approximately the readout resonator frequency, which are partially reflected and partially transmitted. This serves as a probe of the resonator's frequency, which, as we have seen, depends on the state of the qubit. This is how measurement is performed. If the qubits have tunable frequencies, the detuning Δ can be kept large to preserve the qubit's coherence and then decreased when a measurement is performed, in order to increase the dispersive shift. κ controls how fast readout can be performed; a larger κ corresponds to faster readout. However a larger κ also means more coupling between the qubit-resonator system and the outside environment, which causes decoherence. This can be partially mitigated by inserting a bandpass filter at the qubit frequency (see Ref. [148] for a detailed discussion).

1.2.5 Two-qubit gates

Different architectures exist for coupling transmons. IBM uses a conceptually simple cross-resonance method to couple transmons [34]. Each qubit has a microwave drive, and to couple two qubits, one simply drives one qubit (the control) at the resonant frequency of an adjacent qubit (the target). This can be used to implement CNOT gates. However, it can lead to unwanted crosstalk, as other qubits may also be excited.

A more involved way to control the interactions of qubits is by introducing another component called a coupler. One design is the inductively-coupled “gmon” tunable coupler, invented by John Martinis’ group, which became Google Quantum AI [32]. These tunable couplers allow the user to turn off interactions when they are not wanted as well as to achieve the entire family of fSim gates. The circuit for a gmon coupler is depicted in Figure 1.5. Tunable couplers can also be constructed that utilize capacitive coupling between the qubits (Ch. 8 of [128], [177]), which is the coupler design that is implemented in Google’s Sycamore processors [5]. However, the inductive design is described here.

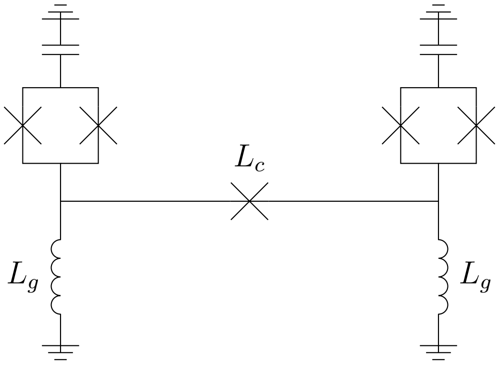


Figure 1.5: A tunable gmon coupler between two qubits.

The two qubits are the SQUIDs (parallel Josephson junctions), in series with their respective capacitors. The inductance, L_c, of the coupler Josephson junction can be tuned by applying an external magnetic flux. In particular, from the Josephson equations (Eq. 1.7), the inductance is

$$L_c = \frac{\bar{\Phi}_0}{I_0 \cos \delta}, \tag{1.31}$$

where δ is the phase difference across the junction, which can be controlled with an external DC magnetic flux. Furthermore, there is a mutual inductance between the two qubits that is controlled by L_c. In particular, assuming that $L_g \ll L_J$ and $L_g \ll L_c$, if a current I_q is sent down through the left qubit, most of it goes straight through the left L_g to the ground, whereas a small fraction of it, $I_{\text{cp}} = I_q L_g/(2L_g + L_c)$ flows through L_c to the right half of the circuit. Most of this (all but $O(L_g)$) flows through the right L_g to ground. The current through the right L_g creates a magnetic flux $L_g I_{\text{cp}}$ through the right SQUID, which is positioned close to it. In particular, the mutual inductance, to leading order in L_g, is

$$M = \frac{L_g^2}{2L_g + L_c}, \tag{1.32}$$

where, as mentioned earlier, L_c is tunable via an external magnetic flux. The mutual inductance creates a coupling between the two qubits, which, to leading order, is:

$$H_{\text{int}} = -\frac{E_J M}{L_J + L_g}\varphi_1\varphi_2. \tag{1.33}$$

Applying the rotating wave approximation, which drops the non-number-conserving terms, the interaction Hamiltonian becomes

$$\begin{aligned} H_{\text{int}} &= -\frac{\omega}{2}\frac{M}{L_J + L_g}\left(a_1^\dagger a_2 + a_1 a_2^\dagger\right) \\ &= g\left(a_1^\dagger a_2 + a_1 a_2^\dagger\right) \end{aligned} \tag{1.34}$$

where a_i is the annihilation operator for qubit i, and ω is the qubit frequency. This Hamiltonian generates swapping between the qubits, i.e. oscillations between the $|01\rangle$ and $|10\rangle$ states at an angular frequency of

$$g = -\frac{\omega}{2}\frac{L_g}{L_J + L_g}\frac{L_g}{2L_g + \bar{\Phi}_0/(I_0\cos\delta)}. \tag{1.35}$$

In addition to swapping, H_{int} shifts the energy of the $|11\rangle$ state by connecting it to the $|02\rangle$ and $|20\rangle$ states. More specifically, if the frequency detuning between the two qubits is $\Delta = \omega_2 - \omega_1$ (where ω_i is the frequency of qubit i), then to second order in g, the energy of the $|01\rangle$ and $|10\rangle$ states shift as

$$\begin{aligned} E_{|01\rangle} &= \omega_1 + \frac{g^2}{\Delta} \\ E_{|10\rangle} &= \omega_2 - \frac{g^2}{\Delta} \end{aligned} \tag{1.36}$$

while the energy of the $|11\rangle$ state shifts as

$$\begin{aligned} E_{|11\rangle} &= \omega_1 + \omega_2 + \frac{2g^2}{\eta - \Delta} + \frac{2g^2}{\eta + \Delta} \\ &= \omega_1 + \omega_2 + \frac{4g^2\eta}{\eta^2 - \Delta} \end{aligned} \tag{1.37}$$

Therefore, in the frame rotating at the qubit frequency, the $|11\rangle$ state acquires a phase at a rate of

$$\omega_{11} = \frac{4g^2\eta}{\eta^2 - \Delta}. \tag{1.38}$$

Thus, gmons naturally generate fSim gates (Eq. 1.6), with $\theta \sim g\tau$ and $\phi \sim g^2\tau$, where τ is the duration of the pulse. In particular, appropriately choosing g and τ to set $\theta = 0$ and $\phi = \pi$, CZ gates can be achieved. One has to be careful to avoid leakage from the $|11\rangle$ state to the $|20\rangle$ and $|02\rangle$ states.

The fact that tunable couplers naturally give rise to a continuous family of native fSim gates was utilized in Ref. [119] (see Section II of the supplementary information to that paper, in particular) and in Refs. [115, 74, 86], which built on the techniques developed in that work.

Transmon-based quantum processors do not have native all-to-all coupling among their qubits. The couplers in Google Sycamore processors enable the qubits to be coupled in a 2D cartesian grid, ideal for implementing the surface code. IBM has chosen instead to couple qubits in what they call a heavy-hex lattice, which consists of connected 12-qubit loops (one for each vertex and each edge of the tiled hexagons). Although gates between nonadjacent qubits are not available natively, they can be decomposed into local gates, in particular by applying SWAP gates until the qubits are adjacent, performing the desired operations, and then swapping them back. This is a disadvantage compared to trapped ion qubits, which feature native all-to-all connectivity.

1.3 Characterizing errors

Unlike ideal quantum computers, real quantum computers are subject to a variety of sources of error. Their coupling to the environment, including through the readout resonator, causes amplitude damping, or decay from the $|1\rangle$ state to the $|0\rangle$ state. This is characterized through the decay time, T_1. Further, they can be subject to various sources of error that suppress the off-diagonal components of the density matrix (or, equivalently for a single qubit, the x and y components of the Bloch sphere representation of the state). These types of errors are known as dephasing and are characterized by the time T_2, the decay constant of the length of the projection of the Bloch vector onto the xy plane. Amplitude damping contributes to T_2, but so do other things. For example, for frequency tunable transmons, fluctuations in the external magnetic flux, caused by noise in the electronics, contributes to T_2 by applying random Z rotations. Furthermore, leakage out of the qubit subspace and miscalibration of gates are sources of errors. Error rates on existing devices can fluctuate on a timescale of hours, so fast characterization of errors and recalibration is important.

The simplest model of an error is a depolarizing channel, which erases all of the information in the state with probability p. An n-qubit depolarizing channel takes the form

$$\Phi_{\text{dep}}(\rho) = (1-p)\rho + \frac{p}{2^n}\mathbb{I}. \tag{1.39}$$

This can be written as a Pauli channel (a quantum channel where all of the Kraus operators are Pauli operators), using the fact that

$$\mathbb{I} = \frac{1}{4^n}\sum_{i=0}^{4^n-1} P_i \rho P_i, \tag{1.40}$$

where P_i is an n-qubit Pauli operator. This channel corresponds to randomly applying Pauli operators, with equal probabilities for each. To prove Eq. 1.40, expand the trace-1 Hermitian matrix ρ in the basis of Pauli operators and note that, for each non-identity single-qubit Pauli operator in the expansion of ρ, half of the P_i will commute with it and the other half will anti-commute.

Therefore, we can also write the depolarizing channel as

$$\Phi_{\text{dep}}(\rho) = \left(1 - p\left(1 - \frac{1}{4^n}\right)\right)\rho + \frac{p}{4^n}\sum_{i=1}^{4^n-1} \hat{P}_i \rho \hat{P}_i. \tag{1.41}$$

This leads to some confusing terminology. Sometimes, reported error rates are the depolarizing probability p, and sometimes, they are the "Pauli error," related to the depolarizing error by

$$p_{\mathrm{Pauli}} = p\left(1 - \frac{1}{4^n}\right). \tag{1.42}$$

There are many techniques for measuring various types of errors in quantum processors. Here I discuss only a few important examples.

1.3.1 Randomized benchmarking

Randomized benchmarking (RB) is a standard way to characterize error rates [49, 40], used at both IBM and Google. At Google, it is typically used for single-qubit error rates, whereas cross-entropy benchmarking (XEB) is used to characterize two-qubit gates. First, a random sequence of Clifford gates is chosen. By the Gottesman-Knill theorem, the Clifford gate that inverts the sequence can be found efficiently and is appended to the sequence. Therefore, ideally, the measurement outcome should be 0. Assuming that the noise channel is a depolarizing channel, the probability of measuring something other than 0, p_{RB}, is related to p by

$$p_{\mathrm{RB}} = p\left(1 - \frac{1}{2^n}\right). \tag{1.43}$$

However, instead of using a single sequence to extract p, one uses random sequences of varying lengths and fits an exponential decay to the probability of measuring 0 as a function of the sequence length. From the decay constant, one extracts the error rate. This allows one to characterize gate errors separately from state preparation and measurement (SPAM) errors.

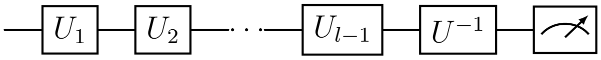


Figure 1.6: A length-l randomized benchmarking circuit. Each of the U_i is a random Clifford gate.

1.3.2 Cross-entropy benchmarking

Randomized benchmarking may not be convenient if one wants to benchmark a two-qubit gate that is not simply related to Clifford gates. For example, if the goal is to benchmark a particular fSim gate, randomized benchmarking may not be convenient. A better choice is to use cross-entropy benchmarking (XEB) [15].

XEB is similar to RB in that it assumes a depolarizing noise channel and characterizes the noise using random circuits of varying lengths. The gate sequence is shown in Figure 1.7. Note that the sequence is not inverted, like it is in RB.

Under a depolarizing channel, the expectation value of any observable $\hat{O}$ decays towards its trace:

$$\langle \hat{O} \rangle_{\text{dep.}} = (1-p)\langle \hat{O} \rangle + \frac{1-p}{2^n} \text{Tr}\left[\hat{O}\right]. \tag{1.44}$$

For XEB, one first performs a classical statevector simulation of the gate sequence (which is possible since it only involves two qubits) and obtain the ideal probabilities

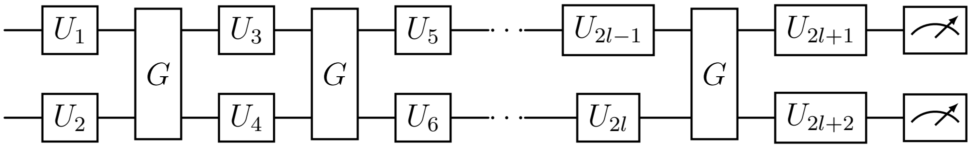


Figure 1.7: A length-l cross-entropy benchmarking circuit. Each of the U_i is a random single-qubit gate, and G is the two-qubit gate that we would like to benchmark.

for each of the four measurement outcomes. The operator $\hat{O}$ is then chosen to be

$$\hat{O} = \mathrm{diag}(\rho_{\mathrm{ideal}}) = \sum_i p_i^{\mathrm{ideal}} |i\rangle\langle i|. \tag{1.45}$$

Eq. 1.44 then becomes

$$\begin{aligned} \sum_i p_i^{\mathrm{obs}} p_i^{\mathrm{ideal}} &= (1-p) \sum_i \left(p_i^{\mathrm{ideal}}\right)^2 + \frac{p}{2^n} \\ \implies p &= \frac{\sum_i p_i^{\mathrm{ideal}} \left(p_i^{\mathrm{ideal}} - p_i^{\mathrm{obs}}\right)}{\sum_i \left(p_i^{\mathrm{ideal}}\right)^2 - \frac{1}{2^n}}, \end{aligned} \tag{1.46}$$

which shows how to obtain the depolarizing channel probability from measurements of the probabilities of the observed bitstrings. As in randomized benchmarking, this is repeated for sequences of varying lengths, and the error per gate is extracted from fitting a decaying exponential to $1-p$. If the single qubit error rate is much lower than the two-qubit error rate, which is typically the case, this error rate can be assigned to the two qubit gates. Otherwise, one may need to subtract the error from the single-qubit gates.

1.3.3 Readout error rates

Readout errors are significant (often percent level) on existing transmon-based quantum processors. Because single-qubit error rates are small compared to readout error rates, the latter can be benchmarked simply by repeatedly preparing the qubit in either $|0\rangle$ or $|1\rangle$ and measuring the probability of obtaining the wrong measurement outcome. The readout error rate when the qubit is prepared in $|0\rangle$ is called e_0. Similarly, e_1 is the readout error rate when the qubit is prepared in $|1\rangle$.

Several factors contribute to readout errors. One is simply misidentification of the $|0\rangle$ and $|1\rangle$ states. As detailed in Ref. [148], the raw measurement from the readout line consists of an in-phase (I) and quadrature (Q) component of the scattered signal (the amplitudes of the sine and cosine components of the signal). The qubit state affects the frequency of the resonator, which affects the phase of the scattered signal. Thus, a region of the IQ plane is associated with the $|0\rangle$ state, while a different region is associated with the $|1\rangle$ state. More specifically, the measured point in the IQ plane is random with an approximately Gaussian distribution, where the two Gaussians (associated with the $|0\rangle$ and $|1\rangle$ state, respectively) have different centers. The tails of the Gaussians overlap, causing some misidentification error. See Sec. 3.5 of Ref. [148] for more details. This misidentification error can be made symmetric or can be made to reduce one type of error at the expense of the other depending on where the boundary between the two regions is drawn. However, other factors, such as decays from the $|1\rangle$ to the $|0\rangle$ state during the measurement, including measurement induced transitions, generally cause e_1 to be higher than e_0.

Typically, multiple qubits share the same readout line. This and other factors can cause correlations among the readout errors, and can also make simultaneous readout errors (i.e. readout error rates when multiple qubits are measured simultaneously) higher than isolated readout error rates. This is typically the case on existing devices. To benchmark simultaneous readout errors, one can prepare the qubits in random bitstring states, measure the qubits, and extract the single-qubit readout error rates.

1.3.4 T_1 and T_2

To measure the amplitude damping rate, T_1, simply prepare the qubit in the $|1\rangle$ state, wait a variable time, and measure. Fit an exponential to the probability of measuring 1 as a function of time.

To measure the dephasing rate, T_2, simply prepare the qubit in the equatorial plane of the Bloch sphere, e.g. by applying a $\sqrt{X}$ or $\sqrt{Y}$ gate, wait a variable amount of time, and measure the projection of the Bloch vector in the equatorial plane. (This is known as a Ramsey sequence.) Fit a decaying exponential to $\sqrt{\langle X\rangle^2 + \langle Y\rangle^2}$; the decay constant is T_2. T_2 can be significantly improved by *dynamical decoupling*, where, instead of leaving a qubit idle, an echo sequence of pulses is applied. This echo sequence can reduce the effect of slow fluctuations in the qubit's frequency caused by varying external magnetic flux. See Section 6.5 of Ref. [33]. Because amplitude damping also contributes to dephasing, $T_2 \leq 2T_1$.

1.3.5 State preparation errors

Qubits may also have state preparation errors, i.e. the initial state may not be $|0\rangle$. Coherent deviations from $|0\rangle$ can be detected by first applying a variable Z rotation and then measuring $\langle X\rangle$ or $\langle Y\rangle$. Oscillations indicate that state preparation was not perfect. See Ref. [101] for more details on how to measure state preparation errors independently from readout errors.

1.4 NISQ Algorithms

The algorithms discussed in Section 1.1 require very low error rates, which will require some form of error correction, i.e. redundantly encoding logical qubits in multiple physical qubits and detecting errors, perhaps by measuring stabilizer operators. This is currently far out of reach. However, other algorithms exist that are being implemented on existing devices and could possibly lead to useful applications. In this section, I describe a few of the these algorithms suitable for noisy intermediate-scale quantum (NISQ) devices.

1.4.1 Trotterized time evolution

One of the promising applications of existing quantum processors is to simulate other quantum systems. Suppose that the system to be simulated is described by a Hamiltonian, H, and we want to implement e^{-iHt}, or an approximation to it. Suppose that there is not a simple exact decomposition of e^{-iHt} in terms of the available single and two-qubit gates, but that $H = H_A + H_B$, where $[H_A, H_B] \neq 0$, and we know how to implement e^{-iH_At} and e^{-iH_Bt} in terms of the basic gates. Then we can achieve an approximation to the desired time evolution as

$$e^{-iHt} \approx \left(e^{-iH_A\delta t}e^{-iH_B\delta t}\right)^{t/\delta t} \tag{1.47}$$

From Taylor expanding both sides, it is easily shown that

$$e^{-iH_A\delta t}e^{-iH_B\delta t} = e^{-iH\delta t} - \frac{1}{2}[H_A, H_B]\delta t^2 + O(\delta t^3). \tag{1.48}$$

Thus, the overall error is

$$\left(e^{-iH_A\delta t}e^{-iH_B\delta t}\right)^{t/\delta t} - e^{-iHt} = -\frac{1}{2}[H_A, H_B]t\delta t. \tag{1.49}$$

The $O(\delta t)$ error can be eliminated by instead using the second-order Trotter expansion:

$$e^{-iHt} \approx \left(e^{-iH_A\delta t/2}e^{-iH_B\delta t}e^{-iH_A\delta t/2}\right). \tag{1.50}$$

Chapter 5 of this thesis implements Floquet dynamics that are related to the Trotterization of the 1D Heisenberg model.

1.4.2 Adiabatic state preparation

In addition to studying the dynamics of quantum systems, it is also useful to study their eigenstates, or mixtures of eigenstates such as thermal states. The goal could either be to study these states themselves, or to prepare them as initial states, e.g. to compute a multi-time correlation function in a particular state.

Adiabatic state preparation utilizes the adiabatic theorem from quantum mechanics, that states that, if the system begins in an eigenstate of H_i, and the Hamiltonian is slowly changed from H_i to H_f, then the system ends up in the corresponding eigenstate of H_f. For example, H_i could be chosen so that its ground state is easy to prepare. Then, a Trotterized time evolution of a time-dependent Hamiltonian that interpolates between H_i and H_f is applied. The final state is an approximation to the ground state of H_f. If the gap between the ground state and the first excited state is small at any time during the evolution, then there is a risk of exciting out of the ground state if the time evolution is not slow enough. On the other hand, a slower time evolution requires deeper circuits and hence more noise.

1.4.3 The variational quantum eigensolver

While adiabatic state preparation provides a circuit to approximately prepare the ground state of a given Hamiltonian, H_f, it may not be the shallowest circuit that does so to a given fidelity, and optimizing for circuit depth is very important on current noisy devices. The variational quantum eigensolver (VQE) attempts to find a shallower circuit to prepare the ground state by postulating an ansatz (i.e. a parameterized quantum circuit, which generates a parameterized trial wavefunction). The energy is measured and is fed into an optimization routine, run on a classical computer, that tunes the variational parameters in order to minimize the energy. After many iterations, this may converge to an approximation to the ground state. The ansatz can be chosen to have a similar structure to the circuit used in adiabatic state preparation, so it is guaranteed to perform no worse than adiabatic state preparation in terms of circuit depth. VQE is implemented in Chapter 2 of this thesis.

CHAPTER 2

VQE AND ERROR MITIGATION

2.1 Introduction

Quantum computation promises to revolutionize physics and chemistry. Our ability to make quantitative predictions about many-body and strongly coupled quantum mechanical systems, such as atoms with many electrons, or systems of quarks and gluons, has, with classical computers, been severely hampered by the exponential memory requirements of storing and operating on many-body wavefunctions. Quantum computers overcome this difficulty by representing the wavefunction in the state of the qubits themselves. Noise, however, has prevented current quantum computers from surpassing the capabilities of classical computers in almost all applications, including quantum simulation.

In the long term, fault-tolerant quantum computation may be achieved using error-correction techniques that employ many physical qubits to encode each logical qubit. In the near term, there is much effort aimed at less costly methods to mitigate or compensate for the effects of noise in current quantum hardware. Some of these error-mitigation techniques involve artificially increasing the noise and extrapolating the output to zero noise [102, 159, 50, 84, 131, 70]. Others involve representing the ideal circuit in terms of noisy ones (the so-called "quasi-probability method") [159, 50]. Still others involve using replica copies of the system to effectively take powers of the density matrix and suppress errors [97, 75], while others attempt to learn the effect of the noise by training on classically simulable circuits [70, 38, 118, 152].

Several proposed error-mitigation methods have successfully modeled the noise as damping the expectation values of observables towards their expectation values in maximally mixed states, i.e., $\langle\hat{O}\rangle_{\text{noisy}} \approx C\langle\hat{O}\rangle_{\text{exact}} + (1-C)\text{Tr}[\hat{O}]/2^n$, where $0 < C < 1$, $\hat{O}$ is an observable, and n is the number of qubits [70, 38, 118, 61, 116, 163, 168, 167, 152]. In this work, we consider traceless observables, so the last term is not present (although its inclusion would not present any significant difficulties). Knowing the damping factor, C, the exact expectation value can be recovered from the noisy one. Various methods for estimating the damping factor are proposed in the works cited above. Here we propose another method and compare the performance of these methods to each other and to zero noise extrapolation when applied to the test problem of measuring the ground state energy of a 1D mixed field Ising spin chain. We find that several methods offer significant improvements over no mitigation, enabling circuits of up to about 25 ansatz layers to yield expectation values within 10% of the correct expectation value. Which method works best is problem-dependant, but for our test problem, we find that estimating the damping factor from the perturbative regime of the Hamiltonian works particularly well.

2.2 Test problem

Our goal is to benchmark error mitigation methods, in order to get a sense for when excessive circuit depth impedes obtaining meaningful results. To benchmark these methods, we consider the problem of finding ground states of the 20-qubit 1D mixed-field Ising model.

The mixed-field Ising model is defined by the following Hamiltonian:

$$H = -J \sum_i Z_i Z_{i+1} - h_x \sum_i X_i - h_z \sum_i Z_i, \tag{2.1}$$

where we impose cyclic boundary conditions on the spins. If $h_z = 0$, this reduces to the transverse-field Ising model, which is exactly solvable using a Jordan-Wigner transformation when the number of qubits is large. In this case, a term arising from the cyclic boundary conditions can be neglected [133]. It is, of course, also solvable in the limits $h_x^2/(J^2 + h_z^2) \ll 1$ and $J^2/h_x^2 \ll 1$. For intermediate values of the field strengths, however, it is non-integrable and quantum chaotic [89, 90]. Approximate methods have been developed to study this model outside of its integrable regime [175]. For the system size studied here (20 qubits), such approximate methods are not necessary since the Hamiltonian can be exactly diagonalized on a classical computer. For a discussion of the perturbative regimes of this Hamiltonian, see Appendix A.1. Fig. A.1 compares the exact ground state energies of this model to those computable in second-order perturbation theory from the analytically solvable limits. Throughout this work, we set $J = 1$, which sets the overall scale of H.

We use an ansatz to parameterize approximate ground states of this Hamiltonian.

In particular, we consider a variant of the Alternating Layered Ansatz (ALT) [30, 122], shown in Fig. A.2 and described in Appendix A.3. We consider both the full ALT ansatz and a constrained version of it, with even cyclic permutation symmetry imposed. The permutation symmetry reduces the number of parameters from $n(l+1)$ to $2(l+1)$, where n is the number of qubits and l is the number of ansatz layers. Fig. A.3 shows that this ansatz is able to capture the ground states of these Hamiltonians as the number of layers is increased. In fact, we do not need to make l very large to capture the ground state, but we do so anyway to study error mitigation.

For future applications of the Variational Quantum Eigensolver (VQE), in which the ansatz circuits are wide and deep enough that they are not classically simulable, the optimization step will need to be performed using measurements from the quantum computer. If the effect of the noise is indeed to damp expectation values of observables, and the parameter dependence of the damping factor does not noticeably shift the location of the minimum, then it is not necessary to apply error mitigation during the optimization routine. In this case, correcting the damping is more relevant after the optimization, when we evaluate the optimized circuit. We demonstrate that our symmetrized ansatz circuits can be optimized for the mixed-field Ising model on 20 qubits using unmitigated measurements from quantum computers (Figs. 2.1 and 2.2). Having shown this, and because our primary goal is to benchmark error-mitigation techniques, we proceed to optimize the remaining circuits classically. In all Figures other than 2.1 and 2.2, the results were obtained with circuits that were optimized classically.

2.3 Error mitigation techniques

We consider the following questions: Given optimized circuits that approximate ground states of the Hamiltonian (2.1), how can we best estimate the damping factor that characterizes the suppression of the expectation value of the Hamiltonian when evaluated on a quantum computer? How deep can we make our ansatz using this method? The remainder of this section contrasts techniques for estimating the damping factor. We do not include the method of Ref. [168], which requires simultaneous measurements of all of the qubits, because readout-error mitigation on a large number of qubits is a significant challenge beyond the scope of this work. We also do not include the motion-reversal method of Ref. [152] because, for circuits of the depths considered here, repeating them and their inverses several times leads to very small fidelities, to which it is impossible to accurately fit an exponential using a reasonable number of measurements. Further, this method relies on the assumption that the circuit and its inverse lead to similar decays in fidelity, which may not be true. We also do not include the quasi-probability method because it requires gateset tomography. We benchmark the following methods. "Extrapolation in circuit depth" is a new method proposed here, whereas the others have been proposed in previous works. We also tested two methods that did not perform as well, which are described in Appendix A.4.

2.3.1 Zero noise extrapolation

In our implementation of zero noise extrapolation (ZNE) [102, 159, 50, 84, 131, 70], we artificially increase the noise by replacing CNOT gates with odd numbers of CNOT gates. For example, if we want to increase the noise by a factor of 3, we replace each CNOT gate by 3 CNOT gates. We can scale the noise by factors other than odd integers by randomly picking CNOT gates to repeat an odd number of times. We then use an exponential fit to extrapolate the measured energy to zero noise. We use three points (one of which is the unmodified circuit) to do this extrapolation.

2.3.2 Calibrating using the perturbative regime

Recently, the Google team observed that, at least for the specific system they studied, the damping factor was independent of the coupling strength [61]. They proposed, therefore, that one can measure the damping factor in a perturbative regime of the Hamiltonian, and then use the same damping factor outside of the perturbative regime to correct the results from the quantum computer.

This method is, of course, limited to Hamiltonians that have a perturbative regime, which doesn't include all Hamiltonians of scientific interest. For example, instances of the Sachdev-Kitaev-Ye (SYK) Hamiltonian do not have a perturbative regime, but simulations of the model may give insights into holography and quantum gravity (e.g., [7, 114, 108, 18, 129]). Nevertheless, many Hamiltonians of interest do have a perturbative regime, so this technique may still prove useful.

The mixed-field Ising model has several different perturbative regimes, discussed in Appendix A.1. For the purposes of this error-mitigation technique, we consider only perturbation in small h_x; we fix $h_z = 0.1$ and consider $0 < h_x \leq 0.5$ to be the perturbative regime and $h_x = 1.5$ to be outside of the perturbative regime. While we might obtain a better estimate of the damping factor by averaging measurements over multiple perturbative regimes, approaching the desired field strength from both sides, multiple perturbative regimes is not a generic feature of Hamiltonians so we do not employ this trick.

2.3.3 Extrapolation in circuit depth

In contrast to Ref. [61], but in agreement with the observation that exponentials work well for ZNE [50], we expect that the damping factor should decay exponentially in the number of ansatz layers and generally observe this to be the case. Further, circuits with small depth may be classically simulable because only qubits in the backwards light cone of the measured qubits need to be included. For 1D ansätze with nearest-neighbor 2-qubit gates, such as the ansatz considered here, the number of qubits needed for exact classical simulation is $\min(m+2l, n)$, where l is the number of ansatz layers, m is the number of qubits on which the observable acts, assumed to be adjacent, and n is the total number of qubits. If 36 qubits are classically simulable, then for the Hamiltonian considered here (Eq. 2.1), with $m = 2$, we can simulate up to 17 ansatz layers for arbitrarily large n. More generally, one needs to consider the size of the backwards light cone as a function of the number of ansatz layers to determine the maximum number of layers that can be simulated classically. This method will not work as well or at all for observables that act on many qubits in the same term, such as the SYK Hamiltonian in either the Jordan-Wigner or Bravyi-Kitaev [17] representation. To test this mitigation method for our system, we fit an exponential to the damping factor for up to 15 ansatz layers and then use the fit to predict the damping factor for up to 50 layers.

2.3.4 Calibrating by omitting 1-qubit gates

Another recent paper [163] proposed estimating the damping factor by eliminating all of the 1-qubit gates in the circuit. The remaining 2-qubit gates, which are all controlled-NOTs, leave the state in $|00\ldots0\rangle$, and the damping factor is estimated as the fidelity of measuring all zeros. For our system, eliminating the single-qubit gates is achieved by setting $\vec{\theta} = 0$. In addition to estimating the damping factor using the fidelity, as suggested by Ref. [163], we also compare the measured energy to the energy in the ideal state to estimate the damping factor.

This method can be regarded as a special case of a method in which the ansatz is evaluated at classically solvable points. For our ansatz, if we pick the Y-rotations to be $\theta_i = n_i\pi/2$, for integers n_i, then the circuit is within the Clifford group and hence classically simulable. Using Clifford circuits to estimate the damping factor has been proposed previously [38]. For generic choices of the n_i, however, the expectation value of the Hamiltonian becomes vanishingly small — and hence impossible to measure with a finite number of shots — as the number of ansatz layers becomes large. The fidelity also cannot be measured without adding more gates to the circuit because the final state is not generically a computational basis state, so estimating the predicted damping factor from the fidelity is not possible. Therefore, these circuits cannot be used to measure the damping factor using the method described in the previous paragraph. We might instead consider picking all of the n_i to be even, so that the state remains a computational basis state. In generic computational basis states, at large n, the Hamiltonian will still have a small expectation value, although not as small as in a generic state. Nevertheless, we could estimate the damping factor by measuring the fidelity of the final state in such a circuit. In the present work, we focus on the case $\vec{\theta} = 0$, which produces the ground state of the free ($h_x = 0$) Hamiltonian, a state close to the states of interest. We leave the study of other Clifford circuits to future work.

In Ref. [163], this method is combined with ZNE. We also benchmark the combination of this method with ZNE. There are two ways of combining these methods. One could apply ZNE to the original circuit and to the calibration circuit (the one with the single-qubit gates removed) separately and then divide the former result by the predicted damping factor from the latter result (we call this method "ZNE first"), or we could divide the measured energy from each noise-scaled target circuit by the estimated damping factor from the corresponding noise-scaled calibration circuit and then extrapolate the quotient to zero noise (we call this method "ZNE last"). We benchmark both methods.

2.4 Experimental techniques

Independent of the error-mitigation technique under consideration, we always do the following to reduce noise:

1. We omit gates outside of the backwards light cone of the measured observable. This technique has previously been called a *light-cone filter* and has been found to be beneficial [116].
2. In order to implement the previous technique to its fullest advantage, and also to reduce the number of qubits measured simultaneously, we measure each Hamiltonian term separately.
3. The qubits in our ansatz are connected in a loop, as are the qubits on the quantum computer. For a loop of n qubits, there are $2n$ ways to map the circuit qubits to the physical qubits that preserve this connectivity (cyclic permutations and a reflection). We average over four of these mappings. Such an average has been called *qubit assignment averaging* [61]. Instead of picking these mappings randomly, we pick configurations for which the damping factor, as predicted by multiplying fidelities, is the largest (closest to one). This ensures that we avoid particularly noisy qubits or gates to the extent possible while also averaging out systematic biases.

4. Additionally, we implement randomized compiling [170, 26, 27, 163], which helps convert coherent errors into depolarizing-channel errors. In this method, every CNOT gate is randomly replaced by one of 16 dressed CNOT gates, to which it is equivalent in the absence of noise. The 16 dressed gates are described succinctly in Figure 1 of [163]. We apply this method when benchmarking error mitigation methods (i.e. Figs. 2.3-2.6) but not when running the VQE experiments (Figs. 2.1-2.2).
5. Finally, we apply readout error mitigation, assuming that readout errors are uncorrelated and that the error rates are as measured by IBM and reported in the backend properties. For details, see Appendices A.5-A.6.

2.5 Results and Discussion

First, we demonstrate that the training phase of VQE can converge even when we don't apply error-mitigation techniques to correct the damping. Then we compare the performance of the various error-mitigation techniques at estimating the damping factor for tuned circuits.

2.5.1 VQE using unmitigated quantum circuit evaluations

Consider the training phase of VQE. In order to improve the convergence of our optimizer, we impose a cyclic permutation symmetry on the ansatz, so that each row of Y-rotations is described by two angles instead of n. This reduces the number of parameters from $n(l+1)$ to $2(l+1)$, where n is the number of qubits and l is the number of ansatz layers. We begin at a random point in parameter space and use the quantum computer to measure the energy. We feed the measured energy into the Simultaneous Perturbation Stochastic Approximation (SPSA) optimizer, running on a classical computer, which picks the updated parameters [1]. The first 50 function evaluations are used to calibrate an SPSA hyperparameter as in [85], following which the optimizer iterates through the SPSA algorithm. At each iteration, SPSA uses two evaluations of the energy, at $\vec{\theta} = \vec{\theta}_i \pm c_i \Delta_i$, and we submit both of these evaluations together in the same job for improved speed.

Figs. 2.1 and 2.2 show the results of this VQE optimization for a 20-qubit mixed-field 1D Ising model on *ibmq_sydney* and *ibmq_toronto*, respectively, with 3–6 ansatz layers. We see that the optimizer is often able to converge to a state close to the true ground state, that is, below the energy of the first excited state. The lowest energy state achieved during our VQE runs has an energy of -33.54. For comparison, the true ground state energy is -33.90, and the first excited state energy is -32.60. We might worry that this convergence would be spoiled by a parameter-dependent damping factor. Indeed, in some of our runs, the damping factor can be seen to increase as the optimizer converges to the ground state, indicating its parameter dependence.

Although the parameter dependence of the damping factor does not prevent convergence in our system, we cannot guarantee that it never will. If the damping factor shifts the locations of the minima, techniques to minimize this effect will be necessary. These could include additional experimental techniques to symmetrize over measurement outcomes, such as exchanging the roles of “1” and “0”, or applying an error mitigation technique during the optimization routine. Many of the error-mitigation techniques discussed here, however, assume either that the damping factor is parameter independent or that the circuit has been optimized. Of the methods that we benchmark, only ZNE could be used during the optimization phase.

One might naturally guess that some of the parameter dependence of the damping factor comes from asymmetric readout errors. We show in Appendix A.6 that, indeed, the asymmetric readout errors lead to a predictable additive shift of expectation values, and hence a parameter-dependent damping factor (Eq. A.6). However, because the shift is constant, its effect is to shift the overall energy by a constant amount, leaving the location in parameter space of the minimum unchanged. For the mixed-field Ising Hamiltonian, this overall shift is approximately $-n(h_x + h_z)(e_1 - e_0) \approx -1.6$ for $n = 20$, $h_x = 1.5$, $h_z = 0.1$, and $(e_1 - e_0) \approx 0.05$. This is a larger relative shift when the energy is closer to zero. It should therefore have the effect of making the damping factor closer to unity at negative energies further from the minimum (closer to zero). We see instead the opposite effect, so this is not the source of the parameter-dependent damping factor.

2.5.2 Comparison of mitigation techniques

Having established that, in at least some cases, error mitigation is not needed to tune the ansatz parameters, we now compare the performance of the error-mitigation techniques, applied to optimized circuits. For the methods that work by estimating a damping factor, we plot the observed damping factor on the same axes as those predicted by the various methods (Fig. 2.3). Additionally, for all of the methods, we plot the relative error of the reconstructed energy as a function of the number of ansatz layers (Fig. 2.4). Finally, we show to how many layers each method allows us to extend our ansatz circuits while successfully mitigating the energy to within 10% of its true value (Figs. 2.5-2.6).

2.6 Discussion and Conclusion

We have used the variational-quantum-eigensolver algorithm to solve a quantum spin model with up to 20 qubits (spins), and have compared various error-mitigation techniques that can be applied to the VQE output to correct the measured ground-state energy. We see that, for studying ground states of the mixed-field Ising model on 20 qubits on *ibmq_toronto* and *ibmq_sydney*, all of these methods provide improvements over no mitigation. Zero noise extrapolation fails for large circuit depths because the measured energy, even with no added noise, is damped very close to zero, and it is difficult to extrapolate these measurements to zero noise. For the run on *ibmq_sydney*, ZNE first fails at 9 layers (Figs.2.5-2.6). Some of the fits used in our implementation of ZNE are shown in Fig. 2.7 and illustrate the difficulty of extrapolating to zero noise when the initial circuit is deep.

Using the Hamiltonian's perturbative regime to estimate the damping factor worked relatively well, perhaps best out of the methods that we tested. In Fig. 2.3, we see that it correctly predicts the damping factor on both devices most of the time even though it, too, seems to fail when the circuit extends to 50 ansatz layers. The reader may have noticed that the error bars in Fig. 2.4 are larger for this method than for the others. This is because the variance of the damping factor over the five Hamiltonians in the perturbative regime contributes to the statistical uncertainty of the predicted damping factor (see Fig. 2.8). This could be reduced by picking more points in the perturbative regime. However, the fact that the spread of the measured damping factors is larger than one would expect from their individual error bars indicates that this method is unlikely to predict the correct damping factor exactly. Further, this method has a significant overhead, requiring the optimization of the ansatz circuits for each of the perturbative regime Hamiltonians in addition to the Hamiltonian of interest. Its applicability is also limited because, unlike ZNE, it can only be applied to optimized circuits and only to systems that have a perturbative regime.

As seen in Fig. 2.3, the damping factor did decay approximately exponentially in the number of ansatz layers. Therefore, the exponential fit to the points with at most 15 layers worked reasonably well out to about 25 layers, which is better than ZNE performed on *ibmq_sydney* but not as well as the "from pert" method worked. Extrapolating in the number of ansatz layers also only works for optimized circuits, but it does not require the Hamiltonian to have a perturbative regime. It does, however, require the Hamiltonian and the ansatz to be local in order for the shallow-depth ansatz circuits to be exactly classically simulable.

The method of removing the single-qubit gates, like ZNE, does not require the ansatz to be optimized and does not place any restrictions on the Hamiltonian or the ansatz. However, it does assume that the damping factor is independent of the ansatz parameters, which we have seen is not true. Consequently, this method, whether combined with ZNE or not, does not perform as well as the other methods (c.f. Fig. 2.3).

In summary, we have seen that error-mitigation methods that work by estimating a damping factor can outperform zero noise extrapolation and extend feasible circuit depths to about 25 ansatz layers. However, these methods are also more specialized, best suited to specific types of Hamiltonians and ansätze and poorly suited to the optimization phase, during which one could still apply ZNE. The Hamiltonian studied in this work was particularly simple and did not require error mitigation during the optimization phase. However, that will not always be the case.

By carefully studying error mitigation techniques and the circuit depths that they enable on current devices, we are progressing toward the day when quantum comptuers will be a tool for learning new things about physics and chemistry systems.

Data and code availability

The code used to perform the numerical simulations, the experimental runs on IBM's hardware, and the data analysis (including error mitigation) is available at `https://github.com/mcmahon-lab/error_mitigation_vqe`. The data used to produce the figures may be made available upon request.

Contributions

This chapter is adapted from Ref. [142]. I designed and performed the experiment. I also wrote and ran the code that was used to classically optimize the ansatz circuits. Peter McMahon and Paul Ginsparg provided support and guidance and helped with manuscript preparation.

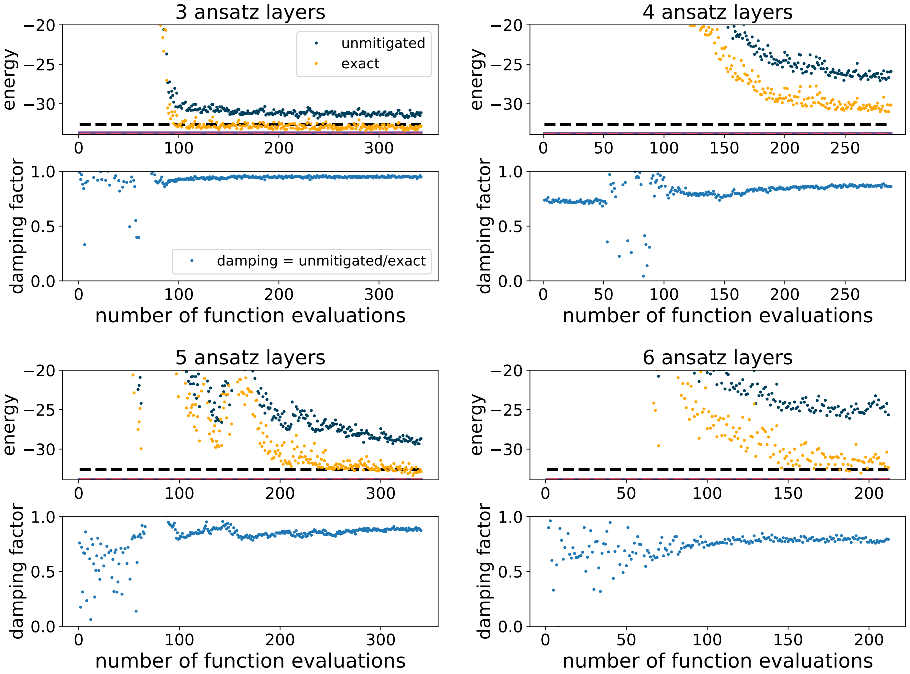


Figure 2.1: VQE for the 20-qubit Ising Hamiltonian, with energy evaluations performed on *ibmq_sydney*, with $h_x = 1.5$ and $h_z = 0.1$. Unmitigated energies are fed into the SPSA algorithm. We impose cyclic permutation symmetry on the ansatz to reduce the number of parameters. Note that SPSA involves computing $E(\vec{\theta}_i \pm c_i\Delta_i)$, not $E(\vec{\theta}_i)$. In the upper panels, the dark-blue dots are the unmitigated energies measured on the quantum computer. These are damped compared to the dark-yellow dots, which show exact classical evaluations of the same circuits, from which we compute the observed damping factor (blue dots in lower panels). The purple horizontal line in each upper panel shows the energy of the classically optimized ansatz circuit, and the magenta dashed line shows the energy of the exact ground state. The horizontal black dashed line indicates the first excited state energy. In all trials except the 4-ansatz-layer trial, the optimizer succeeds in finding a state whose energy is below the first excited state energy (we would pick the point with the lowest energy, not necessarily the last point).

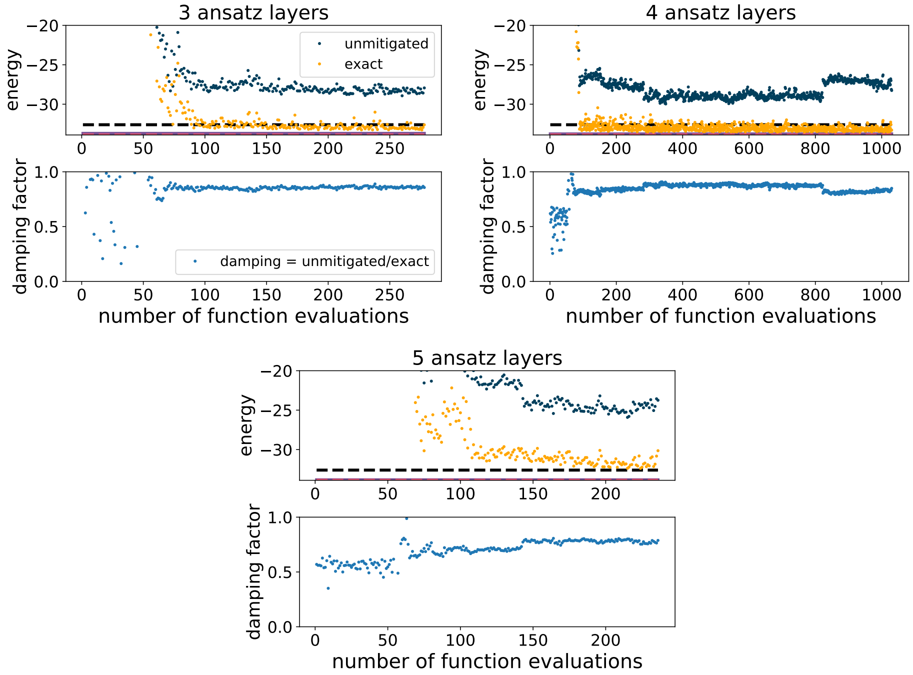


Figure 2.2: Same as Fig. 2.1 except on *ibmq_toronto*. The discontinuities, particularly apparent in the 4-layer experiment, coincide with daily recalibrations of the device. For the 3- and 4-layer trials the optimizer succeeds in obtaining states with energies below the first excited state energy.

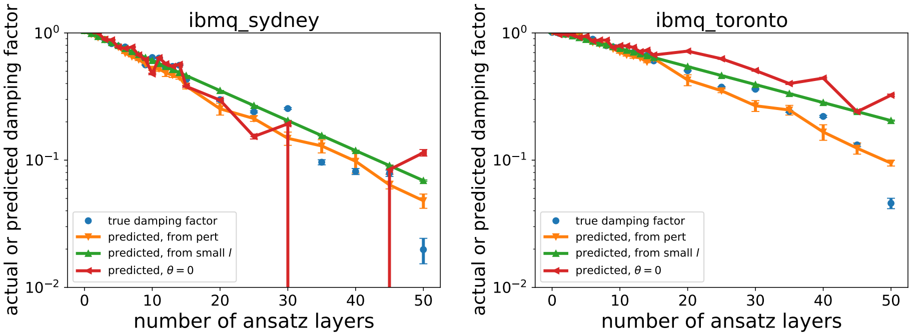


Figure 2.3: Measured and predicted damping factors for ground states of the 20-qubit mixed field Ising model with $h_x = 1.5$ and $h_z = 0.1$. The methods used to predict the damping factor are described in Secs. 2.3.2, and 2.3.3, and 2.3.4, respectively. For the first method, we take $h_x = \{0.1, 0.2, 0.3, 0.4, 0.5\}$ to represent the perturbative regime and average the measured damping factors for those five circuits. ZNE is absent from this plot because it does not predict a damping factor. When computing energies, we apply readout error mitigation using the reported readout error rates, as discussed in Appendix A.6. We see that the "from pert" method correctly predicts the damping factor for circuits up to about 25 layers and that all of the methods fail by 50 layers. Error bars indicate statistical errors only (see Appendix A.2).

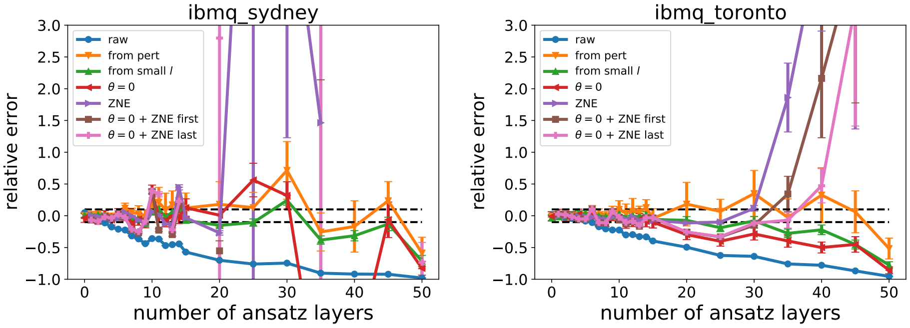


Figure 2.4: Relative error in the measured ground state energy of the 20-qubit mixed field Ising model with $h_x = 1.5$ and $h_z = 0.1$, mitigated using each of the methods studied in this chapter.

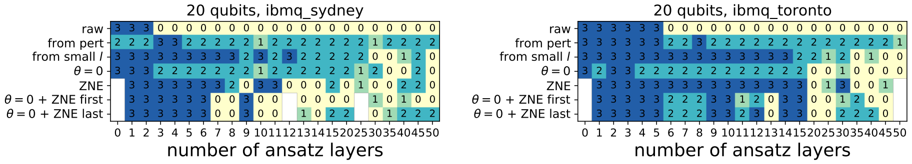


Figure 2.5: Effectiveness of error mitigation. “3” means that the reconstructed energy is within 10% of the correct value. “2” means statistically indeterminate at 1 standard deviation. “1” means statistically indeterminate at 2 standard deviations. “0” means that the reconstructed energy deviated by more than 10% from the true energy. In this figure, the methods involving calibration circuits do not include separate readout error mitigation.

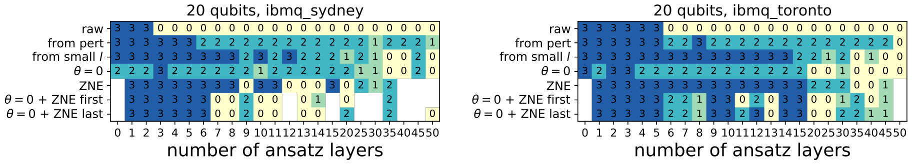
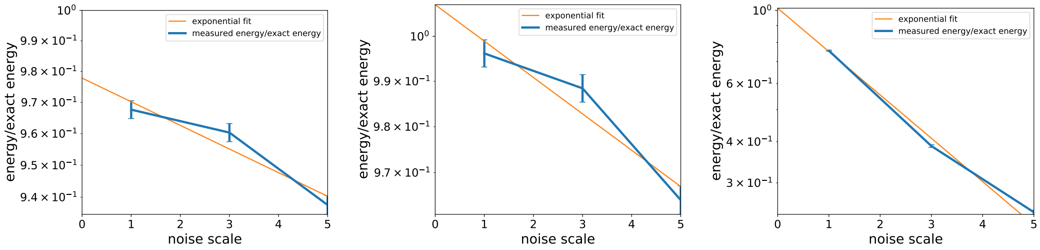


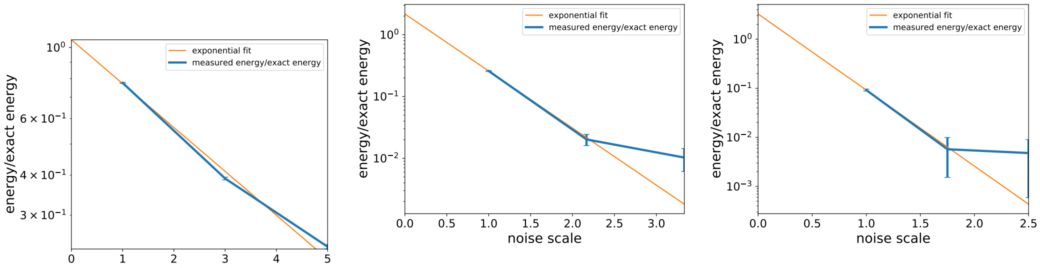
Figure 2.6: Same as Fig. 2.5, but applying readout error mitigation (Appendix A.6) to all energy evaluations, including in calibration circuits.



(a) 1 ansatz layer, without readout error mitigation

(b) 1 ansatz layer, with readout error mitigation

(c) 6 ansatz layers, without readout error mitigation

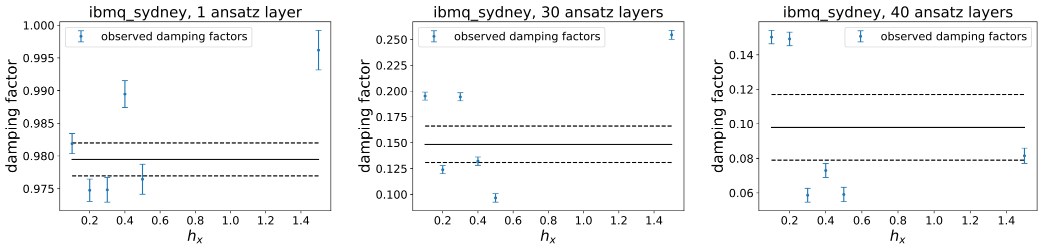


(d) 6 ansatz layers, with readout error mitigation

(e) 30 ansatz layers, without readout error mitigation

(f) 40 ansatz layers, without readout error mitigation

Figure 2.7: Some of the exponential fits used for zero noise extrapolation on *ibmq_sydney*. It is not clear that readout error mitigation, as we perform it, is beneficial for this method. In fact, for 30 and 40 ansatz layers, some of the readout-mitigated energies were negative, causing the exponential fits to fail. In future work, we intend to to explore better ways of performing readout error mitigation. It is also clear why this method fails at large circuit depth. The measured energy is very small even before we artifically increase the noise, and it is therefore difficult to extrapolate back to zero.


ibmq_sydney, 1 ansatz layer

ibmq_sydney, 30 ansatz layers

ibmq_sydney, 40 ansatz layers

Figure 2.8: Some of the extrapolations used when estimating the damping factor from the perturbative regime. The solid horizontal line indicates the predicted damping factor (the mean of the damping factors for $h_x \leq 0.5$), and the dashed horizontal lines indicate the 1σ uncertainty of the mean. Readout error mitigation has already been applied in these plots.

CHAPTER 3

DISSIPATIVE STATE PREPARATION

Chapter 2 described the preparation of ground states of the 1D mixed-field Ising model using VQE. Ref. [115] discusses the preparation of similar states using a very different method, dissipative cooling, which mimics how cooling happens in nature. The system is coupled to an environment, in this case ancilla qubits that are reset between cycles. The interaction is designed to remove quasiparticles from the system and hence cool it towards its ground state. Although the removal of quasiparticles is not perfect, the resulting state of the system resembles the ground state in several important ways, quantified by long-range correlations, mutual information, and quasiparticle content.

This chapter details my contribution to Ref. [115]. Although I am an author of that paper, I did not write the text, so it is not reprinted here. Readers of this thesis are encouraged to read Ref. [115] for the full context.

My contributions to this work were twofold. First, I contributed the code that the first author used to compute the ground state energy of the 1D transverse field Ising model (TFIM), used in Figures 2 and 3 of the paper. Second, I performed numerical simulations of VQE, using code that I had written for the project discussed in Chapter 2 as a comparison. Further, although I did not write the manuscript, I did provide detailed comments on a draft of it, which were integrated into the final version.

3.1 Analytical solution of the transverse field Ising model

This section shows how to compute the ground state energy of the transverse field Ising model. A similar calculation was first performed by Ref. [133]. The Hamiltonian that we would like to solve is

$$H = -J\left(\sum_i Z_i Z_{i+1} + g\sum_i X_i\right). \tag{3.1}$$

We will solve this by doing a Jordan-Wigner transformation. Breaking convention slightly, define the Majorana fermions:

$$\begin{aligned}\chi_j &= Z_i \prod_{j<i} X_j \\ \tilde{\chi}_j &= Y_i \prod_{j<i} X_j,\end{aligned} \tag{3.2}$$

which satisfy $\{\chi_i, \chi_j\} = 2\delta_{ij}$, $\{\tilde{\chi}_i, \chi_j\} = 2\delta_{ij}$, and $\{\tilde{\chi}_i, \tilde{\chi}_j\} = 2\delta_{ij}$, where the curly brackets are anti-commutators. We can write the Hamiltonain in terms of these as

$$H = -iJ\left(\sum_i \tilde{\chi}_i \chi_{i+1} - g\sum_i \tilde{\chi}_i \chi_i\right). \tag{3.3}$$

Next, we want to find a new set of Majorana fermions in terms of which the Hamiltonian becomes that of free fermions. Note that an orthogonal transformation of Majorana fermions preserves the anticommutation relations. That is, if we define new Majorana fermions $\psi_i = O_{ij}\chi_j$ (with implicit summation over repeated indices), then

$$\begin{aligned}\{\psi_i, \psi_j\} &= O_{ik}O_{jl}\{\chi_k, \chi_l\} \\ &= 2(OO^T)_{ij} \\ &= 2\delta_{ij},\end{aligned} \tag{3.4}$$

so the anti-commutation relations are preserved under orthogonal transformations.

Now, we can write the Hamiltonian as

$$H = -i\tilde{\chi}_i \mathcal{H}_{ij} \chi_j, \tag{3.5}$$

where

$$\mathcal{H}_{ij} = J\delta_{j,i+1} - Jg\delta_{ij}. \tag{3.6}$$

We want to find a new set of Majorana fermions ψ_i and $\tilde{\psi}_i$ that diagonalize H, i.e. eliminate the hopping term. Let's define

$$\begin{aligned} \chi_i &= O_{ij}\psi_j \\ \tilde{\chi}_i &= \tilde{O}_{ij}\tilde{\psi}_j. \end{aligned} \tag{3.7}$$

(This is different from our previous O and ψ.) In terms of these, the Hamiltonian becomes

$$\begin{aligned} H &= -i\tilde{O}_{ik}\tilde{\psi}_k \mathcal{H}_{ij} O_{jl}\psi_l \\ &= -i\tilde{\psi}_i (\tilde{O}^T \mathcal{H} O)_{ij} \psi_j. \end{aligned} \tag{3.8}$$

Next, suppose that we have found the singular value decomposition of $\mathcal{H}$ and can write it as

$$\mathcal{H} = \tilde{O}\Sigma O^T, \tag{3.9}$$

where Σ is diagonal. (We will denote the diagonal elements of Σ as λ_i.) Then we can write

$$H = -i\sum_i \lambda_i \tilde{\psi}_i \psi_i. \tag{3.10}$$

To relate this back to more familiar free fermions, define the fermionic creation and annhilation operators:

$$\begin{aligned} a_i &= \frac{1}{2}\left(\psi_i + i\tilde{\psi}_i\right) \\ a_i^\dagger &= \frac{1}{2}\left(\psi_i - i\tilde{\psi}_i\right), \end{aligned} \tag{3.11}$$

which satisfy the canonical anticommutation relation $\{a_i, a_i^\dagger\} = \delta_{ij}$. The fermionic number operator for species ψ_i is

$$a_i^\dagger a_i = \frac{1}{2}\left(1 - i\tilde{\psi}_i\psi_i\right), \tag{3.12}$$

so we can finally write the Hamiltonian as

$$H = -\sum_i \lambda_i \left(1 - 2a_i^\dagger a_i\right). \tag{3.13}$$

The ground state energy is therefore

$$E_0 = -\sum_i \lambda_i. \tag{3.14}$$

Now, suppose we want the expectation values of the individual ZZ and X terms. To find them, we will write these operators in the ψ basis. We have

$$\begin{aligned} Z_i Z_{i+1} &= i\tilde{\chi}_i\chi_{i+1} \\ &= i\tilde{O}_{ij}O_{i+1,k}\tilde{\psi}_j\psi_k \\ &= \tilde{O}_{ij}O_{i+1,k}(a_j - a_j^\dagger)(a_k + a_k^\dagger) \\ \langle 0|Z_i Z_{i+1}|0\rangle &= \sum_j \tilde{O}_{ij}O_{i+1,j} \\ &= (\tilde{O}O^T)_{i,i+1}. \end{aligned} \tag{3.15}$$

Similarly,

$$\begin{aligned} X_i &= -i\tilde{\chi}_i\chi_i \\ \langle 0|X_i|0\rangle &= -(\tilde{O}O^T)_{ii}. \end{aligned} \tag{3.16}$$

These results allow the ground state energy, as well as the individual ZZ and X terms to be computed efficiently. I wrote the following Python code to implement this solution (the sparse arrays are just used for convenient initialization; they are converted to numpy arrays):

```python
import numpy as np
from scipy.linalg import svd
from scipy.sparse import coo_array


def fermion_H(n, g, J=1):
    i = np.append(np.arange(n - 1), np.arange(n))
    j = np.append(np.arange(1, n), np.arange(n))
    data = np.append(J * np.ones(n - 1), -J * g * np.ones(n))
    H = coo_array((data, (i, j)))
    return H


def ising_ground_state_energy(n, g, J=1):
    # returns the ground state energy of the transverse field Ising model
```

```
    H = fermion_H(n, g, J).toarray()
    l = svd(H, compute_uv=False)
    return -sum(l)


def ising_ground_state(n, g, J=1):
    H = fermion_H(n, g, J).toarray()
    Otilde, l, OT = svd(H)
    E = -sum(l)
    M = Otilde @ OT
    ZZ = np.array([M[i, i + 1] for i in range(n - 1)])
    X = -np.diag(M)
    return E, ZZ, X
```

3.2 Comparison of VQE and cooling

Having described both VQE and cooling, it is informative to see how they compare when applied to the same problem. To do this, I used the same ansatz as in Chapter 2, which is illustrated in Figure A.2, and optimized it classically using the gradient-based GPU optimizer that I wrote for Chapter 2.

First, consider the 18-qubit antiferromagnetic chain with $g/J = 1$, for which results are shown in Figure 3 of Ref. [115], including the lower panel of Fig. 3.D. In particular, in Eq. 3.1, we pick $J = -1$, $g = -1$ (corresponding to rotating $Z \to X$, $X \to -Z$ from Eq. 1 of Ref. [115]). The exact ground state energy is $E_0 = -22.562$, and the gap between the ground state and first excited state is $E_1 - E_0 = 0.170$. For VQE, we minimize the standard cost function,

$$C(\vec{\theta}) = \left\langle \psi(\vec{\theta}) \middle| H \middle| \psi(\vec{\theta}) \right\rangle . \tag{3.17}$$

The results are shown in Figure 3.1. The text in each of the two panels shows the energy that was achieved, the fidelity with the true ground state, the expectation value of the parity, $\prod_i X_i$, and the maximum single-site $|\langle Z_i \rangle|$. Plotted is the same long-range correlation function shown in Figure 3.D of Ref. [115]. In particular,

$$Y_j P_{j+1,k-1} Y_k = Y_j \left(\prod_{i=j+1}^{k-1} X_i \right) Y_k \tag{3.18}$$

(since we are using $Z \to X$, $X \to -Z$ compared to Ref. [115]).

It appears that at least 19 layers are needed in order to generate the long-range correlations (here “layer” means the same thing as in Chapter 2 and Figure A.2). Nevertheless, in both cases, the energy achieved was very close to the ground state energy, well below the energy of the first excited state. This illustrates that E/E_0 can be made close to unity while other quantities, such as long-range correlations, and, in this case also the single-site $\langle Z_i \rangle$, differ significantly from what they are in the true ground state. The precise energy measurements needed to distinguish $E = -22.52$ from $E = -22.56$ could prove difficult on real devices, where both many shots and high fidelities are required for this level of precision.

Next, consider the same chain with $g/J = 0.6$, for which results are shown in the

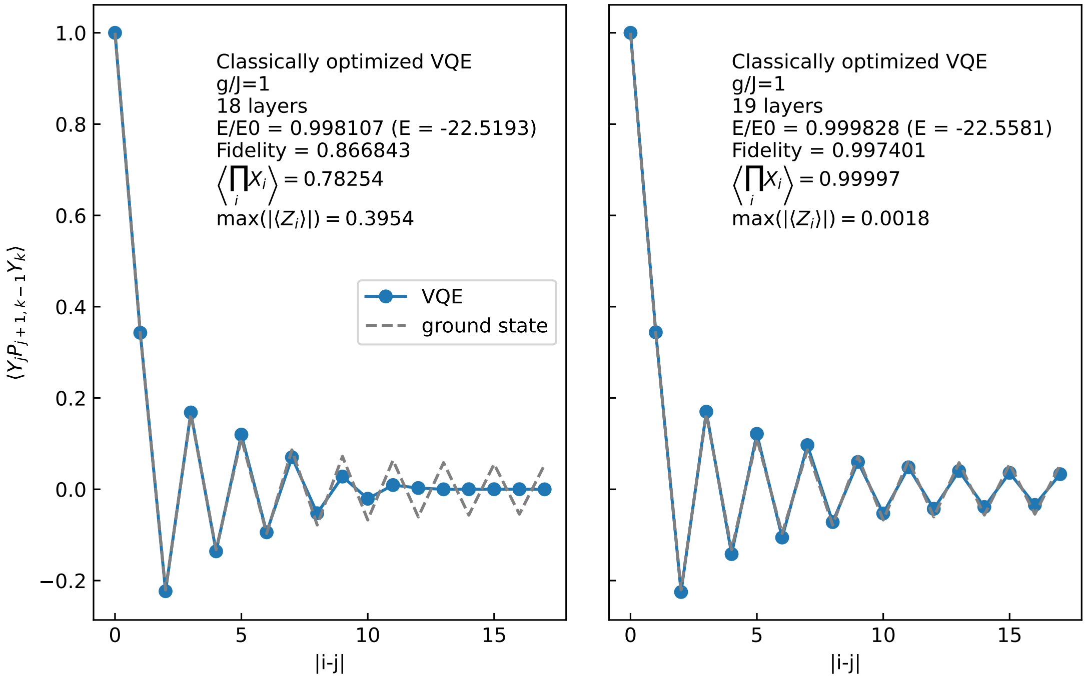


Figure 3.1: Classically-simulated VQE applied to the antiferromagnetic 1D TFIM with $g/J = 1$. The long-range correlation function from Fig. 3D of Ref. [115] is shown.

upper panel of Fig. 3.D in Ref. [115]. The ground state is nearly degenerate. The lowest two energy eigenstates are approximately $\frac{1}{\sqrt{2}}(|0101\ldots\rangle \pm |1010\ldots\rangle)$. The ground state energy is extensive; $E_0 \sim n$ (specifically $E_0 = -18.8498$), whereas the gap between the lowest two energy levels is approximately g^n (specifically 0.0001). The small gap makes it nearly impossible for VQE to identify the ground state from within the nearly degenerate subspace without any extra input.

In particular, if I simply perform VQE with Eq. 3.17, I find that it tends to converge to states close to either $|0101\ldots\rangle$ or $|1010\ldots\rangle$. $\langle Z_i \rangle$ exhibits an antiferromagnetic pattern, whereas in the true eigenstates, $\langle Z_i \rangle = 0$. These states lack the long-range correlations exhibited in Fig. 3.D of Ref. [115].

However, this is selling VQE a bit short, since it is straightforward to modify the cost function to enforce that $\langle Z_i \rangle \approx 0$, a property that we know by symmetry the ground state should have:

$$C(\vec{\theta}) = \left\langle \psi(\vec{\theta}) \right| H \left| \psi(\vec{\theta}) \right\rangle + \lambda \left(\sum_i \left(\left\langle \psi(\vec{\theta}) \right| Z_i \left| \psi(\vec{\theta}) \right\rangle \right)^2 \right), \tag{3.19}$$

where $\lambda > 0$. Using this cost function, I find that, indeed, VQE is able to converge to an eigenstate. However, it is still unable to distinguish between the ground state and the first excited state. This is illustrated in Figure 3.2. Two independent optimization routines are performed, one with the 18-layer ansatz and one with the 19-layer ansatz. The classically-performed optimization routine is allowed to run for several hours and then is stopped, perhaps before it has fully converged. It is clear that the 18-layer case shown on the left converged to the first excited state, whereas the 19-layer case shown on the right coverged to the ground state, even though the 18-layer case actually had a lower energy. This again highlights the importance of not relying solely on energy as the criterion of convergence, especially when the gap is small.

Suppose that we did not know which of these two states was the ground state. As we have seen, measuring the energy gives the wrong answer in this case. We might additionally measure the parity $\langle \prod_i X_i \rangle$, which could tell us the answer if we know what the parity of the ground state is supposed to be, i.e. in Figure 3.2, we can select the state with positive parity. However, parity measurements are difficult

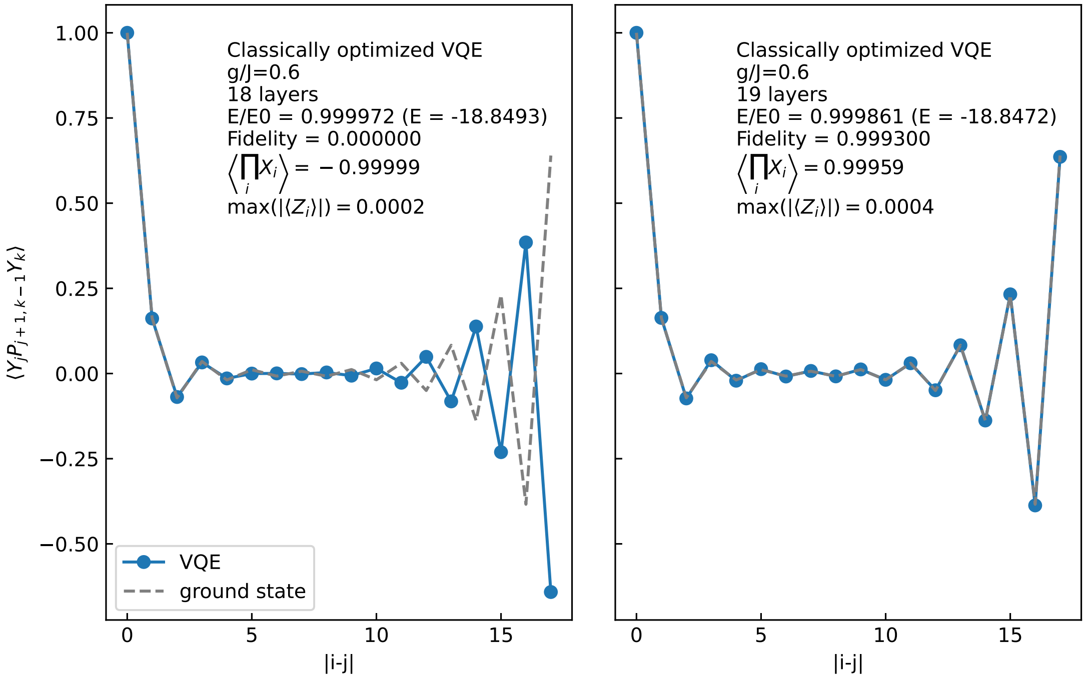


Figure 3.2: Classically-simulated VQE applied to the antiferromagnetic 1D TFIM with $g/J = 0.6$ and $\langle Z_i \rangle \approx 0$ enforced with a penalty term. The long-range correlation function from Fig. 3D of Ref. [115] is shown.

on many qubits because if there is an error on any one of the qubits, that flips the sign of the parity. From Eq. A.10, we see that, if $e_0 + e_1 \approx 2\%$, then we expect the expectation value of the parity measurement to be suppressed by a factor of about 0.98^n, which for 18 qubits is 0.7. This damping factor is not so close to 0 that we cannot tell the sign of the parity. However, as we scale up the number of qubits, we require more shots to be able to tell the sign of the parity. Typical readout error rates on a Sycamore processor are shown in Figure C.3 for the experiment described in Chapter 5.

It is striking that, even though dissipative cooling performs much worse than VQE in terms of the energy, it is able to pick out the ground state from within the nearly-degenerate subspace and reproduce produce the long-range correlation function, which is difficult for VQE if there is a small gap and there is no easy way to add penalty terms to pick out the eigenbasis. For the 1D TFIM, adding penalty terms to pick out the eigenbasis from within this subspace is easy, but that may not be true for all gapless systems. Further, as we have discussed, distinguishing the ground state from the first excited state, even if VQE is picking out eigenstates as in Figure 3.2, may be difficult. Finally, the dissipative cooling algorithm naturally produces eigenstates of the Floquet unitary rather than of the Hamiltonian, which may be desired in some cases. Generating Floquet eigenstates using VQE is not straightforward.

CHAPTER 4

CHARACTERIZING CORRELATED READOUT ERRORS

4.1 Introduction

Quantum computers promise to transform computational chemistry and physics by enabling more accurate simulations and by rendering previously intractable computations within reach. Classical computers are severely limited in their ability to simulate many-body quantum mechanical systems with long-range entanglement, and analytical methods based on perturbation theory generically fail at strong coupling. Quantum computers, however, promise to overcome these challenges by encoding the quantum state of the system directly in the state of the qubits, the two level systems used as the basis for digital quantum computation.

The main reason why quantum computers have not yet succeeded in providing an advantage over classical techniques for studying these systems is that the circuit depth is severely limited by noise, particularly by high two-qubit error rates and decoherence. In the long-term, the goal is to use error correction to achieve fault-tolerant quantum computing. In the near term, however, a variety of error-mitigation methods have been proposed that aim to undo the effects of the noise on expectation values of interest [142].

Another challenge has also hindered quantum advantage on superconducting transmon-based quantum computers: state preparation and measurement (SPAM) errors. SPAM error rates on these devices are high, about 1–10%, depending on the device and the qubit, which makes the simultaneous readout of many qubits unreliable. This effect is particularly relevant for measuring fermionic operators, which act on many qubits simultaneously after being written as Pauli operators using a Jordan-Wigner or Bravyi-Kitaev transformation. Various creative techniques have been proposed to mitigate correlated measurement (also known as readout) errors (e.g. [126, 154, 67, 132, 110, 11, 72]).

A more common approach to readout error mitigation consists of measuring the measurement fidelity matrix (MFM, also known as the response matrix), $R_{ij} = P(\text{measure } j|\text{prepare } i)$, and applying its inverse to the measured counts. This is feasible if one assumes that readout errors are uncorrelated (e.g. [16]) but is exponentially difficult in general. It is also feasible if one allows for readout errors to have a local or quasi-local correlation structure [68].

In this work, we benchmark correlated readout errors on current IBM quantum computers and test the extent to which it is necessary to include these correlations in error mitigation. In particular, we measure sparse Sachdev-Ye-Kitaev (SYK) Hamiltonians [176] in low-energy product states. Although product states are not good approximations to the ground state, they can be constructed without any two-qubit gates, so the error is dominated by SPAM errors. Further, terms in the SYK Hamiltonian act on many qubits, so this observable is particularly sensitive to SPAM errors. Even so, we find that uncorrelated readout error mitigation is able to reproduce the correct energy to within statistical uncertainty in almost all cases.

4.2 Test problem: Sparse SYK

The Sachdev-Ye-Kitaev model was introduced by Alexei Kitaev in 2015 as a simple example of the holographic principle; in the large-N and low temperature limit, it is dual to Jackiw-Teitelboim gravity [95, 93, 94, 111, 112, 113]. It is described by the ensemble average over Hamiltonians of the form

$$H = i^{\frac{q}{2}} \sum_{i_0 < \ldots < i_{q-1}} j_{i_0 \ldots i_{q-1}} \chi_{i_0} \cdots \chi_{i_{q-1}}, \tag{4.1}$$

where the χ_i are Majorana fermions: $\{\chi_i, \chi_j\} = \delta_{ij}$, and the $j_{i_0 \ldots i_{q-1}}$ are Gaussian-random numbers with mean 0 and variance $\frac{J^2(q-1)!}{N^{q-1}}$. q is an even number that must be at least four for the model to be dual to JT gravity, and N is the number of fermions, which, as we will see, is twice the number of qubits needed to represent the Hilbert space. J is a constant that sets all of the physical scales. For this work, we fix $q = 4$. Note that the interactions are all-to-all

We can write this as a qubit Hamiltonian using a Jordan-Wigner or Bravyi-Kitaev transformation [17]. The Jordan-Wigner transformation allows one to represent $2n$ Majorana fermions using n qubits: $\chi_i = \frac{1}{\sqrt{2}} X_i Z_{i-1} \ldots Z_0$ and $\tilde{\chi}_i = \frac{1}{\sqrt{2}} Y_i Z_{i-1} \ldots Z_0$. The reader can verify that this reproduces the correct anticommutation relations. When written in terms of Pauli operators using the Jordan-Wigner transformation, the reader can see that the terms in the highly-nonlocal SYK Hamiltonian (4.1) act on extensively many qubits; there are Pauli-Z strings between the non-adjacent fermions. The Bravyi-Kitaev transformation also produces large Pauli operators, although not quite as large as those produced by Jordan-Wigner, as illustrated in

Figure B.1.

There is significant interest in implementing SYK Hamiltonians on a quantum computer (e.g. [114, 7, 108, 157]). Doing so could let us study aspects of the SYK model and quantum gravity that are not analytically tractable. It is also an important component of a recent experiment that claims to create a state dual to a traversable wormhole [18, 129, 151, 78].

One challenge in studying SYK this way is that the number of Hamiltonian terms is very large: $\begin{pmatrix} N \\ q \end{pmatrix}$. However, Ref. [176] argued that it is only necessary to keep $O(N)$ terms from this Hamiltonian in order to recover the interesting physics of SYK (see also [25]). If the number of terms kept is N_{terms}, then the new variance of the couplings should be chosen as

$$\frac{J^2(q-1)!}{N^{q-1}} \frac{\begin{pmatrix} N \\ q \end{pmatrix}}{N_{\text{terms}}} \approx \frac{J^2 N}{q N_{\text{terms}}}. \tag{4.2}$$

The interesting physics of the SYK model, including its duality with JT gravity, emerges at low temperatures. To study this regime on a quantum computer, it is necessary to prepare a low-temperature thermal state. The most feasible near-term proposals for doing this are variational techniques in which the energy must be measured at each iteration (e.g. [157, 114]). Low-temperature SYK thermal states have not yet been achieved for system sizes large enough to exhibit the characteristic fast-scrambling dynamics that led Kitaev to consider SYK as a model of holography.

Many challenges prevent the production of these thermal states on current quantum computers, including, prominently, the practical limitations on circuit depth. However, because the Hamiltonian terms act on so many qubits, readout errors are also quite significant. In this chapter, we consider the easier of these challenges, namely that of mitigating the readout errors.

4.3 Characterizing and mitigating readout errors

Existing quantum computers have state preparation and measurement (SPAM) errors, the latter of which are also known as readout errors. These errors impede the measurement of operators that act on many qubits simultaneously, such as the SYK Hamiltonian (4.1). Further, these readout errors may be correlated, making them more difficult to mitigate. Here, we study the readout errors on several IBM quantum processors, characterizing their correlations, and studying readout error mitigation applied to measurements of sparse SYK Hamiltonians.

4.3.1 Uncorrelated readout errors

Before including correlations, we consider a simpler model of readout errors. In particular, we assume that the response matrix for the n-qubit system is a tensor product of n 2×2 response matrices:

$$R = \otimes_q R_q, \qquad R_q = \begin{pmatrix} 1 - e_0^{(q)} & e_1^{(q)} \\ e_0^{(q)} & 1 - e_1^{(q)} \end{pmatrix}. \tag{4.3}$$

The single-qubit error rates, e_0 and e_1, are measured by preparing the qubits in random bitstring states and extracting the probability of bit-flip errors. We further condition on the qubit having been prepared in the $|0\rangle$ state in the preceeding bitstring, since we have found that state preparation errors can be higher when the previous state of the qubit was $|1\rangle$. This procedure assumes the error rates associated with readout are much higher than those associated with state preparation and single-qubit X gates, which is generally the case. The inverse of the response matrix can then be efficiently constructed and applied to measured bitstrings.

To test the efficacy of this method, we prepare low-energy product states and measure the expectation value of the SYK Hamiltonian, Eq. 4.1. More specifically, we consider sparse SYK (Eq. 4.2), with N_{terms} random terms and coefficients. We convert this to a spin Hamiltonian using either a Jordan-Wigner or Bravyi-Kitaev transformation. We then use a classical optimizer to find the product state with the lowest energy. The use of product states allows us to run circuits without two-qubit gates, thus isolating the effect of readout errors. Further, we divide the Hamiltonian terms into sets that can be measured simultaneously (see Appendix B). Next we run two sets of circuits: (1) the random bitstring circuits to characterize the readout error rates and (2) the circuits to prepare the low-energy product state and measure its energy, both in the same job. Running them in the same job means that they will be executed in the order: circuit 0, shot 0, circuit 1 shot 0, . . . , which allows us to exclude any effects from error rates drifting over time. Finally, we process the measured bitstrings, extracting the unmitigated and mitigated SYK energy.

Results from various 27-qubit quantum processors are shown in Figure 4.1. Here we consider sparse SYK with 54 Majorana fermions and either 25 or 100 terms. We find that uncorrelated readout error mitigation is able to reproduce the correct energy to within statistical uncertainty. Thus, correlated readout error mitigation is not necessary in this case.

Next, we repeat the experiment on a 65-qubit quantum processor, *ibmq_brooklyn*. We consider sparse SYK with 130 Majorana fermions (65 qubits) and 25 terms. As shown in Figure 4.2, we find that, somewhat remarkably, uncorrelated readout error mitigation is able to transform raw energy measurements, sometimes even with the wrong sign, to mitigated values that are within 1σ of the true energy. Results are shown for two sparse-SYK instances, one of which is qubitized via a Bravyi-Kitaev transformation and the other of which is qubitized using a Jordan-Wigner transformation. Readout errors have a more significant effect in the Jordan-Wigner case, as expected since the Pauli operators act on more qubits. This also leads to smaller error bars in the Bravyi-Kitaev case, and there we see what may be a small systematic deviation from the ideal energy. The VQE optimization routine requires precise energy measurements in order to be able to identify the gradient. Therefore, small systematic errors could potentially lead to poor performance of the algorithm. Therefore, we now focus on characterizing and attempting to invert, correlated readout errors on this 65-qubit device.

4.3.2 Characterizing correlated readout errors

First, we make a simplifying assumption about the structure of the correlated errors. We assume that

$$P(b'_{n-1}b'_{n-2}\dots b'_0|b_{n-1}b_{n-2}\dots b_0) = \prod_i P_i(b'_i|b_{n-1}b_{n-2}\dots b_0), \tag{4.4}$$

where $P(b'_{n-1}b'_{n-2}\dots b'_0|b_{n-1}b_{n-2}\dots b_0)$ is the probability of observing bitstring $b'_{n-1}b'_{n-2}\dots b'_0$ given that the qubits were prepared in the computational basis state corresponding to the bitstring $b_{n-1}b_{n-2}\dots b_0$. In the expression on the right-hand side, $P_i(b'_i|b_{n-1}b_{n-2}\dots b_0)$ is the probability that qubit i is measured to be b'_i given that the qubits are prepared in $b_{n-1}b_{n-2}\dots b_0$ (marginalizing over the other qubits). We have found this to be a good model of errors on the devices that we studied.

We identify correlated readout errors as follows:

1. Data collection:
    (a) Randomly choose $n_b \leq 2^n$ unique n-bit bitstrings. Let's call these b_{ij}, where $i \in \{0, \dots, n_b\}$ indicates which bitstring it is and $j \in \{0, \dots, n\}$ labels the qubit.
    (b) Generate a length n_b Python list of Qiskit quantum circuits that prepare n qubits in the computational basis states corresponding to the bitstrings generated in (a).
    (c) Run these circuits in a single job on the desired quantum processor, using s shots, the maximum delay time between circuits, and disabling excited

state promotion (ESP) readout [125, 124]. If this data is going to be used for error mitigation, then ideally it should be run in the same job as the circuits that the mitigation is going to be applied to.

2. Data analysis:
   (a) For each qubit q and each given state of that qubit ($g \in \{0, 1\}$), do the following:
      i. For each of the prepared bitstrings such that $b_{i,q} = g$ and $b_{i-1,q} = 0$, compute the number of times (out of s) that qubit q was measured to be $1 - g$, marginalizing over the measured states of the other qubits. We find that conditioning on the qubit having been prepared in 0 in the previous state is necessary to eliminate the effects of state preparation errors, which otherwise can lead to a higher error rate when the previous state of qubit q was 1. This generates a list of approximately $n_b/4$ numbers that, when considered as fractions out of s, each approximate the error rate on qubit q, given that it is prepared in the state g. If no correlations are present, then these numbers are sampled from a binomial distribution with s weighted coin-flips and probability p.
      ii. Perform a hypothesis test to check whether list of numbers found in i. is consistent with a binomial distribution. In practice, we check whether the variance is significantly greater than that predicted by a binomial distribution: $p(1 - p)s$, using the mean of the list found

in i., divided by s to estimate p. If no significant deviation from a binomial distribution is found, then we don't identify any correlations here and continue the for loop (a). Otherwise:

A. Iterate over the number of control qubits, n_c, starting with 1 and going up to $n_{c,\max}$, the user-defined maximum number of control qubits. Within this loop, iterate over the possible choices of control qubit(s), of which there are $\begin{pmatrix} n-1 \\ n_c \end{pmatrix}$. For each choice of control qubits, split the numbers (sample error rates) found in i. according to the state of the control qubit(s). Test whether all of the sub-vectors are consistent with binomial distributions. If yes, accept these control qubits. Otherwise, continue.

The actual procedure is slightly more complicated in that we also iterate over the number of distinct error rates among the 2^{n_c} states of the control qubits, favoring fewer distinct error rates.[1] Further, we don't stop looking for control qubits once we have found candidate control qubits; we continue to see if there are any other candidates with the same number of control qubits and number of distinct error rates. The p-values from the hypothesis tests are all scaled up appropriately to account for the fact that we are doing many hypothesis tests. We consider the probability that a binomial distribution will ever produce a sample variance as large as that observed,

---

[1] For example, if there are two control qubits, there are a maximum of 4 distinct error rates on the affected qubit, one for each joint state of the two control qubits; however there could be fewer distinct error rates. For example, it could be that the error rate on the affected qubit is larger when the control qubits are in the state 11, while the other 3 states produce the same error rate on the affected qubit.

in any of n_t trials, where n_t is the number of ways of partitioning the vector from (2.a.i) that we are iterating over.

4.3.3 Mitigating readout errors

The characterization procedure described in Section 4.3.2 can also be used for error mitigation. In particular, correlation maps like that shown in Figure 4.6 can be used to identify the tensor product structure of the $2^n \times 2^n$ response matrix. In particular, the correlation map can be decomposed into sets of qubits, which we call islands, such that there are no correlations between the islands. This is illustrated in Figure 4.7. The response matrix is then a tensor product of the response matrices in the islands. For the case shown in Figure 4.7, the largest island has size 17, so the problem of inverting a $2^{65} \times 2^{65}$ matrix is thus reduced to that of inverting a $2^{34} \times 2^{34}$ matrix. While this is a significant reduction in complexity, unfortunately, it is still infeasible.

However, this could be the starting point for identifying the correlations that most strongly affect the readout error rates. Including only the strongest correlations in the response matrix could lead to smaller correlation islands and make the problem tractable. We leave this for future work.

4.4 Conclusions

There are a number of challenges that make SYK a particularly difficult system to study using programmable quantum processors. If one wants to perform VQE on a system that is large enough for quantum advantage, correlated readout errors may pose a challenge. This challenge may be overcome either with readout error mitigation strategies that attempt to include the most important correlations, as outlined in the previous section, or with improvements in device performance so that the correlations are a negligible effect, as is already the case on the 27-qubit devices studied in this chapter. However, other challenges may prove less surmountable. In particular, we found in the Appendix (Fig. B.2) that the circuit depths required to approximate the ground state may be much deeper than those needed for the 1D spin chains considered in Chapters 2 and 3, which is not surprising if one motivates VQE by considering adiabatic state preparation and notes the complexity of the SYK Hamiltonian. Circuit depth is severely limited by both two-qubit gate errors and coherence times. While it is possible that these challenges may be addressed with a clever sparsification of the Hamiltonian, fermion-to-qubit mapping, or variational ansatz, we do not see SYK as a likely candidate for quantum advantage in the near term.

Contributions

This work was performed in collaboration with Raphael Pooser and Titus Morris of Oak Ridge National Laboratory. We thank Haggai Landa and Paul Ginsparg for helpful discussions. ER gratefully acknowledges financial support from US DOE grant DE-SC0020360 and access to IBM quantum-computer time via the *AFRL IBM Q Network Hub*, facilitated by Laura Wessing and Jon Maggiolino. ER also thanks Peter McMahon for access to computing resources. The views expressed are those of the authors, and do not reflect the official policy or position of IBM or the IBM Quantum team. We used Qiskit [161] to interact with IBM's devices. The introduction to this chapter contains text that ER wrote for a DOE SCGSR proposal.

Code availability

The code for this thesis chapter is available at https://github.com/eliottrosenberg/correlated_SPAM. Circuits to measure readout errors can be run with `readout_circuits_rand_execute`, and the readout correlation map can be constructed with `readout_partition(job)`, where `job` is the IBMQ job with the random circuits (both in `readout_characterization.py`). This returns a `readout_partition` object that contains the correlation map. The parameter `max_n_control` sets the maximum number of control qubit when constructing the map. The method `measure_H` allows one to measure a Hamiltonian (or other operator written in terms of Pauli operators) and apply readout error mitigation, as long as the islands are small enough. For uncorrelated readout error mitigation, simply set `max_n_control = 1`.

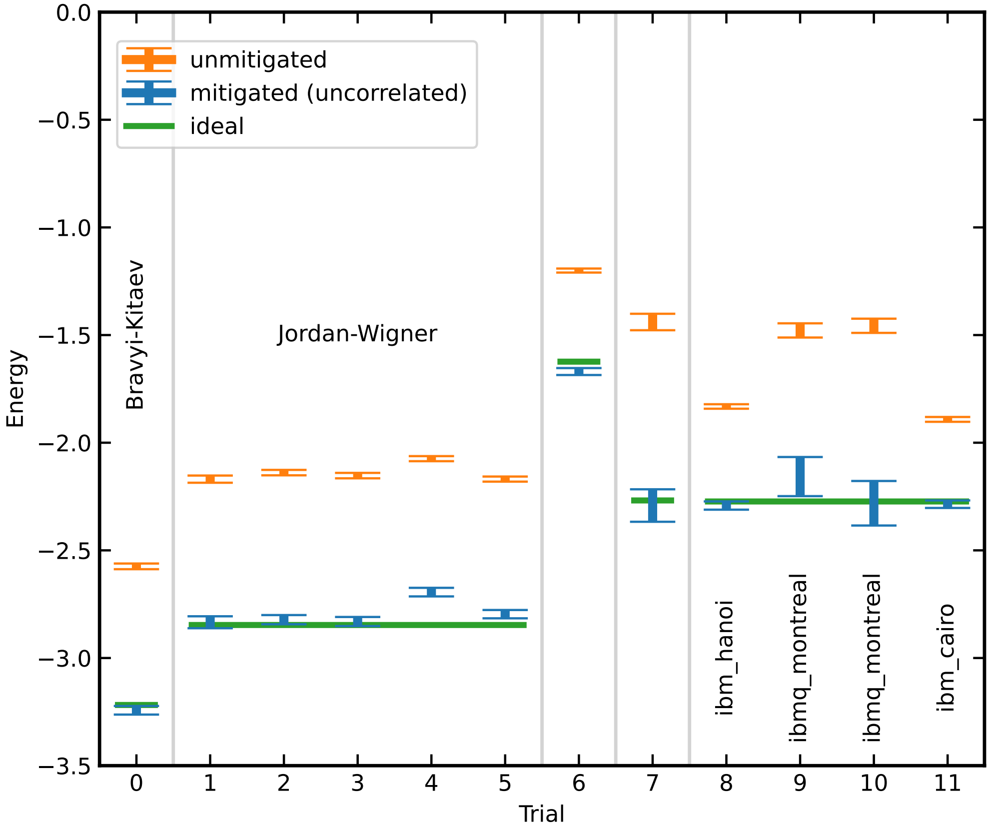


Figure 4.1: Uncorrelated readout error mitigation for sparse SYK with 54 Majorana fermions (27 qubits). Vertical lines separate different random instances of sparse SYK. The instance shown in trial 0 is converted to a spin Hamiltonian via a Bravyi-Kitaev transformation, whereas the rest are converted to spin Hamiltonians via Jordan-Wigner transformations. The Hamiltonians used for instances 0–6 contain 25 terms, whereas the instances used in trials 7-11 contain 100 terms. Unless otherwise indicated, the experiments are performed on the 27-qubit device *ibm_hanoi*. As described in the text, the lowest-energy product state is identified classically (energy shown in green). It is prepared and its energy measured on the quantum processor (measured energy shown in orange). Uncorrelated readout error mitigation is applied to the measurement results, and the mitigated energy is shown in blue. All error bars are $\pm 1\sigma$.

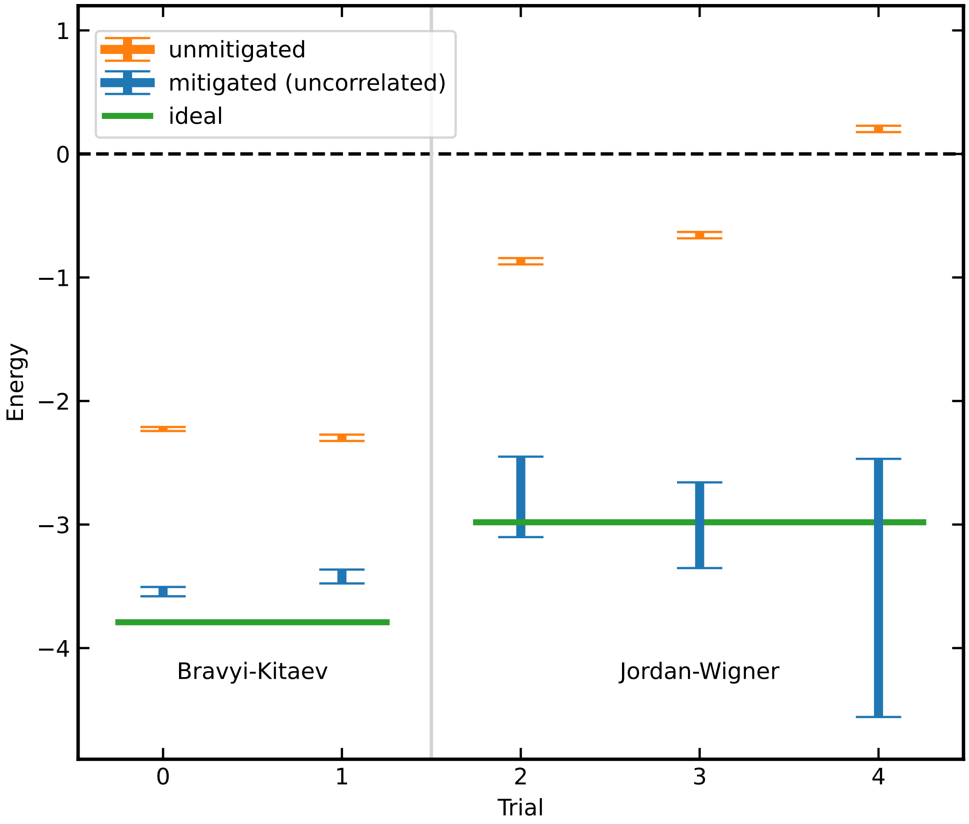


Figure 4.2: Same as Figure 4.1 except for sparse SYK with 130 Majorana fermions (65 qubits), run on the 65-qubit device *ibmq_brooklyn*. Both sparse SYK instances have 25 terms. As indicated in the plot, one is transformed to a spin Hamiltonian using a Bravyi-Kitaev transformation and the other using a Jordan-Wigner transformation.

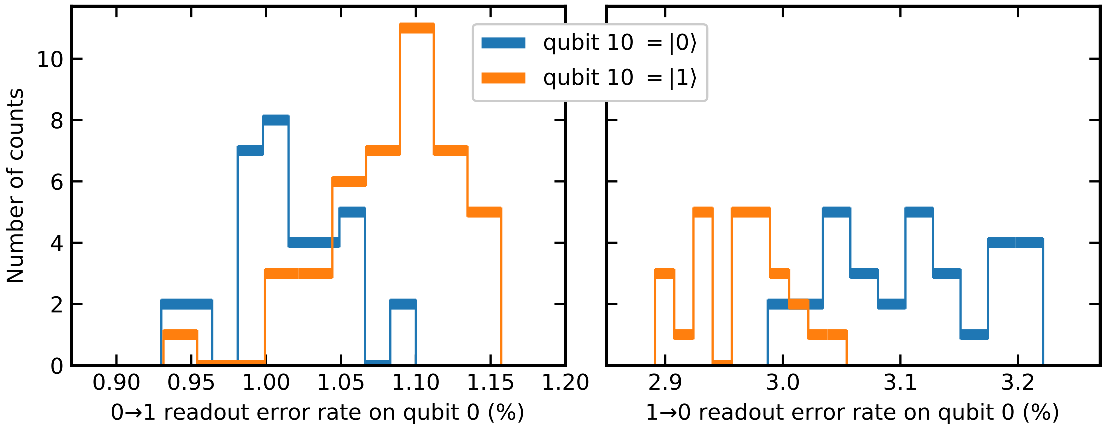


Figure 4.3: The readout error rate on qubit 0 depends on the state of qubit 10. This correlation was identified using the procedure described in the text.

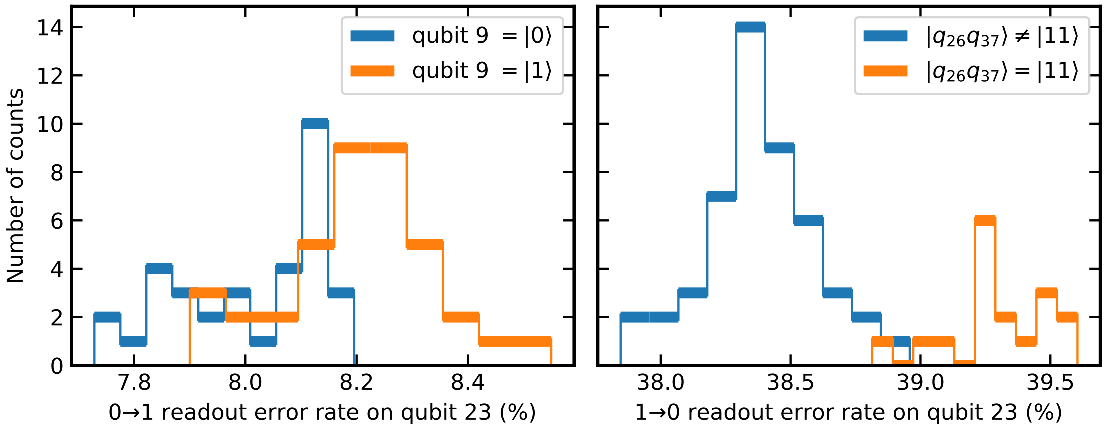


Figure 4.4: The readout error rate on qubit 23 depends on qubit 9 if it is prepared in the 0 state (left) but on both qubits 23 and 37 if it is prepared in the 1 state (right).

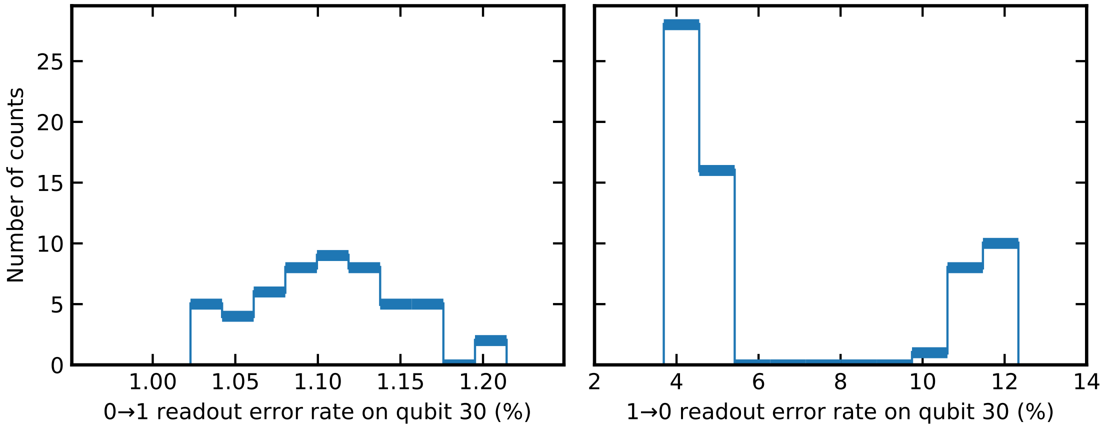


Figure 4.5: The readout error rate on qubit 30. When prepared in the 1 state (right), the readout error rate on least two distinct values depending on the prepared state of the other qubits. Out of the 63 prepared states of the other qubits, 19 result in an error rate above 9%, whereas the other 44 result in an error rate below 6%. However, this splitting is not accounted for by the state of any three other qubits on the device. This suggests a correlation with more than three control qubits.

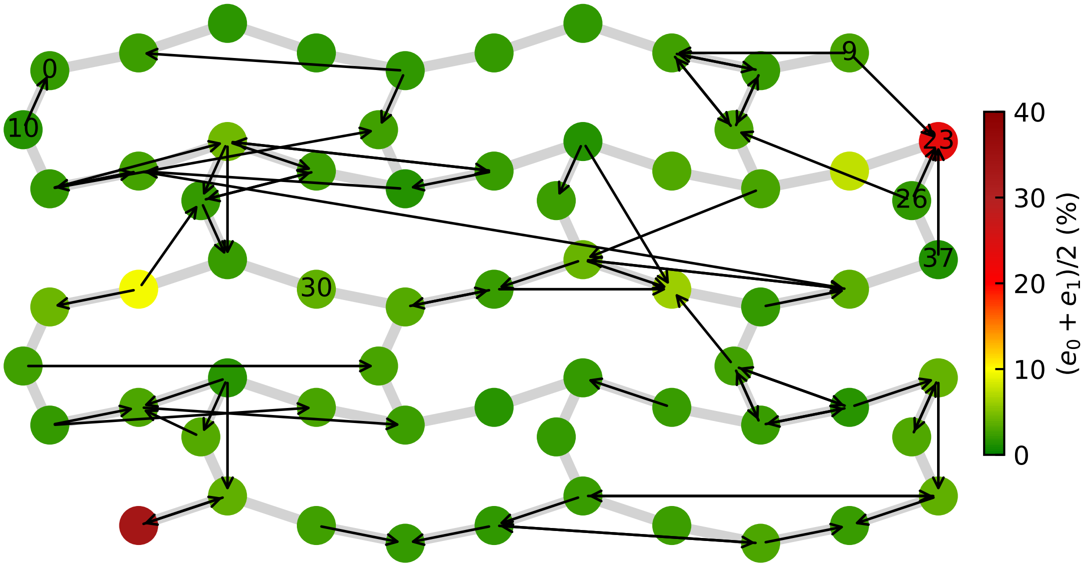


Figure 4.6: Correlation graph of readout errors on *ibmq_brooklyn*. The qubits highlighted in Figures 4.34.5 are labeled. Arrows indicate correlations in the readout error rates. Grey lines indicate where native two-qubit gates can be applied. The colors of the qubits indicate their average readout error rate.

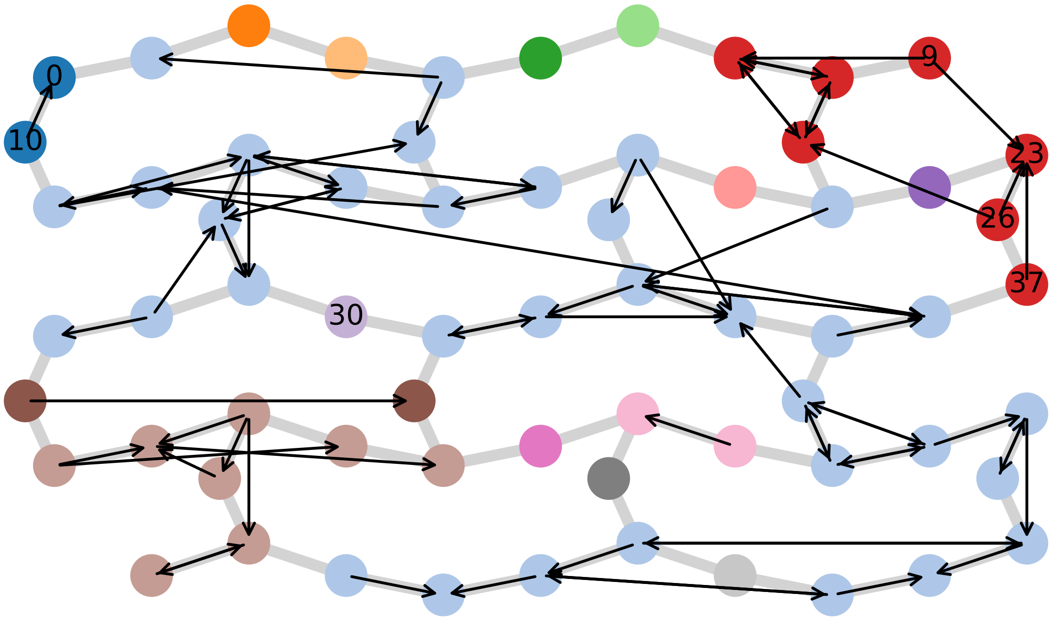


Figure 4.7: Same as Figure 4.6 but with qubits colored according to their island.

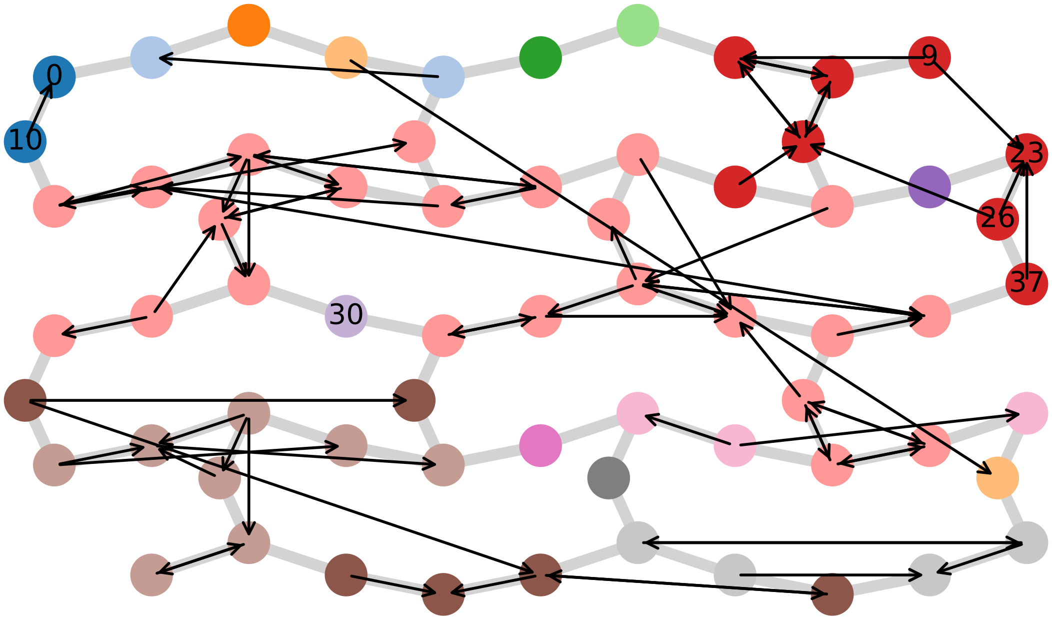


Figure 4.8: Same as Figure 4.7 but with when multiple options exist for possible control qubits, the maximum likelihood option (smallest variance of the error rate) is picked, whereas in Fig. 4.7, the most local option is picked.

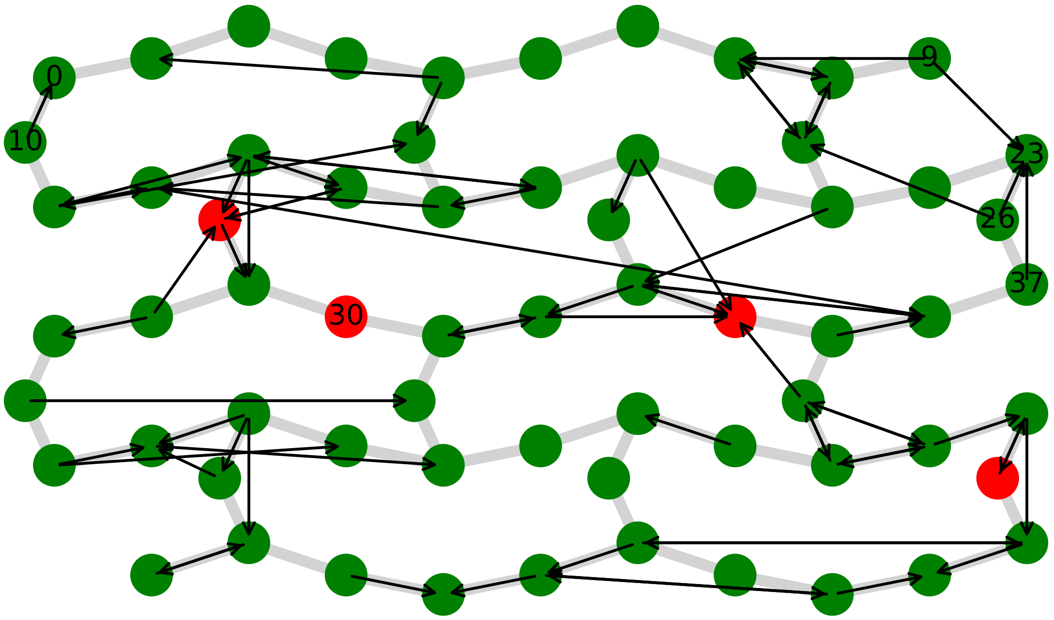


Figure 4.9: Same as Figure 4.7 but indicating in red which qubits exhibit multiple error rates that are not controlled by the states of at most three other qubits. An example of such a qubit is shown in Figure 4.5.

CHAPTER 5

THE HEISENBERG MODEL AND THE KPZ UNIVERSALITY CLASS

In statistical physics, the notion of universality is a powerful assertion; it implies that systems with entirely different microscopic interactions can share the same emergent macroscopic description due to having certain basic physical properties in common. It is a triumph of this assertion that, for instance, the Ising model prevails in our understanding of the zero-temperature phase transitions in a wide class of systems [83, 144]. The basic ingredients commonly affecting universality classes are the collective behavior of constituent elements, symmetries, conservation laws, and dimensionality, as described by the renormalization group (RG) theory [130]. In contrast to rather well-understood low-temperature universality classes, which are determined by ground-state physics, we have limited knowledge of the universality classification of dynamical phases of matter at finite temperatures, where contributions from the entire spectrum must be considered [135, 60].

It has been observed in several dynamical systems that the long-time behavior permits a few-parameter hydrodynamical description, suggesting the existence of universality [155, 52, 82, 184, 171, 56]. The emergence of a hydrodynamical description relies on reaching local, and subsequently, global equilibrium [109, 23]. This fate is less certain in systems with an extensive set of conserved quantities, i.e., integrable systems, which are known to evade thermalization, and their universal behaviors are discussed in the framework of generalized hydrodynamics [28, 13, 44, 24, 42].

Distinct microscopic models or dynamics belong to the same universality class if they share a single scale-invariant limit under a RG flow [83, 130]. A universality class is commonly characterized by scaling exponents and scaling functions, and it is rather implausible to extract them all experimentally. Therefore, experiments, e.g., on quantum processors, cannot prove that a set of observed dynamics belongs to a given class, but they can *falsify* a universality conjecture [136] by examining its predictions. They can also probe numerically and theoretically challenging regions of the parameter space, which has proven advantageous for studying universal behaviors [155, 52, 82, 184, 171, 56].

Superconducting quantum processors offer high wavefunction sampling rates, which enabled them to outperforming classical computers in sampling tasks [5, 173]. On these processors one can go beyond mean expectation values and provide "snapshots" of an observable, which allows for measuring quantum fluctuations and the probability distribution of the observable. The capability of collecting full counting statistics could have fundamental consequences for our understanding of dynamical universalities. In particular, it is commonly assumed that the scaling functions and exponents of the first few moments uniquely determine a universality class, and there have not yet been any instances where the higher moments of an observable have led to a different classification.

Spin dynamics in the one-dimensional (1D) XXZ model have been the subject of numerous recent studies [169, 13, 104, 156, 147, 77, 62, 41, 63, 43, 45, 46, 76, 150, 79, 171]. This integrable model describes nearest-neighbor exchange interactions

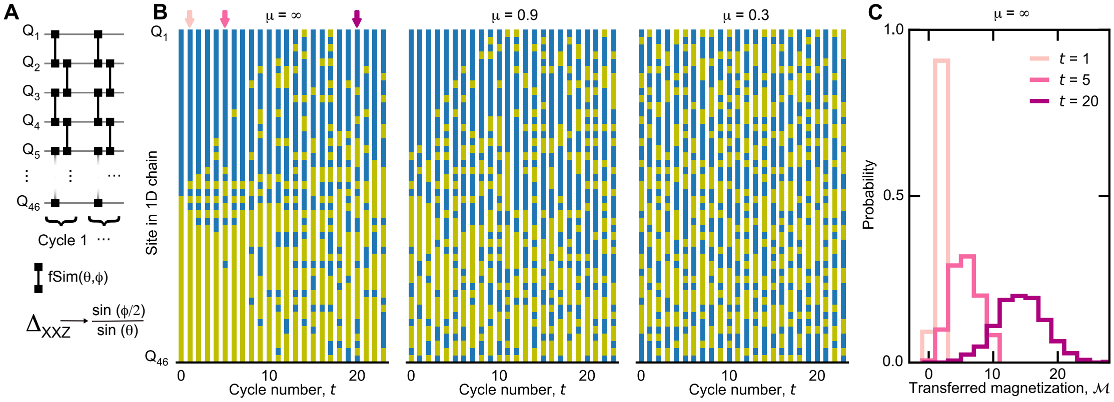


Figure 5.1: **Domain wall relaxation in the Heisenberg XXZ spin chain**. (**A**) Schematic of the unitary gate sequence used in this work, where fSim gates are applied in a Floquet scheme on a 1D chain of $N_Q = 46$ qubits. (**B**) Relaxation dynamics as a function of site and cycle number for $\mu = \infty, 0.9$, and 0.3 for initially prepared domain-wall states with $2\langle S^z\rangle = \pm\tanh\mu$. (**C**) Histogram showing the probability distribution of transferred magnetization after $t = 1$, 5 and 20 cycles (arrows in **B**) for $\mu = \infty$.

between spin-1/2 particles, with the Hamiltonian [71]

$$\hat{H} = \sum_i \left(S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + \Delta S_i^z S_{i+1}^z\right), \tag{5.1}$$

where S^x, S^y, and S^z are spin-1/2 operators, and Δ is the anisotropy parameter. When $\Delta = 1$, this system is the Heisenberg model, a paradigmatic model of quantum magnetism that possesses a global SU(2) rotational symmetry. The spin dynamics in the Heisenberg model exhibit characteristics consistent with the Kardar-Parisi-Zhang (KPZ) universality class, which was originally introduced to describe the stochastic, nonlinear dynamics of driven interfaces and has proven to apply to a wide range of classical systems [104, 156, 147, 77, 62, 41, 63, 43, 45, 46, 76, 150, 171]. The KPZ-like behavior of the spin dynamics is surprising due to the absence of stochasticity and nonlinearity in the Heisenberg model.

In a 1D chain of $N_Q = 46$ superconducting qubits, we simulate this spin model by periodic (Floquet) application of high-fidelity 2-qubit unitary fSim(θ, ϕ) gates (Fig. 5.1A, see SM and Ref. [127]). Here, θ sets the amplitude of hopping between adjacent qubit lattice sites, and ϕ is the conditional phase angle imparted when two spin excitations are adjacent to each other. Within each cycle, two-qubit fSim(θ, ϕ) gates are applied between all neighboring pairs in the chain, resulting in the cycle unitary:

$$\hat{U}_F = \prod_{\text{even bonds}} \text{fSim}(\theta, \phi) \prod_{\text{odd bonds}} \text{fSim}(\theta, \phi). \tag{5.2}$$

In the limit $\theta, \phi \to 0$, $\hat{U}_F$ is the Trotter–Suzuki expansion of the XXZ Hamiltonian (5.1), with $\Delta = \sin(\phi/2)/\sin(\theta)$. Away from this limit, there is no unique Hamiltonian associated with $\hat{U}_F$, but Eqs. (5.1) and (5.2) still share symmetries and are both integrable by the Bethe ansatz [14, 29, 164, 36, 107].

To study dynamics under the unitary evolution (5.2), we generate domain-wall initial states with an adjustable contrast parameter μ (Fig. 5.1B). Specifically, we initialize the chain in a set of product states such that the left and right halves have average magnetization $\pm\tanh(\mu)$, respectively:

$$\rho(t = 0) \propto (e^{2\mu S^z})^{\otimes N_Q/2} \otimes (e^{-2\mu S^z})^{\otimes N_Q/2}. \tag{5.3}$$

When $\mu \to \infty$, the system approaches a pure domain-wall state with the two sides fully magnetized in opposite directions. Only when $\mu = 0$, the initial state is an infinite-temperature thermal state that preserves SU(2) symmetry. When $\mu \neq 0$, the magnetization is preferentially along the z-axis, breaking the SU(2) rotational

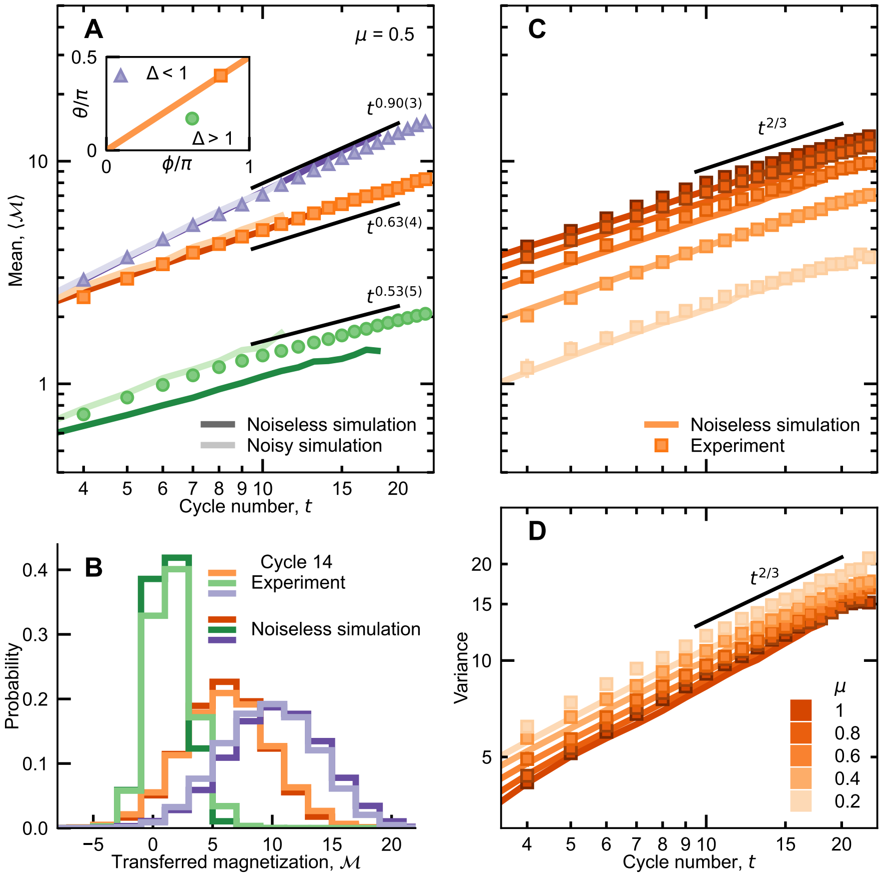


Figure 5.2: **Mean and variance in various transport regimes**. (**A**) Mean of transferred magnetization $\langle \mathcal{M}(t) \rangle$ as a function of cycle number for initial states with $\mu = 0.5$ and for $\Delta = 0.16$ (purple triangles), 1 (orange squares), 1.6 (green circles). Light and dark curves show simulations with and without noise, respectively. The $\langle \mathcal{M}(t) \rangle$ can be fit to $t^{1/z}$ and gives $z = 1.12 \pm 0.04$ in the ballistic, $z = 1.6 \pm 0.1$ in the superdiffusive, and $z = 1.9 \pm 0.2$ in the diffusive regime. The inset illustrates three different regimes characterized by $\Delta = \sin(\phi/2)/\sin(\theta)$ with the orange line being the isotropic Heisenberg limit. (**B**) Histogram showing the probability distribution of measured $\mathcal{M}$ for values of Δ studied in **A** at cycle 14. Light and dark lines show experimental data and noiseless simulation results, respectively. (**C**,**D**) Mean and variance of $\mathcal{M}$ for $\Delta = 1$ and $0.2 \leq \mu \leq 1$ (brighter to darker squares). With increasing μ, the mean increases, while the variance decreases.

symmetry of the Heisenberg model.

A natural measure of spin transport is the total transferred magnetization, $\mathcal{M}(t)$, defined as twice the net number of excitations that have crossed the middle of the chain after t cycles. In our experiment, we sample over initial bitstring states with probabilities given by Eq. (5.3). For each initial state, we prepare the qubits in that state and then apply t cycles of fSim gates. Let $N_{R,1}(b)$ be the number of excitations ("1"s) in the right half of bitstring b. The transferred magnetization $\mathcal{M}$ is the stochastic variable defined by

$$\mathcal{M}(t)/2 = N_{R,1}(b_t) - N_{R,1}(b_i), \tag{5.4}$$

where b_i is the initial bitstring, sampled from Eq. (5.3), and b_t is the associated final bitstring sampled at t. For example, if the initial bitstring is 111010 and the final bitstring is 110110, then the transferred magnetization is 2. Since the dynamics are number-conserving, the transferred magnetization is also the net number of zeros that have crossed from the right to the left. Repeating the experiment many times, we construct the probability distribution of $\mathcal{M}$, $P(\mathcal{M})$.

Fig. 5.1B shows measurement instances for three values of μ. The left column in each panel shows an instance of the initial state for the given μ, and the subsequent columns show typical bitstrings evolved from that state. As excitations (spin flips) propagate through the chain, smaller domains become more probable. In Fig. 5.1C, we show histograms of $\mathcal{M}$ at different times, starting in a pure ($\mu = \infty$) domain wall. Due to locality of the circuit, $|\mathcal{M}(t)|$ is upper-bounded by $2\,t$. Consequently, the distribution is narrow and centered around a small value at $t = 1$, since only a few excitations have crossed the middle of the chain, and becomes wider at later times.

In the context of spin transport, the first and second (variance) moments of $\mathcal{M}$ have been extensively studied both theoretically and experimentally [13, 104, 156, 147, 77, 107, 62, 41, 63, 43, 45, 46, 76, 150, 171, 88]. Taking advantage of our tunable fSim gates, we explore how these two moments depend on the anisotropy parameter, Δ. Fig. 5.2A shows the mean of $\mathcal{M}$ over time for values of Δ equal to 0.16 (purple), 1 (orange) and 1.6 (green), and an initial domain wall height of $\mu = 0.5$. We observe markedly different scaling behaviors in the three regimes. Eliminating the initial transient cycles, we fit a power law, $\langle \mathcal{M} \rangle \sim t^{1/z}$, to the data over cycles 10–23 and extract scaling exponents of $z = 1.12 \pm 0.04$, $z = 1.6 \pm 0.1$ and $z = 1.9 \pm 0.2$. These are in close agreement with theoretical predictions for the ballistic ($z = 1$)[183], superdiffusive ($z = 3/2$)[104], and diffusive ($z = 2$)[145, 39] behaviors, respectively. Observation of superdiffusive propagation for isotropic interactions ($\Delta = 1$), measured here and also in other works [169, 104, 156, 147, 77, 41, 62, 63, 45, 43, 46, 76, 150, 171, 88], has been interpreted as a signature of the KPZ universality class.

Numerical simulation of these domain-wall dynamics are shown with solid dark lines in Fig. 5.2A. A variety of numerical simulations often rely on approximation schemes, which could lead to inaccurate results. In contrast, here we perform *exact* statevector sampling up to cycle 18 without any approximations. This is achieved by taking advantage of the fact that $\langle\, \mathcal{M}(t)\, \rangle$ only depends on the spins within the light cone of width $2t$, and can thus be determined exactly by simulating shorter chains. This simplification also allows for arriving at analytical results for all moments of $\mathcal{M}$ at early cycles. Nevertheless, the computational cost grows exponentially, and with the resources used here, the simulations at cycles 14, 16, and 18 take about 1, 2, and 14 hours, respectively (see SM).

Importantly, the slight discrepancies between the observed and predicted exponents are also seen in the exact statevector simulations up to $t = 18$ cycles (colored lines), suggesting that these deviations primarily stem from finite-time and large-μ effects and are less affected by experimental imperfections. Indeed, the exact scaling exponents are only expected in the long-time limit and as $\mu \to 0$. By simulating the effects of noise in our system (lighter lines in Fig. 5.2A), we find that these are almost negligible for $\Delta < 1$ and $\Delta = 1$, and somewhat larger for $\Delta > 1$, due to the lower values of $\mathcal{M}$ in the diffusive regime. This effect is also noticeable in the magnetization transfer distributions in Fig. 5.2B. Since the distribution is narrower in the $\Delta > 1$ case, the noise has a larger effect on the shape of the distribution here than for the other two values of Δ. The error in this case is found to be predominantly caused by combined occurrences of T1 errors and $0 \to 1$ readout errors, which are not eliminated by post-selection (SM). By including this effect in the simulation, we find good agreement in all three regimes.

Superdiffusive transport, $\langle \mathcal{M} \rangle \sim t^{2/3}$, at $\Delta = 1$ is a characteristic of systems within the KPZ universality class. Moreover, numerical studies found that the spin-spin correlation function coincides with the KPZ scaling function [106, 137],

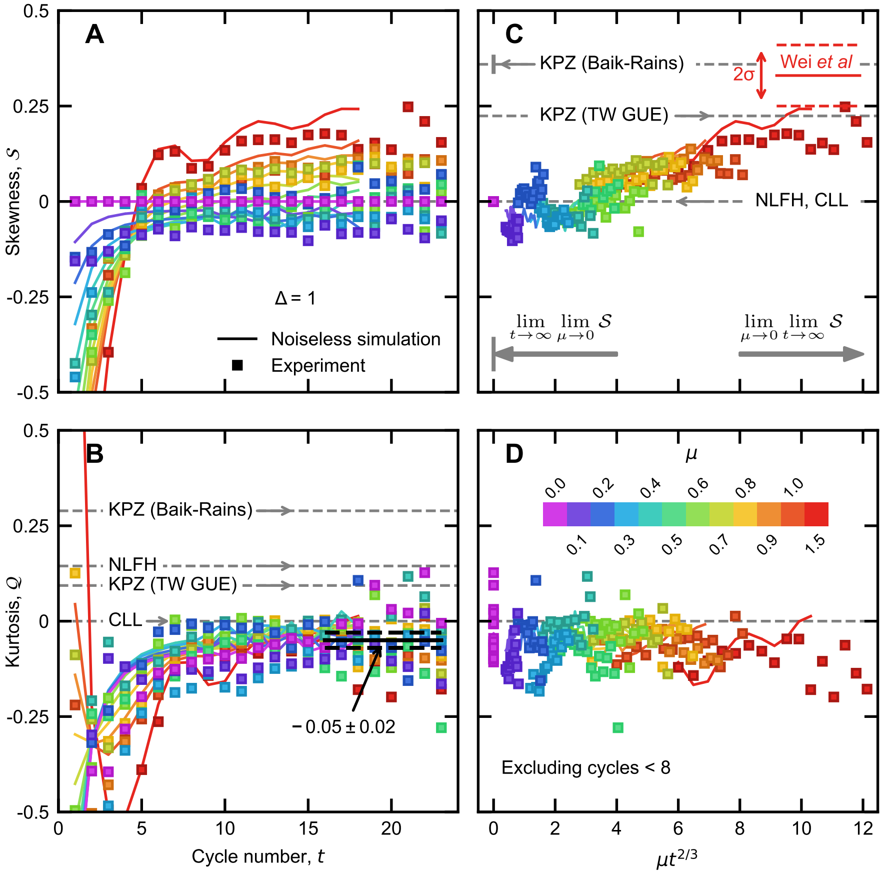


Figure 5.3: **Skewness and excess kurtosis of transferred magnetization**. (**A**) Skewness of transferred magnetization distribution $\mathcal{S}$ as a function of t, for $\Delta = 1$ and various $\mu \in [0, 1.5]$. Experimental data and noiseless simulation results are shown with squares and lines, respectively. We symmetrize the $\mu = 0$ probability distribution, after which the skewness is exactly 0. (**B**) Same as (**A**), but for excess kurtosis of the transferred magnetization distribution $\mathcal{Q}$. (**C, D**) Analogous to (**A**) and (**B**), but with the x-axis re-scaled as $\mu t^{2/3}$ and excluding data points for which $t < 8$. Dashed horizontal gray lines indicate predictions based on the KPZ universality class (TW GUE: Tracy-Widom Gaussian-unitary-ensemble), and nonlinear fluctuating hydrodynamics (NLFH) model [123]. The red horizontal lines show $\mathcal{S}$ measured in Ref. [171] ($\mu = 1.53$), and the 1σ confidence interval.

which has led to the conjecture that near-equilibrium spin transport in the Heisenberg model belongs to the KPZ universality class [106, 100, 178]. This universality class is associated with the classical nonlinear stochastic KPZ equation $\partial h/\partial t = \nu\nabla^2 h + \lambda(\nabla h)^2 + \eta(x,t)$, which was originally introduced [143] to describe the dynamics of driven interfaces as a height field $h(x,t)$, where ν, λ, η set the strength of the smoothening diffusion, roughening nonlinear growth, and stochasticity terms, respectively. The conjecture asserts that at late times the magnetization profile behaves similarly to $\partial h(x,t)/\partial x$. Consequently,

$$\lim_{\mu\to 0} \mathcal{M}(t) \ \longleftrightarrow\ 2h(0,t) - h(-\infty,t) - h(\infty,t). \tag{5.5}$$

To further examine the universality class of the Heisenberg spin dynamics, two aspects are of particular importance. First, since the universal behavior is expected to depend on whether the system is in equilibrium, it is essential to measure the dependence on μ. Second, while the scaling exponent of the mean is consistent with the KPZ universality class, further insights can be gained by examining higher moments (the "full counting statistics") of $P(\mathcal{M})$. Due to the reduced signal-to-noise ratio, measuring higher moments at small μ is experimentally challenging. We utilize our fast sampling capability to measure $P(\mathcal{M})$ as a function of μ and t (Refs. [99, 65]). Figs. 5.2C,D show the temporal evolution of $\mathcal{M}$ and its variance for various values of μ ranging from 0.2 to 1. We find that, at small μ, the dynamical exponents of both the mean and the variance are close to 3/2. For larger μ, the dynamical exponent of the mean approaches 5/3, consistent with recent numerical results [104] (SM).

Next, we extract the skewness $\mathcal{S}$ and kurtosis $\mathcal{Q}$ of $P(\mathcal{M})$,

$$\mathcal{S} = \frac{\alpha_3}{\alpha_2^{3/2}}, \quad \mathcal{Q} = \frac{\alpha_4}{\alpha_2^2} - 3, \tag{5.6}$$

where $\alpha_k = \left\langle (\mathcal{M} - \langle \mathcal{M} \rangle)^k \right\rangle$ is the k^{th} moment. In Fig. 5.3A, we show the temporal dependence of $\mathcal{S}$ for μ ranging from 1.5 to 0.1. Consistent with Ref. [171], $\mathcal{S}$ approaches to about 0.3 for $\mu > 1$. However, as μ is reduced towards the equilibrium point, we observe that $\mathcal{S}$ goes to zero. Fig. 5.3B shows that for later cycles, the initial strong time dependence of $\mathcal{Q}$ weakens. By averaging over cycles 16 to 23, we obtain a kurtosis of -0.05 ± 0.02 (Fig. 5.3B).

In order to test the KPZ universality conjecture, one needs to study the infinite-time ($t \to \infty$) and near-equilibrium ($\mu \to 0$) limits. These limits are experimentally inaccessible. However, if there exists a function $f(\mu, t)$ such that the moments are functions of $f(\mu, t)$, then one may be able to extrapolate measured values at finite μ and t to these unattainable limits. We empirically find that the zero crossing of $\mathcal{S}$ scales as $t_0 \sim \mu^{-1.49}$ (SM), suggesting that $\mathcal{S}$ may be a function of $\mu t^{2/3}$. Indeed, after excluding the initial transient behavior, $\mathcal{S}$ does appear to be a single-valued function of $\mu t^{2/3}$ (Fig 5.3C, SM).

KPZ has been conjectured to apply to high-temperature thermal states at late times, corresponding to taking $\mu \to 0$ first and then $t \to \infty$ (Ref. [106]). In this case, $P(\mathcal{M})$ should become the Baik-Rains distribution [138]. However, this distribution is skewed (Table 5.1), whereas our measurements suggest that $\mathcal{S} = 0$ (Fig. 5.3C), as is also dictated by symmetry.

One might also search for KPZ universality away from $\mu = 0$, corresponding to a different order in taking these noncommuting limits. When taking $t \to \infty$ first, the appropriate probability distribution to compare $P(\mathcal{M})$ against is the Tracy Widom (TW) distribution [138], which has $\mathcal{S}$ of about 0.22. This order of limits corresponds to large $\mu t^{2/3}$ in Fig. 5.3C, where we indeed find $\mathcal{S}$ consistent with this distribution, as also seen in earlier experiments [171]. However, $\mathcal{Q}$ of the TW GUE (Gaussian unitary ensemble) distribution is 0.09, whereas we find $\mathcal{Q} = -0.05 \pm 0.02$. The emergence of KPZ dynamics in this order of limits is further ruled out by numerical and theoretical predictions that the dynamics become diffusive ($z = 2$) on a timescale $t \sim 1/\mu^3$ (Refs. [99, 64]).

One could consider taking the two limits simultaneously in a way that the dynamics do not become diffusive, e.g., by holding $\mu t^{2/3}$ constant. The correct distribution to compare against is TW GUE in this case as well. If we take the limit with $\mu t^{2/3}$ fixed at a large value, we find $\mathcal{S}$ consistent with TW GUE, but the measured $\mathcal{Q}$ is still inconsistent with the TW GUE prediction of 0.09, ruling out KPZ dynamics on the timescales accessible in the experiment. While it remains possible that KPZ dynamics will emerge at much later times (i.e., $\mathcal{Q}$ will increase to 0.09), we see no evidence or motivation for this.

An outstanding question is why only lower-point observables seem to behave consistently with KPZ universality. Intriguingly, other systems have been identified that exhibit similar behavior. One such system is a nonlinear fluctuating hydrodynamic (NLFH) model with two coupled stochastic modes [21, 22, 123], which predicts $\mathcal{S} = 0$,

	$\langle \mathcal{M} \rangle$	σ^2	$\mathcal{S}$	$\mathcal{Q}$
Experiment	$t^{2/3}$	$t^{2/3}$	$0\,*$	-0.05 ± 0.02
KPZ (Baik-Rains)	$t^{2/3}$	$t^{2/3}$	0.36	0.29
NLFH	$t^{2/3}$	$t^{2/3}$	0	0.14
CLL	$t^{2/3}$	$t^{2/3}$	0	$\in [-0.03, 0.03]$

Table 5.1: Comparison of the experimentally observed moments with the theoretically predicted values from various models. The experimental kurtosis value is averaged over cycles $16 - 23$ and $\mu = 0 - 0.4$, and the errors are computed using the jackknife method. Green (red) entries indicate agreement (disagreement) between the data and theory. See supplement for details regarding KPZ predictions. *The skewness is 0 due to symmetrization of our data.

consistent with the Heisenberg spin chain. However, it suggests $\mathcal{Q} = 0.14$, differing from what we observe, perhaps because not all aspects of the model are universal. Another such system is the classical Landau-Lifshitz (CLL) magnet [99, 98], which predicts $\mathcal{S} = 0$ and a $\mathcal{Q}$ that is negative and close to zero at these time scales [98]. These are consistent with our experimental results. It is rather surprising that this classical system is so successful in capturing the behavior of a quantum spin chain with enhanced quantum fluctuations due to confinement [59].

Studies of the universal aspects of quantum dynamics have attracted notable interest recently; accordingly, a complete classification of their universal properties is lacking. Our findings suggest these classifications could involve unanticipated subtleties. Our first result, also observed by others, is the superdiffusive transport characterized by $\langle \mathcal{M} \rangle \sim t^{2/3}$, shown in Fig. 5.2A. While this anomalous diffusion is suggestive of the known KPZ universality classes, this classification is not compatible with our second finding—the vanishing of $\mathcal{S}$ and $\mathcal{Q}$ near equilibrium (Figs. 5.3C, D).

Despite the apparent success of the CLL model, a full understanding requires the development of a systematic spacetime RG framework that could establish the origin of the KPZ-like behavior starting from the microscopic dynamics of the Heisenberg model. Quantum processors have the potential to help with such RG studies. For example, the multi-scale entanglement re-normalization ansatz (MERA) applies ideas of RG flow to tensor networks and quantum circuits [165, 53].

Our observations are rooted in the interplay of integrability, quantum fluctuations, and symmetry and have proved to be challenging to describe using an effective quantum field theory. The observed discrepancies with KPZ predictions suggest that the infinite-temperature dynamics in the Heisenberg chain—if universal—belong to a yet-to-be-discovered dynamical universality class.

Contributions

This chapter is an adaptation of [143]. I implemented the experiment, calibrated the gates, developed error mitigation techniques, collected experimental data, developed numerical simulation techniques, ran the numerics, made the figures, and contributed to the interpretation and writing. The experiment was proposed by Tomaz Prosen, Sarang Gopalakrishnan, and Vedika Khemani, who also provided theory support. Pedram Roushan selected the proposal for me to work on and directed the project. Trond Andersen collected additional experimental data, contributed to writing and figure preparation, and was actively involved in discussions of interpretation. Rhine Samajdar performed additional numerics and contributed to manuscript preparation and interpretation. The project was made possible by the Google Quantum AI team, which provided infrastructure and support.

CHAPTER 6

MEASUREMENT INDUCED PHYSICS

Like Chapter 3, this chapter details my contributions to a project on which I was not the lead author, namely version 2 of Ref. [74], of which I am the third author. This work uses a Sycamore processor to study the dynamics of monitored circuits. As illustrated in Figure 6.1, it takes advantage of a mapping from the monitored circuit (panel a) to one with only terminal measurements (panel c). We are interested in the following a particular trajectory among the mid-circuit measurements, which, in the circuit that is actually implemented (panel c), corresponds to post-selecting on a particular measurement outcome for the upper qubits. This collapses the remaining qubits into a pure state $|\psi_m\rangle$. Obtaining decent statistics to characterize $|\psi_m\rangle$ requires many shots because we need to post-select on the same measurement outcomes each time.

Finally, the second Rényi entropy is measured on subsystems of $|\psi_m\rangle$ using randomized measurements, following the protocol of Refs. [51, 81, 48, 19, 149]. A transition from extensive to sub-extensive entanglement is observed as the circuit parameters are varied (panel e). This utilizes an error mitigation technique whereby the subsystem entropy due to decoherence is subtracted off. The measured Rényi entropies are in good agreement with numerical simulations (see the supplemental material to Ref. [74]).

My contribution to this project was to calibrate fSim gates and collect the experi-

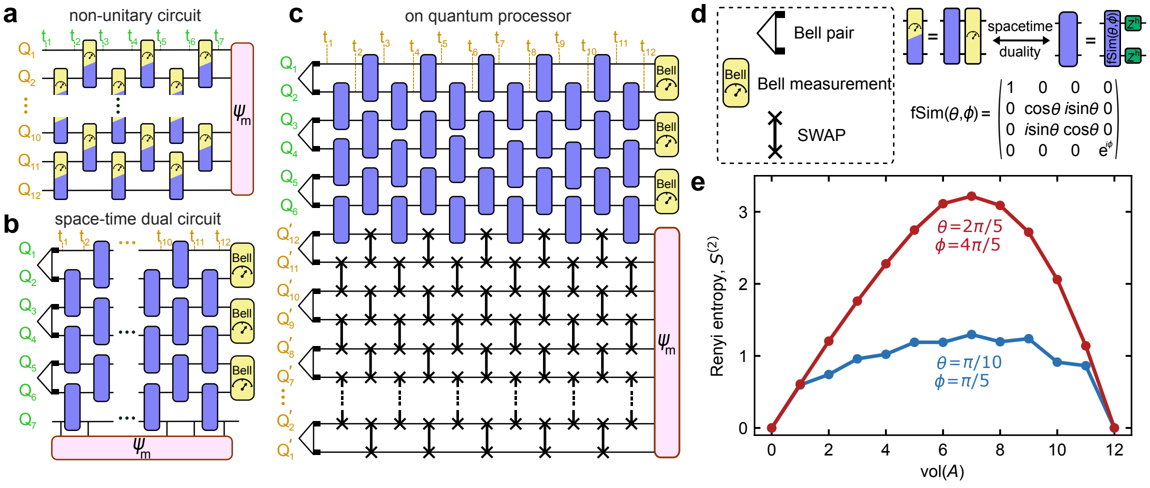
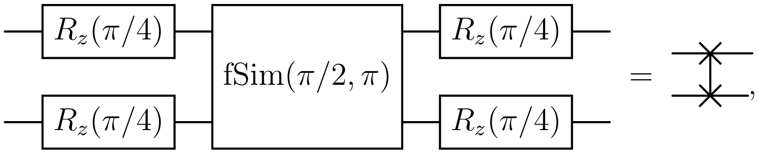


Figure 6.1: An experiment to study the dynamics of a monitored circuit, highlighting a tunable crossover from volume-law to area-law entanglement.

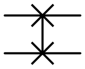
mental data that appear in panel c. I used the same code and techniques to calibrate fSim gates for this project as for Chapter 5, which are detailed in Appendix C.1.3.

There are two important differences between gate calibration for this work and for the project described in Chapter 5. One is the additional challenge of calibrating multiple choices of fSim angles simultaneously. The circuit depicted in Figure 2 of Ref. [74] involves three different two-qubit gates: (1) CZ gates, used to initialize the Bell pairs, (2) SWAP gates, used to extend the output state $|\psi_m\rangle$ over space instead of time, and (3) fSim(θ, ϕ), which implements the dynamics; the fSim angles are chosen here to be either $(\theta, \phi) = (0.1\pi, 0.2\pi)$ or $(\theta, \phi) = (0.4\pi, 0.8\pi)$. The other distinction is that the 1D chain used in this experiment contains only 18 qubits, whereas the chain used in Chapter 5 contains 46 qubits.

My strategy in performing this experiment was to begin by calibrating the 46 qubits used in Chapter 5 and then pick an 18-qubit sub-chain based on the single- and two-qubit gate fidelities, measured using RB and XEB, respectively, as well as on readout error rates (see Section 1.3). The ordering of the qubits within the chain (i.e. which end to call qubit 0) was determined based on the fidelities of the SWAP and fSim(θ, ϕ) gates in their respective portions of the chain (fSim(θ, ϕ) gates are performed on the first 7 qubits, and SWAP gates are performed on the last 12 qubits). CZ gates are implemented as fSim($0, \pi$), and SWAP gates are implemented as fSim($\pi/2, \pi$), plus single qubit Z rotations. In particular,



(6.1)

where  is a SWAP gate, and the equality is up to a global phase. In order to reduce leakage, a longer rise time of 8 ns was used for the SWAP gate, as described in the supplement to Ref. [74]. The resulting two-qubit error rates are shown in Figure 6.2.

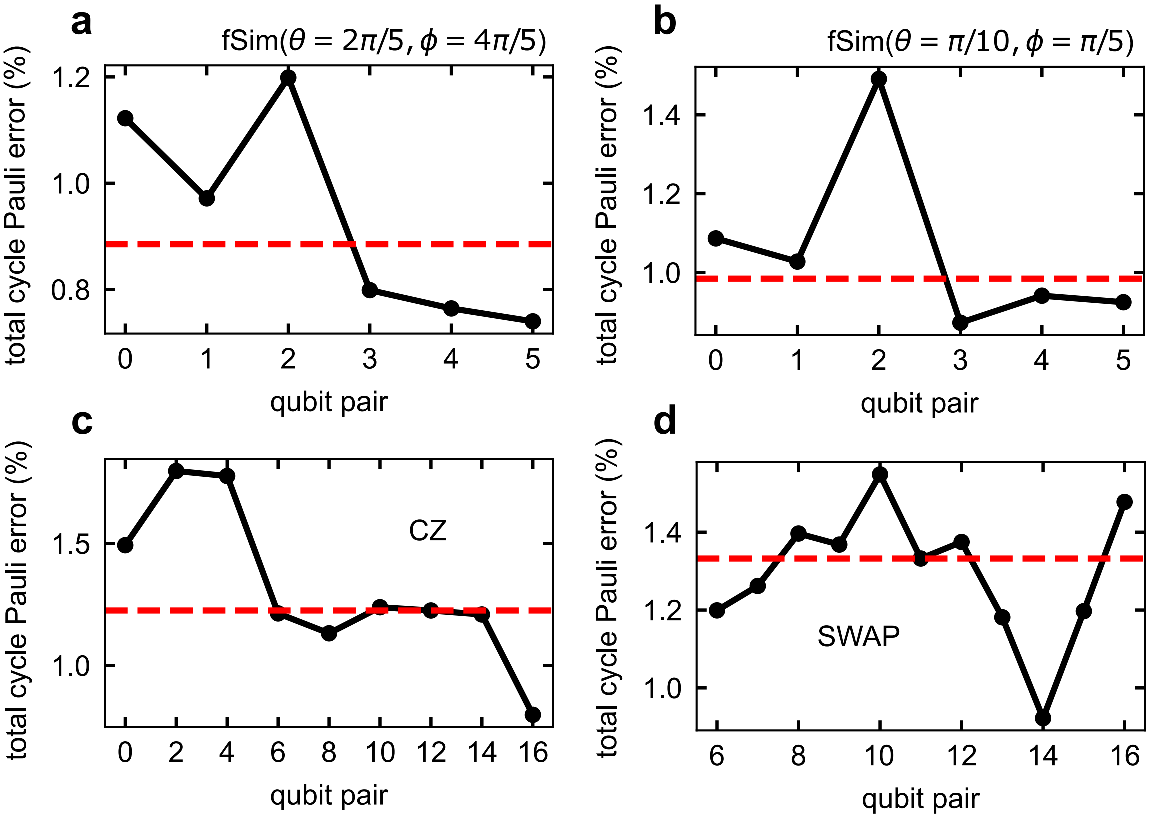


Figure 6.2: The error rates of the two-qubit gates used in Figure 6.1. The red dashed line indicates the median.

CHAPTER 7

CONCLUSION

This thesis has considered a number of projects that all work toward the goal of achieving useful quantum advantage, that is, of using an existing or near-term quantum processor to solve a problem that is both of interest to the community and cannot be solved using classical methods. What constitutes "useful" is open to some interpretation, as is the level of proof required to demonstrate quantum advantage. Chapter 5 of this thesis is perhaps the closest that has been achieved. The experiment detailed there is useful in that it addresses an outstanding question in the scientific community, and the quantum hardware exhibits an advantage over the classical techniques tested so far (see also Ref. [146]). However, there may well be other classical techniques that can solve the problem efficiently [31]; testing such techniques should be a high priority for determining whether this experiment might have achieved quantum advantage or whether an incremental scaling up of the experiment might be able to do so. After outlining how this scaling up might be achieved, I take a broader view and offer reflections on the directions that I find most promising for the next 2–5 years.

7.1 Immediate follow-ups to Chapter 5

If, after thoroughly investigating known classical simulation techniques, it turns out that the higher moments of the transferred magnetization are indeed difficult to compute classically, several improvements can be made to the experiment to increase the number of qubits and cycles. In Appendix C, I show that achievable circuit depths are limited by amplitude damping and that the algorithmic relaxation time, T_1^A, is approximately a factor of two worse than one would expect from the measured singe-qubit amplitude damping rates at their idle frequencies. One explanation for the shorter algorithmic relaxation time has to do with the frequency trajectory that the qubits follow over the course of a cycle. As illustrated in Figure C.2, the qubit's lifetime is a function of its frequency. The idle frequency is carefully picked with this in mind, but over the course of a cycle, the qubit is at other frequencies that are not carefully chosen with T_1 in mind. By optimizing the frequency trajectory, it may be possible to improve the algorithmic relaxation time and hence the number of cycles that can be achieved in the experiment.

Further, errors that make it past the post selection appear to be dominated by a combination of amplitude damping and $0 \to 1$ readout errors. This explains the observed discrepancies between the $\Delta > 1$ experimental data and noiseless simulations (Figs. C.5, C.14). Although the $\Delta = 1$ data, which is most relevant to the KPZ conjecture, appears more robust to these errors, suppressing this effect may nevertheless be beneficial. This can be achieved by an alternate readout calibration that favors low $0 \to 1$ error rates at the expense of higher $1 \to 0$ error rates. Although such a calibration would reduce the number of counts that pass the post-selection, it would increase the quality of the post-selected data. Balancing these competing considerations amounts to balancing the systematic and statistical sources of error, which may be carefully investigated in order to improve the experiment.

7.2 Strategies for the next 5 years

Next, I step back and make some high-level observations and suggestions about strategies that seem promising for applications of quantum processors in the next 5 years. I do not comment on efforts towards error correction, which, for example, Google is actively pursuing [4, 2], instead focusing on so-called NISQ applications that do not require error correction.

7.2.1 Application first or method first?

There are two competing, but perhaps complementary, strategies for how to achieve useful quantum advantage on near-term hardware. One strategy is what I call "method first," in which error mitigation itself is the goal; the work is either meant to be application-independent, or the application is chosen to showcase the error mitigation technique. This strategy is exemplified by IBM's work in this area (e.g. Refs. [12, 91]), as well as by Chapters 2 and 4 of this thesis. The method-first philosophy is in line with the community's approach to error correction, which is application-agnostic. The opposite philosophy, exemplified by many of Google's NISQ experiments, including those described in Chapters 3, 5, and 6, is what I would characterize as "application first," in the sense that the particular application is the primary goal and any error mitigation is specifically tailored to that application.

Both strategies have their merits. The method-first approach has led to the development of powerful, general techniques for error mitigation based on scalably characterizing the noise and then either inverting it [12] or amplifying it and extrapolating back to zero [91], along with a number of other techniques described in Chapter 2. Eventually, in order to demonstrate useful quantum advantage, these techniques will need to be applied to a carefully-chosen problem so that the quantum processor teaches us something interesting, i.e. eventually, the application will need to become the focus.

The application-first approach has the advantage that often, application-specific error mitigation techniques offer an advantage over more general ones. For example, the error mitigation employed in Chapter 5 of this work takes advantage of number-conservation in the ideal dynamics, as well as the causality structure of the circuit, whereas the error mitigation employed in Chapter 6 (described in detail in Ref. [74]) takes advantage of how the entanglement between the system and its environment is expected to grow with system size. These techniques might be considered too simple or too problem-specific to be studied for their own sake, but they can still be highly effective for these particular problems, and in the search for any useful application of existing or near-term quantum processors, problem-specific techniques like these might pay off.

Chapter 4 of this thesis provides an example of some of the pitfalls that can arise when taking the method-first approach. In this chapter, I have a specific application in mind (VQE for SYK), but focus on trying to mitigate correlated readout errors generally rather than on solving the test problem end-to-end. This approach hindered me a bit in working on this project. I was immersed in the technical details of characterizing and mitigating correlated readout errors before asking how important this was and whether it would actually enable VQE for SYK. It was only after developing the techniques to characterize and mitigate correlated readout errors that I discovered that, on the 27-qubit devices, correlated readout errors did not have a significant effect on the observable of interest (Fig. 4.1). On the 65-qubit device, they may have had a small but statistically significant effect (Fig. 4.2). However, on this device, the correlation structure was too complicated to easily invert (Fig. 4.7). It is possible that it could be made tractable by only considering the most important correlations, but even so, it seems that VQE for SYK likely requires deeper circuits than current gate fidelities and coherence times support (Fig. B.3). An application-first approach would have forced me to think about these considerations from the beginning and only devote significant time and effort to solving the correlated readout error problem if it enabled me to study something interesting end-to-end.

I do not intend to diminish the value of method-first work, which has led to the development of a variety of promising error mitigation techniques that may well find useful application. However, in these early days of quantum processors, when the community is still trying to demonstrate any useful application of the technology, it makes sense to hand-pick applications that are well suited to the technology, tailoring the error mitigation to the application. This focuses the technical effort on precisely what is needed to solve the problem and also allows one to take advantage of features that may not be present in the general case.

7.2.2 Which applications?

Having expressed a preference for beginning with a particular application in mind, I now offer some advice on picking an application.

Consider the circuit depth.

Until error correction leads to fault-tolerance, perhaps the most important consideration when picking an application is that the fidelity of the quantum state decays exponentially due to T_1, T_2, and gate errors. No matter what error mitigation techniques one employs, the signal is exponentially suppressed and therefore requires an exponentially large number of measurements in order to be distinguished from random noise. A non-error-corrected quantum processor is not going to provide an exponential speedup over classical computers (unless one finds a problem that maps to a constant-depth circuit and is hard classically).

Quantum advantage is rather a game of two competing exponentials: (1) the exponential cost, in runtime and memory, of classically simulating quantum circuits versus (2) the exponential growth of the number of shots needed on a quantum processor to measure an exponentially suppressed signal. The hope is that, although the cost is exponential on both sides, the quantum processor may come out ahead. This has already been shown to be the case for random circuit sampling (the so-called "quantum supremacy" experiment) [5, 174, 182, 120]. The goal of the NISQ project for the next few years, before error correction is available, will be to identify other, more useful, applications where quantum processors provide an advantage.

Although some quantum systems are more interesting to study than others, I recommend starting with a relatively simple system rather than the one that is most interesting. As described earlier in this chapter, I began my quantum computing career thinking that I would simulate the Sachdev-Ye-Kitaev model. However, as discussed in Appendix B, this system is very difficult to simulate, compared to the local spin chains that have been successfully simulated (although not yet definitively beyond the capabilities of classical computers). My advice is to focus on systems that have only $O(1)$ circuit depth per Trotter step, at least until quantum advantage has been demonstrated for these systems. Spin chains where the coupling in the dynamics matches the connectivity of the qubits are particularly well-suited to existing devices. Further, there is no need to remain in the Trotter limit; faster dynamics can be achieved by considering Floquet systems instead of Hamiltonian systems, which helps reduce circuit depth. I suggest noting the circuit depths that have been achieved in state-of-the-art experiments to date (e.g. [120, 91], and Chapter 5 of this work, or whatever has been achieved to date) and avoiding anything significantly deeper than that (absent any innovation that makes deeper circuits possible).

Consider classical simulability.

The conventional wisdom is that 1D is easy and 2D is hard for classical simulations. While there may be exceptions to this rule, it will generally be beneficial to take advantage of any 2D connectivity present in the device in order to make the circuits more difficult for classical methods. Further, if the quantum processor supports multiple types of entangling gates, consider which is more difficult to simulate classically. For example, CZ and CNOT gates are only rank-2, so applying a CZ gate in a tensor network simulation does not increase the bond dimension by as much as a generic two-qubit gate, which is rank-4. Understanding the capabilities and limitations of classical simulation techniques is crucial for identifying an application that may exhibit quantum advantage and can help avoid the pitfalls of Ref. [91], which claimed quantum advantage and was almost immediately followed by three papers that demonstrated that their observables could be computed classically [87, 10, 160].

Consider usefulness.

Although the most interesting quantum systems may be out of reach for now, there are outstanding questions about simpler systems that one may hope to answer using near-term devices. Questions about local spin systems may not be impactful enough to justify billions of dollars of investment, but that is not the goal; the investment will not pay off before error correction. The goal is simply to find something that is more useful than random circuit sampling or whatever applications have been demonstrated to date.

7.2.3 Considerations when running the experiment

After picking an application, it is time to design the experiment. This subsection discusses various strategies to improve the quality of the experimental data.

Calibration, mitigation, and error avoidance

Error mitigation is not the only way to reduce the effects of noise. Another way is through improved calibration. The tunable-frequency and tunable-coupler architecture at Google provides many extra knobs to adjust compared to the fixed-frequency architecture with cross-resonance based coupling that is available at IBM. Because of this, calibration is a more significant component of the experiments in Chapters 5 and 6, performed on Google hardware, than in of those in Chapters 2 and 4, performed on IBM hardware.

In order to calibrate the hardware for a particular experiment, it is important to understand what types of errors are dominant in that experiment, because sometimes, one type of error can be reduced at the expense of another type of error. An error budget can be learned through simulations with noise, as in Section C.2.3 for Chapter 5. In this particular experiment, I find that the dominant sources of noise are amplitude damping and $0 \to 1$ readout errors. Therefore, improving the algorithmic T_1 could benefit the experiment even if it is at the expense of increased gate errors. As discussed earlier in this chapter, reducing the $0 \to 1$ readout error rate could also benefit the experiment.

In addition to calibration and error mitigation, careful circuit design can improve an experiment. I call this error avoidance. For example, in Chapter 5, if the initial bitstring is more than half-full (more 1s than 0s), the physical states corresponding to $|1\rangle$ and $|0\rangle$ are interchanged, which improves the algorithmic T_1 (Fig. C.1). Similarly, in Chapter 2, gates outside the backwards light-cone of the measured operators are omitted, reducing the effects of cross-talk. Simple improvements such as these in experimental design can go a long way in improving overall fidelities.

Benchmarking

On existing transmon-based devices, error rates fluctuate over the course of hours due to what are known as TLSs (e.g. [103, 179]). Therefore, it is important to regularly benchmark the performance of the qubits being used in an experiment. This includes monitoring readout error rates, single-qubit error rates, and two-qubit error rates, the latter of which can be characterized using either randomized benchmarking or cross-entryopy benchmarking, discussed in Chapter 1. This helps identify when part of the device is broken so that it can be either recalibrated or avoided. Measured error rates should be compared against those reported in state-of-the-art papers (e.g. [120, 91, 4] or whatever papers are current) and not perform the experiment without attaining comparable error rates.

Simulations

Classical simulations are critical to NISQ experiments. They serve three purposes: (1) validation of the experiment, (2) understanding the dominant sources of noise, and (3) understanding whether the experiment is a candidate for quantum advantage, and if so, what resources (i.e. number of qubits, circuit depth) are needed. Even if one considers (3) before beginning the experiment, it is worthwhile to return to (1) and (2) later. I recommend using measured error rates as inputs to the noisy simulations. The simulations enable identification of the types of noise that contribute most significantly to systematic errors in the measured observables. This knowledge can be used to reduce these sources of error, possibly at the expense of others. When measurements are dominated by statistical errors, I recommend taking data faster, possibly at the expense of increasing systematic errors (c.f. Sec. 7.1), as long as doing so does not make the systematic errors larger than the statistical errors.

7.3 Outlook

We are fortunate to live in a time when quantum processors are just starting to extend beyond the capabilities of classical computers in a limited set of problems. With further improvements in device fabrication and calibration, as well as improvements in experiment design and error mitigation, I am confident that we will see increasingly useful and increasingly difficult-to-simulate applications of quantum processors in the near future, although such applications will likely be restricted to certain easy-to-simulate systems such as locally coupled spins (or possibly nonlocally coupled spins on a trapped ion device where native nonlocal gates can be performed). Although other systems with more direct commercial applications will likely remain out of reach in the near future, near-term devices hold the potential to inform our understanding of important topics in condensed matter physics, such as thermalization, many body localization, and the behavior of monitored quantum systems. This is tremendously exciting for the physics community but falls short of the world-changing promises of quantum computing. For those promises to be realized, we will need error correction, which various research groups, such as Google Quantum AI [2], are actively pursuing.

APPENDIX A

APPENDIX TO CHAPTER 2

A.1 Perturbation theory for the mixed-field Ising model

The mixed-field Ising Hamiltonian (Eq. 2.1) is analytically solvable in several limits, as discussed in the main text. In the small-h_x limit, it is diagonal in the computational basis, and the ground state (assuming $J > 0$ and $h_z > 0$) approaches $|00\ldots00\rangle$. The ground state energy, to second order in h_x is

$$E_{\text{gs}} = -n\left(h_z + J + \frac{h_x^2}{2h_z + 4J} + O(h_x^3)\right). \tag{A.1}$$

Further, in the large h_x limit, the ground state approaches $|++\ldots++\rangle$, where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. In this limit, the ground state energy is

$$E_{\text{gs}} = -n\left(h_x + \frac{2h_z^2 + J^2}{4h_x} + O\left(h_x^{-2}\right)\right). \tag{A.2}$$

Finally, in the small h_z limit, the Hamiltonian approaches the transverse-field Ising model, which, in the large-n limit, is exactly solvable via a Jordan-Wigner transformation [133]. Deriving an analytical expression for the ground state energy in the small h_z limit is beyond the scope of this work. Instead, we numerically compute the energy to second-order in perturbation theory in this limit (i.e. using the exact numerical eigenvectors from $h_z = 0$).

The perturbative results are compared to the exact ground state energies in Fig. A.1.

A.2 A note about error bars

Throughout Chapter 2, error bars indicate statistical errors only. There are, of course, a variety of systematic errors as well. However, the present goal is to correct these systematic errors, so it is more appropriate for error bars to only show statistical errors, which result from having a finite number of counts. We consider the variation among the qubit assignments used in the qubit assignment averaging to be a systematic error, as the same choices of qubit assignments are used in the calibration circuits as in the circuit that we wish to correct. However, we consider the variation over the different "perturbative" points in the method introduced in Sec. 2.3.2 to be statistical error, since a larger sample of such points can always be taken.

A.3 Variational ansatz circuit

As described in the main text, our ansatz consists of CNOTs between adjacent qubits and Y-rotations on all qubits (Fig. A.2). CNOT is a basis gate on IBM's quantum computers. However, the Y-rotations and the change of basis required to measure X_i must also be written in terms of basis gates. We use the following basis-gate decompositions:

$$\begin{aligned} R_y(\theta) &= R_z(\pi)\sqrt{X}R_z(\pi+\theta)\sqrt{X} \\ HR_y(\theta) &= R_z(\pi)\sqrt{X}R_z(3\pi/2-\theta)\sqrt{X}R_z(-\pi). \end{aligned} \tag{A.3}$$

ibmq_toronto and *ibmq_sydney* include 12- and 20-qubit loops, the topologies studied here, so we have not needed to apply any gates between non-neighboring qubits.

A.4 Poorly performing methods

In addition to the methods discussed in the main text, we also benchmark two methods that do not perform as well. In one of these methods, which we call "multiplying fidelities," we multiply $(1 - e_i)$, where e_i is the reported error rate for gate i, over all of the gates appearing in the backwards light cone of the observed qubits. This is then taken as an estimate of the overall damping factor. In Fig. A.4, it is clear that this method predicts more suppression than actually exists. The other poorly performing method is to use the noise model in Qiskit Aer, which includes local depolarizing channels on all of the gates, thermal relaxation errors, and readout errors. This method is not scalable since it requires simulating the circuit, but we benchmark it for comparison. It consistently predicts less suppression than exists.

A.5 Readout error mitigation

In the standard technique for readout error mitigation on n measured qubits, one prepares the qubits in each of the 2^n computational basis states and measures the probabilities of reading out the different outcomes (e.g. [6]). This results in a $2^n \times 2^n$ matrix, which is inverted to perform the mitigation. When n is large, this method

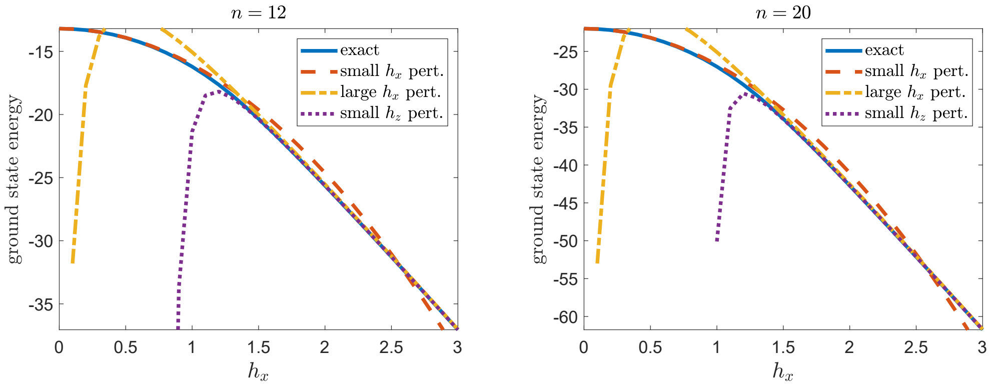


Figure A.1: Exact ground state energy of the mixed-field Ising model (Eq. 2.1) and predictions from second-order perturbation theory, for $h_z = 0.1$. Relevant to this work, we see that $h_x \leq 0.5$ is in the small h_x perturbative regime, while $h_x = 1.5$ is outside of it (at least to second order). For the small-h_z limit, we cheat slightly and use exact diagonalization to compute the eigenstates and eigenenergies when $h_z = 0$. We thus include the cyclic boundary term that is neglected in analytic treatments (cf., Eq. 2.4 in [133], and the following discussion), so the small-h_z line should be regarded as slightly better than one could do analytically, but approaching the analytical solution as the number of spins increases.

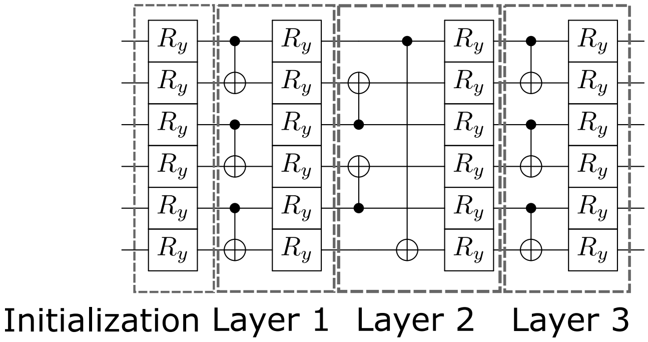


Figure A.2: The Alternating Layered Ansatz that we use throughout this work, shown for 6 qubits and 3 layers. When we impose partial cyclic permutation symmetry, we demand that, within a layer of y rotations, all of the rotations on even qubits have the same angle, and all of the rotations on odd qubits have the same angle.

becomes intractable and one must make simplifying assumptions, such as ignoring correlations among the qubits or among sets of qubits. For the Mixed-Field Ising Hamiltonian (Eq. 2.1), we only need to measure at most two qubits simultaneously, so the standard readout error mitigation method is easily tractable. Nevertheless, we assume uncorrelated readout errors and use the methods developed in Appendix A.6 to perform the mitigation. We use the readout error rates measured by IBM during their approximately daily calibrations [3]. We have found that this works about as well as doing the full readout error mitigation using our own calibrations.

A.6 Uncorrelated readout errors

In this appendix, we show that the effect of uncorrelated readout errors is to reduce the expectation value of a Pauli operator by the factor $(1 - 2e_m)$, for each qubit measured, where $e_m = (e_0 + e_1)/2$ is the average readout error, and offset it by a small additive amount. Some similar results appear in [58].

Suppose that we want to measure a Pauli string of length N by applying readout gates to all of the qubits. If the readout gates were noiseless, we would measure outcome q with probability $f(q)$. The noiseless expectation value of the Pauli operator is then $\langle P\rangle = \sum_q P(q)f(q)$, where $P(q) = \pm 1$ is the bit parity of q. Now, suppose that the measurement gates add noise such that the probability of a 0 getting measured as a 1 is e_0 and the probability of a 1 getting measured as a 0 is e_1. Let $\tilde{f}(q)$ be the probability distribution of the outcomes, including this readout error. Let

$\langle \tilde{P} \rangle = \sum_q P(q)\tilde{f}(q)$.

The main result that we will establish is that

$$\langle \tilde{P} \rangle = \sum_q P(q) f(q) (1-2e_0)^{n_0(q)} (1-2e_1)^{n_1(q)}, \tag{A.4}$$

where $n_0(q)$ is the number of bits in q that are 0, and $n_1(q)$ is the number of bits in q that are 1. If the two error rates are equal ($e_0 = e_1 = e_m$), Eq. A.4 becomes

$$\langle \tilde{P} \rangle = \langle P \rangle (1-2e_m)^N, \quad e_0 = e_1 = e_m. \tag{A.5}$$

However, we can also simplify Eq. A.4 without assuming the error rates are equal. For $N = 1$, $\langle P \rangle = f(0) - f(1)$, so Eq. A.4 becomes

$$\langle \tilde{P} \rangle = \langle P \rangle (1 - e_0 - e_1) + e_1 - e_0, \quad N = 1. \tag{A.6}$$

For larger N, $\langle P \rangle$ does not determine $f(q)$. However, the largest and smallest $\langle \tilde{P} \rangle$ for a given $\langle P \rangle$ will always result from letting $f(q)$ be nonzero for at most two choices of $n_1(q)$, one with n_1 even and the other with n_1 odd. Define these values as n_1^+ and n_1^-, respectively. Define $n_0^\pm = N - n_1^\pm$ and $f_\pm$ to be the corresponding probabilities. Then $\langle P \rangle = f_+ - f_-$, and so Eq. A.4 becomes

$$\langle \tilde{P} \rangle = \left(\frac{1+\langle P \rangle}{2}\right) (1-2e_0)^{n_0^+} (1-2e_1)^{n_1^+} - \left(\frac{1-\langle P \rangle}{2}\right) (1-2e_0)^{n_0^-} (1-2e_1)^{n_1^-} \tag{A.7}$$

Without loss of generality, we assume that $e_1 \geq e_0$. We also assume that $e_1 < 0.5$. Then $\langle \tilde{P} \rangle$ is maximized by picking $n_1^+ = 0$ and n_1^- to be either N or $N-1$, whichever is odd. Similarly, the smallest $\langle\tilde{P}\rangle$ is obtained by picking $n_1^+ = N$ or $N-1$, whichever is even, and $n_1^- = 1$. For the case of $N = 2$, this becomes

$$(1-2e_1)\left[\langle P\rangle(1-e_1-e_0)-(e_1-e_0)\right] \leq \langle\tilde{P}\rangle \leq (1-2e_0)\left[(1-e_0-e_1)\langle P\rangle + e_1 - e_0\right]. \tag{A.8}$$

The midpoint between these two bounds gives an estimate of $\langle\tilde{P}\rangle$ when $N = 2$:

$$\langle\tilde{P}\rangle \approx \langle P\rangle(1-e_0-e_1)^2\langle P\rangle + (e_1-e_0)^2, \quad N = 2. \tag{A.9}$$

Note that the y-intercept is suppressed compared to the $N = 1$ case (Eq. A.6). Indeed, it becomes more suppressed as we go to higher N. Numerically, we find that, for $N \gtrsim 3$, for randomly chosen $f(q)$,

$$\langle\tilde{P}\rangle \approx (1-e_0-e_1)^N\langle P\rangle, \quad N \gtrsim 3 \tag{A.10}$$

where this becomes a better approximation as N increases. We present numerical evidence for this result in Figs. A.5–A.6.

A.6.1 Proof of Eq. A.4

Consider two bit strings, both of length N, q and q'. Define the following quantities:

$$\begin{aligned} k_0(q,q') &= \sum_{i=0}^{N-1} \mathbb{1}_{q_i'=0}\mathbb{1}_{q_i=1}, \qquad & u_0(q,q') &= \sum_{i=0}^{N-1} \mathbb{1}_{q_i'=0}\mathbb{1}_{q_i=0}, \\ k_1(q,q') &= \sum_{i=0}^{N-1} \mathbb{1}_{q_i'=1}\mathbb{1}_{q_i=0}, \qquad & u_1(q,q') &= \sum_{i=0}^{N-1} \mathbb{1}_{q_i'=1}\mathbb{1}_{q_i=1}. \end{aligned} \tag{A.11}$$

Then, upon adding readout error, the probability $\tilde{f}(q)$ of measuring bit string q is

$$\tilde{f}(q) = f(q)(1-e_0)^{n_0(q)}(1-e_1)^{n_1(q)} + \sum_{q' \neq q} f(q') e_0^{k_0(q,q')} e_1^{k_1(q,q')} (1-e_0)^{u_0(q,q')}(1-e_1)^{u_1(q,q')}. \tag{A.12}$$

The expectation value of bit parity in the distribution $\tilde{f}$ is

$$\begin{aligned}
\langle \tilde{P} \rangle =& \sum_q P(q) f(q) (1-e_0)^{n_0(q)}(1-e_1)^{n_1(q)} \\
&+ \sum_q \sum_{q' \neq q} f(q') P(q') (-1)^{k_0(q,q')+k_1(q,q')} e_0^{k_0(q,q')} e_1^{k_1(q,q')} \\
&\quad \times (1-e_0)^{n_0(q')-k_0(q,q')}(1-e_1)^{n_1(q')-k_1(q,q')} \\
=& \sum_q P(q) f(q) (1-e_0)^{n_0(q)}(1-e_1)^{n_1(q)} \\
&+ \sum_{q'} \sum_{k_0=0}^{n_0(q')} \sum_{k_1=0}^{n_1(q')} \begin{pmatrix} n_0(q') \\ k_0 \end{pmatrix} \begin{pmatrix} n_1(q') \\ k_1 \end{pmatrix} f(q') P(q') e_0^{k_0} e_1^{k_1} \\
&\quad \times (1-e_0)^{n_0(q')-k_0}(1-e_1)^{n_1(q')-k_1}(-1)^{k_0+k_1} \mathbb{1}_{k_0+k_1>0} \\
=& \sum_{q'} \sum_{k_0=0}^{n_0(q')} \sum_{k_1=0}^{n_1(q')} \begin{pmatrix} n_0(q') \\ k_0 \end{pmatrix} \begin{pmatrix} n_1(q') \\ k_1 \end{pmatrix} f(q') P(q') e_0^{k_0} e_1^{k_1} \\
&\quad \times (1-e_0)^{n_0(q')-k_0}(1-e_1)^{n_1(q')-k_1}(-1)^{k_0+k_1} \\
=& \sum_q P(q) f(q) (1-2e_0)^{n_0(q)}(1-2e_1)^{n_1(q)},
\end{aligned} \tag{A.13}$$

which is Eq. A.4.

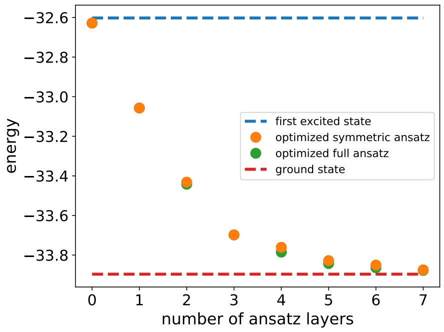


Figure A.3: Ansatz performance in representing ground states of the mixed-field Ising Hamiltonian on 20 qubits, with $h_x = 1.5$ and $h_z = 0.1$. The symmetric ansatz is able to represent the ground states nearly as well as the full ansatz with the same number of layers. Further, relatively few layers are needed to capture the ground state. However, we experiment with many more layers in Chapter 2 and this appendix so as to benchmark error mitigation techniques.

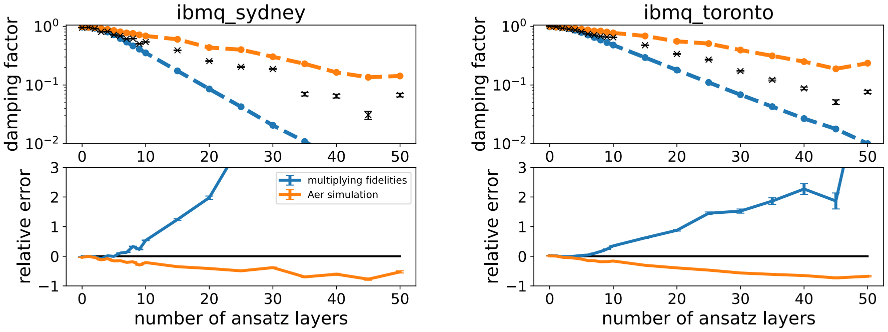


Figure A.4: The measured damping factor (black “x”s) versus the number of ansatz layers, for classically optimized circuits that are then run on *ibmq_sydney* or *ibmq_toronto*. The circuits are optimized to the ground state of the 12-qubit mixed-field Ising Hamiltonian with $h_x = 1.5$ and $h_z = 0.1$. The blue line shows the damping factor that is predicted from multiplying the fidelities of the gates in the backwards light cone of the measured qubits. The orange line shows the prediction from a Qiskit Aer noise model with local depolarizing errors, thermal relaxation errors, and readout errors. In both cases, the error rates are obtained from the device’s calibration. For deep circuits, multiplying the gate fidelities consistently predicts too much suppression, whereas the Qisit Aer model with local gate errors predicts too little suppression.

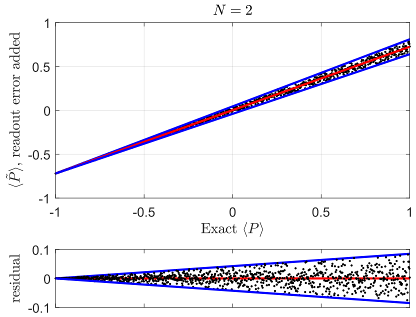


Figure A.5: In this Figure and in Fig. A.6, we generate random probability distributions $f(q)$ that give the desired expected values of bit parity $\langle P \rangle$, which are equally spaced between -1 and 1. $\langle \tilde{P} \rangle$ is then calculated using Eq. A.4. The resulting data are compared to Eq. A.9, and the residuals $\hat{R}$ are plotted. In this Figure, two qubits are measured, and the effect of the readout error is bounded by Eq. A.8 (blue lines) and approximated by A.9 (red line). We set $e_1 = 0.1$ and $e_0 = 0.05$.

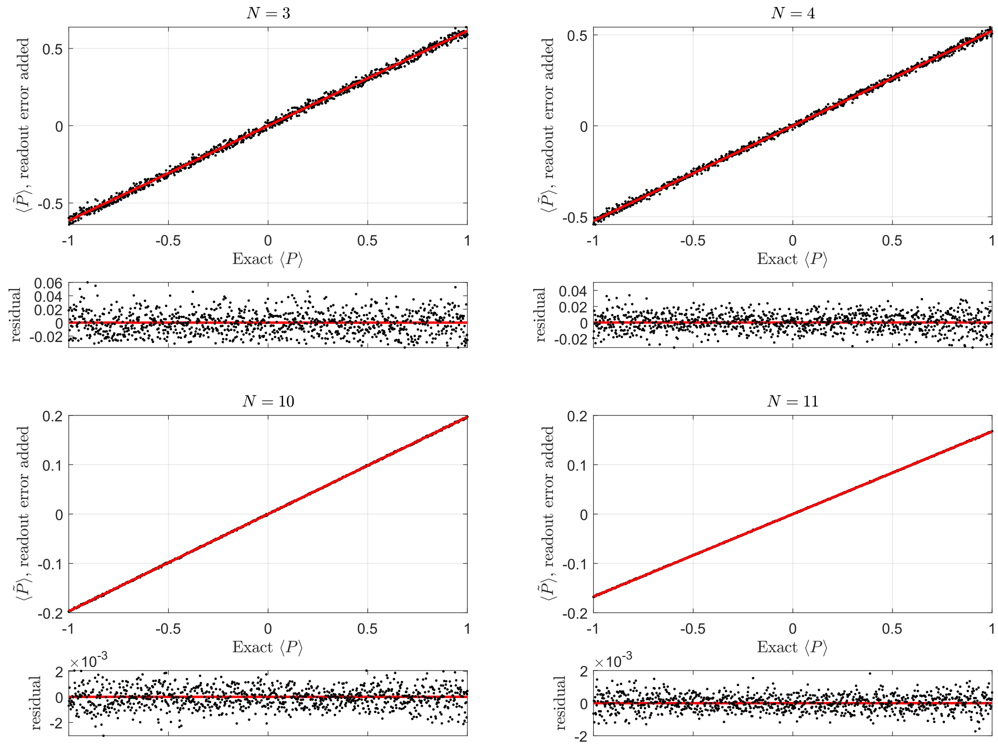


Figure A.6: Same as Fig. A.5 but reading out 3, 4, 10, and 11 qubits, respectively. Eq. A.10 (red line) is a good approximation and becomes better as the number of measured qubits increases.

APPENDIX B

APPENDIX TO CHAPTER 4: SYK

B.1 VQE for SYK

Some states are easier to prepare than others. For example, an infinite-temperature initial state can be implemented by randomly sampling initial bitstrings, requiring only a single layer of X gates. However, the duality between SYK and Jackiw-Teitelboim (JT) gravity emerges at low temperatures. A low-temperature initial state might require a variational technique similar to VQE [114]. Other proposals for generating thermal states include Refs. [35, 181, 8, 180, 9, 172, 158, 73, 37]. The dissipative cooling protocol studied in Chapter 3 may also be useful for generating thermal states. As a proxy for the difficulty for generating a low-temperature thermal state, this appendix considers the difficulty of generating an approximation to the ground state.

The difficulty of generating the ground state is related to the difficulty of performing time evolution. Adiabatic state preparation, dissipative cooling, and even phase estimation all involve implementing Trotterized time evolution with the target Hamiltonian, which is difficult for SYK, $O\left(N^{7/2}t + N^{5/2}t\,\mathrm{polylog}(N/\epsilon)\right)$, before adding the cost of transpiling to the connectivity of the device, which might add another factor of N on a 1D device or $\sqrt{N}$ on a 2D device [7]; for comparison the 1D spin chains studied elsewhere in this thesis only have a cost of $O(Nt)$ in the number of gates and $O(t)$ in depth. VQE is motivated by adiabatic state preparation, so if the latter is difficult, the former is also likely (but not guaranteed) to be difficult. In this appendix, I consider some of the particular difficulties involved in performing VQE for SYK.

The first thing to note is that the SYK Hamiltonian (Eq. 4.1) has $\begin{pmatrix} N \\ q \end{pmatrix} \sim N^4$ terms, which is significantly more than the 1D spin chains studied in Chapters 2, 3, and 5, which have $O(N)$ terms. This affects the number of measurements that must be performed in VQE, as well as the number of rotations that must be performed in order to implement a Trotter step. Although this isn't how I implement it in Chapter 2, VQE for the 1D TFIM, can be performed with only two measurement configurations per energy measurement (measure all of the qubits in the X basis and in the Z basis). In contrast, because Majorana fermions on different sites do not commute with each other, many of the terms in Eq. 4.1 cannot be measured simultaneously. Terms that overlap by an even number of Majoranas commute, whereas those that overlap by an odd number do not. I want to divide these terms into the smallest number of bins such that all of the terms within a bin mutually commute. This will tell us the minimum number of circuits per VQE iteration.

This appendix presents a method for grouping the Hamiltonian terms. The first step is to write the terms as Pauli operators. This can be done using a Jordan-Wigner transformation (also used in Chapter 3) or a Bravyi-Kitaev transformation [17], for example. Python code is available at https://github.com/eliottrosenberg/correlated_SPAM/blob/main/pauli_algebra.py, which conveniently implements these transformations. Example usage is as follows

```
from pauli_algebra import *

whichMajorana = 5

N = 10

encoding= 'jordan_wigner'

pauli_op, coeff = majorana(whichMajorana, N, encoding)

print(pauli_op)

print(coef)
```

which outputs

```
[3 3 2 0 0]

1
```

This means that $\chi_5 = Y_2 Z_1 Z_0$. (The convention in the Jordan-Wigner transformation used here is that even Majoranas begin with an X and odd Majoranas begin with a Y, then they have trailing Zs down to Z_0.)

One can also input a product of Majorana fermions. For example, to write the product $\chi_5 \chi_8$ in terms of spin operators, you can do

```
whichMajorana = [8,5]

N = 10
```

```
encoding= 'jordan_wigner'
pauli_op, coeff = majorana(whichMajorana, N, encoding)
print(pauli_op)
print(coef)
```

which outputs

```
[0 0 1 3 1]
1j
```

This means that $\chi_5 \chi_8 = iX_2 Z_3 X_4$. The convention is that in a product of operators, the rightmost one is the 0th one, which is the 0th element of the list (which confusingly appears on the left when printed). The output `pauli_op` indicates, for each of the $N/2$ sites, which type of Pauli operator acts on that site (0 for $\mathbb{I}$, 1 for X, 2 for Y, and 3 for Z). The other output, `coef`, keeps track of the coefficient on the output Pauli operator, which will be a fourth root of unity.

The same functions work with the Bravyi-Kitaev encoding. For example,

```
pauli_op, coef = majorana(5,10,'bravyi_kitaev')
print(pauli_op)
print(coef)
```

outputs

```
[0 3 2 1 0]
1
```

meaning that, in the Bravyi-Kitaev encoding with N Majorana fermions, $\chi_5 = Z_1 Y_2 X_3$. You can input a product of Majorana fermions as well. For example, `pauli_op, coef = majorana([8,5],10,'bravyi_kitaev')` returns `pauli_op = [0 3 2 2 1]` and `coef = -1j`, meaning that $\chi_5 \chi_8 = -i Z_1 Y_2 Y_3 X_4$.

This tool makes it simple to convert the SYK Hamiltonian into a spin Hamiltonian. In order to understand the difficulty of measuring $\langle H_{\text{SYK}} \rangle$ on a quantum processor, it is important to quantify the sizes of the Hamiltonian terms, i.e. the number of sites on which they act nontrivially. Figure B.1 shows the distribution of sizes these terms after converting the fermions to spin operators using either a Jordan-Wigner or Bravyi-Kitaev transformation. The Bravyi-Kitaev transformation leads to significantly smaller terms, a clear advantage for implementation on NISQ devices. However, these terms still act on about 10-12 sites, which is a significant challenge compared to the spin Hamiltonians studied elsewhere in this thesis, the terms of which only act on at most two sites.

In order to perform VQE, expectation values of these terms must be measured. The average effect of readout errors on these measurements is given by Eq. A.10. In particular if $e_0 + e_1 \approx 2\%$ and P contains 12 terms, I expect the expectation value of the operator to be damped to about 78% of its ideal value due to readout errors. This is a significant effect that must be compensated for using readout error mitigation, as detailed in the main text, Chapter 4.

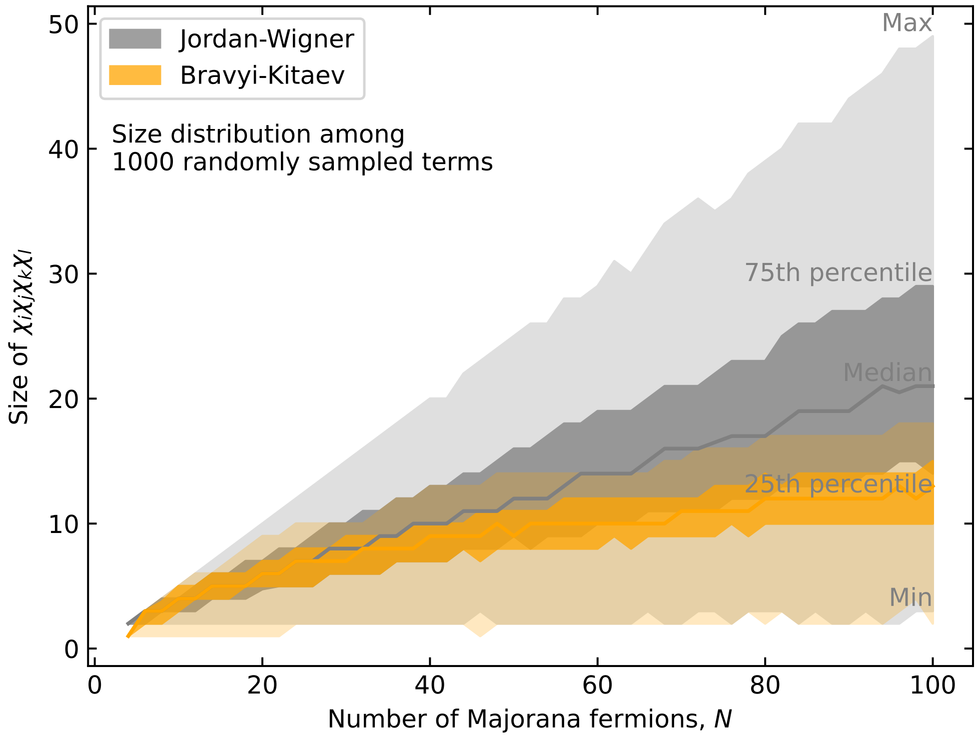


Figure B.1: The number of sites on which terms in the SYK Hamiltonian act nontrivially after qubitization via either a Jordan-Wigner or Bravyi-Kitaev transformation, as a function of system size.

Next, I return to the issue of grouping the Hamiltonian terms. Code for performing this grouping is available on Github[1]. The strategy is as follows:

1. I start with a list, `remaining`, of the Pauli operators that I have yet to measure. Sort this list by size and pick the largest Pauli operator, `pauli_to_measure`; I will measure in a basis that lets us measure this operator. However,

---

[1] https://github.com/eliottrosenberg/correlated_SPAM/blob/main/SYK/generate_SYK_Hamiltonian.m (written in MATLAB, sorry; blame Stuart Geman for teaching me to use MATLAB in undergrad)

`pauli_to_measure` may contain identity operators on some sites, and I will want to fill those in with either X, Y or Z so that I measure as many terms simultaneously as I can. To do this I:

2. first remove `pauli_to_measure` from `remaining`, as well as any other Pauli operators that I have already covered (i.e. those that commute with `pauli_to_measure` and do not act on any sites that `pauli_to_measure` does not act on). Next,
3. if the number of bases to assign, `num_qubits_to_assign`, is small (at most 9):
    (a) search through all of the `3**num_qubits_to_assign` possible assignments, and select the one that maximizes the number of terms in `remaining` that are measured (this is implemented by the function `find_optimal_assignment`). Otherwise:
    (b) do the following, implemented by `find_good_assignment`: of the terms in `remaining` that commute with `pauli_to_measure`, find the one that fills in the maximal number of unassigned bases (i.e. fills in the maximum number of identity factors in `pauli_to_measure`). Fill in `pauli_to_measure` accordingly and remove any terms that are now measured by this choice from `remaining`. Repeat until there is no change in `pauli_to_measure`.
4. The basis in which to measure, `pauli_to_measure` is now fully constructed. Add it to the list of bases in which to measure.
5. Repeat steps 1–4 until `remaining` is empty.

The results of this grouping algorithm are shown in Figure B.2. Evidently, the number of measurement configurations required (as found by this algorithm) depends on the mapping from fermions to qubits; somewhat fewer measurement configurations are required for the Jordan-Wigner representation than for the Bravyi-Kitaev representation. Of course, the commutation relations among the Hamiltonian terms are independent of the qubitization scheme used, but because I am measuring each qubit in a local Pauli basis, the representation does matter (e.g. Z_1Z_0 and X_1X_0 commute, but I would still measure them separately).

Unfortunately, I find that the fractional benefit from grouping decreases as the system size increases. Further, the reduction in the number of measurements is only about a factor of 2, not enough to make the $\begin{pmatrix} N \\ 4 \end{pmatrix}$ terms per VQE iteration viable. A possible workaround is to study sparse SYK instead of SYK [176, 25], which is the solution implemented by [78].

So far, I have only considered the challenges associated with performing the energy measurements. A separate challenge is whether a shallow ansatz, suitable for NISQ processors, can approximate the ground state. To assess this, I picked a single instance of the SYK Hamiltonian with 24 Majorana fermions (i.e. a particular choice for the random coefficients in Eq. 4.1). The Hamiltonian was not sparsified; all $\begin{pmatrix} 24 \\ 4 \end{pmatrix} = 10,626$ terms were included. I used exact diagonalization to find the ground state energy and then classically optimized ansatz circuits to try to minimize the energy, $\langle\psi(\vec{\theta})|H|\psi(\vec{\theta})\rangle$. The ansatz and code used to optimize it were the same as

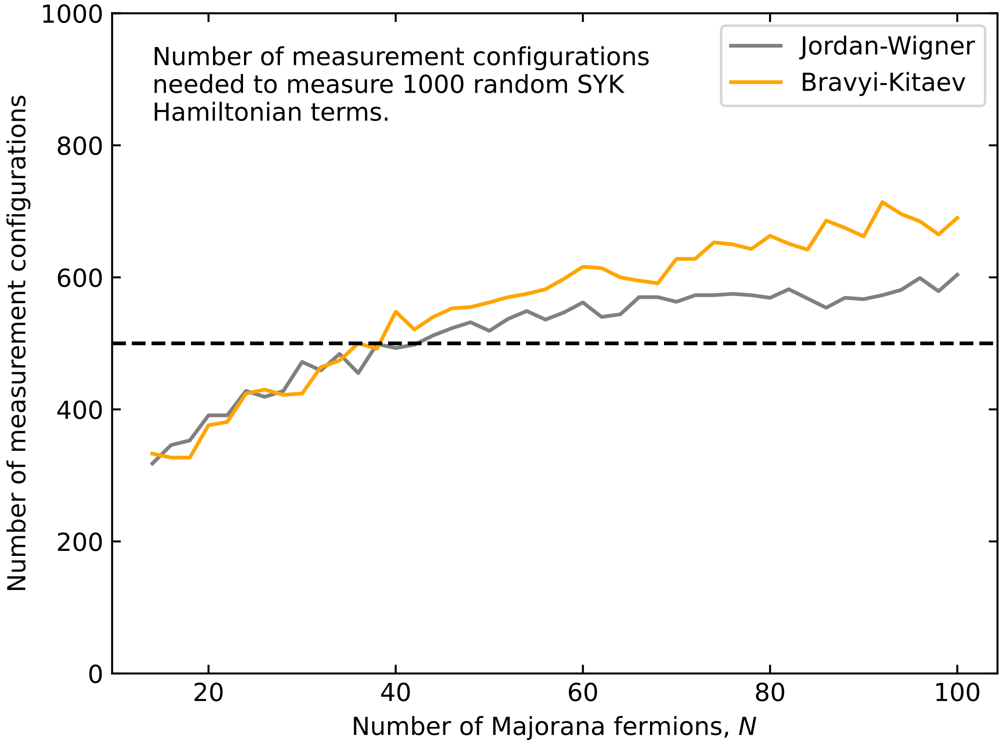


Figure B.2: The number of measurement configurations required to measure 1000 random SYK terms.

used in Chapters 2 and 3; the ansatz is depicted in Figure A.2. This was repeated for both the Jordan-Wigner and Bravyi-Kitaev encodings of the fermions into qubits. The results are shown in Figure B.3. Unfortunately, after 19 ansatz layers, the energy has only gotten about halfway to the true ground state energy (specifically, $E/E_{\rm gs} = 0.504$ for the Bravyi-Kitaev encoding and $E/E_{\rm gs} = 0.505$ for the Jordan-Wigner encoding at 19 ansatz layers). The poor performance compared to what was achieved for 1D Ising models in Chapters 2 and 3 ($E/E_{\rm gs} \approx 1$), is perhaps not surprising; as suggested earlier in this appendix, we might expect the performance of VQE to be related to the performance of adiabatic state preparation, and adiabatic state preparation certainly requires a much deeper circuit for SYK than it does for the Ising models, given the cost of time evolution that was discussed earlier in this chapter.

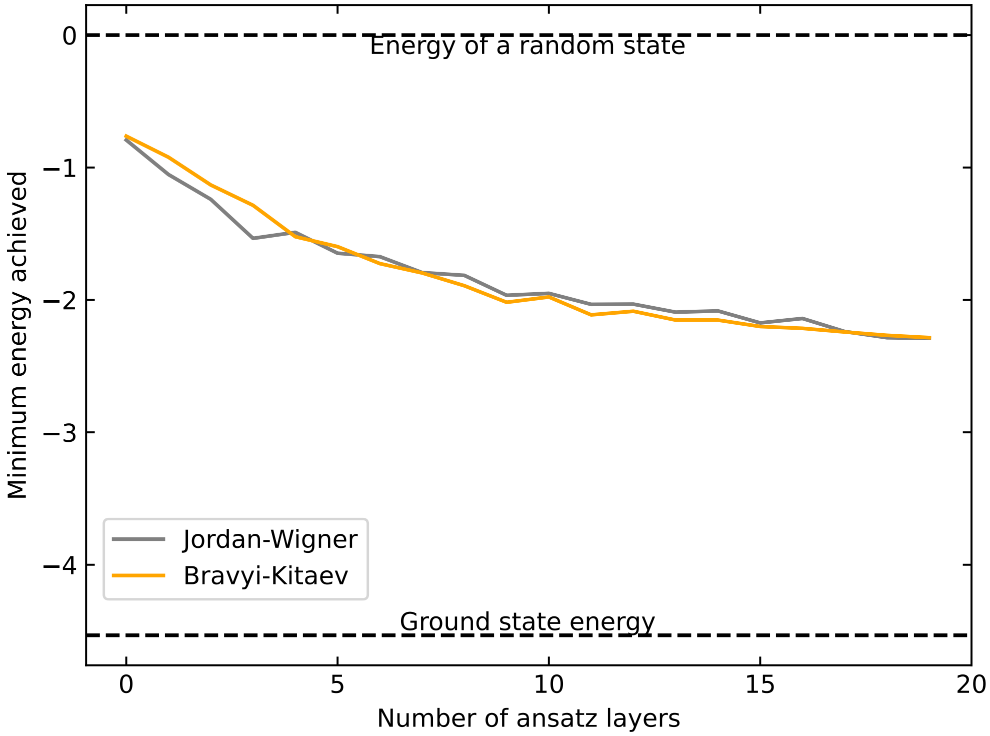


Figure B.3: The performance of classically-optimized VQE for an instance of SYK with 24 Majorana fermions. Repeated for both Jordan-Wigner and Bravyi-Kitaev qubitizations. The true ground state energy is also shown, which is computed using exact diagonalization.

As we have seen, SYK poses many challenges that would need to be overcome in order for a nontrivially-sized system to be simulated on an existing or near-term quantum processor. It is possible that by sparsifying the Hamiltonian sufficiently, these challenges could be rendered surmountable. Regardless of the level of sparsification, one challenge that will remain for a VQE implementation is the need to accurately measure the expectation values of large Pauli operators, which I address in the main text, Chapter 4.

APPENDIX C

APPENDIX TO CHAPTER 5

C.1 Experimental techniques and device characterization

C.1.1 Overview

The experiments are performed using 46 frequency-tunable superconducting transmon qubits. The qubits are prepared in a random bitstring state according to the probabilities set by the initial imbalance μ: qubits on the left side of the chain are prepared in $|1\rangle$ with probability $p = e^{\mu}/(e^{\mu} + e^{-\mu})$, otherwise $|0\rangle$, and qubits on the right are prepared in $|0\rangle$ with probability p, otherwise $|1\rangle$. The system is then evolved with alternating layers of fSim gates [57, 127, 119], which implment a Floquet XXZ model. Finally, all 46 qubits are measured in the computational basis. Because ideal fSim gates are number-conserving, we post-select on the measured bitstrings having the correct number of 1s, effectively mitigating against photon loss, which otherwise causes the number of 1s to decay. After sampling over N_{states} initial bitstring states, we compute the expectation value of the kth power of the transferred magnetization as

$$\langle \mathcal{M}(t)^k \rangle = \frac{1}{N_{\text{states}}} \sum_i \frac{1}{N_i^{\text{counts}}(t)} \sum_j N_{ij}^{\text{counts}}(t) \left(2\left(N_1^R(j) - N_1^R(i)\right)\right)^k, \qquad \text{(C.1)}$$

where i is the initial bitstring and j is the measured bitstring, $N_i^{\text{counts}}(t)$ is the total number of counts that survive postselection after t cycles when the initial state is i, and $N_{ij}^{\text{counts}}(t)$ is the number of times the bitstring j is measured after t cycles when the initial state is i. $N_i^{\text{counts}}(t) = \sum_j N_{ij}^{\text{counts}}(t)$. $N_1^R(i)$ is the number of 1s in the right half of the binary representation of i. Moments are computed as

$$\begin{aligned} \alpha_k(t) &= \left\langle \left(\mathcal{M}(t) - \langle \mathcal{M}(t) \rangle\right)^k \right\rangle \\ &= \sum_{i=0}^{k} \begin{pmatrix} k \\ i \end{pmatrix} \left\langle \mathcal{M}(t)^{k-i} \right\rangle \left(- \langle \mathcal{M}(t) \rangle\right)^i , \end{aligned} \tag{C.2}$$

where the second line is written in terms of the experimentally measured quantities, Eq. (C.1). Finally, the skewness $\mathcal{S}(t)$ and kurtosis $\mathcal{Q}(t)$ are computed as

$$\begin{aligned} \mathcal{S}(t) &= \frac{\alpha_3(t)}{\alpha_2(t)^{3/2}} \\ \mathcal{Q}(t) &= \frac{\alpha_4(t)}{\alpha_2(t)^2} - 3. \end{aligned} \tag{C.3}$$

Statistical uncertainties of each of these quantities are computed using the remove-one jackknife method, wherein one initial state is removed from the sample, and the variation of the quantity of interest (e.g. of $\langle \mathcal{M}(t) \rangle$, $\alpha_2(t)$, $\mathcal{S}(t)$, or $\mathcal{Q}(t)$), depending on which state is removed, is used to estimate the statistical uncertainty of that quantity. The jackknife method is also used to estimate bias, which is found to be negligible compared to the statistical uncertainties. The jackknife method is described in more detail in Section C.1.4.

In addition to post-selecting on number conservation, we apply several additional error avoidance and mitigation techniques:

(1) We post-select on the causal possibility of the observed bitstring. For example, if an initial bitstring on 8 qubits is 11011000, then it is not possible, in a noiseless system, to observe 01011001 at cycle 1. The rightmost 1 appeared acausally (likely by readout error) and hence the bitstring is filtered out even though it contains the correct number of 1s. We have an efficient algorithm (described in Section C.1.2 for checking whether a given observed bitstring is causally possible after t cycles from a given initial bitstring. This filtering mostly affects the earliest few cycles, for which the number of causally connected bitstrings is small. This filtering is a small effect compared to the number-conserving post-selection, which keeps an exponentially decaying number of bitstrings as a function of cycle number (see Figure C.1A).

(2) Noting that the effects of amplitude damping (T_1) on our experiment are worse in initial bistring states with more 1s, when an initial state is more than half-full, i.e. the number of 1s is greater than 23, we relabel the $|0\rangle$ and $|1\rangle$ states, i.e. we start in the initial bitstring $\bar{b}_i$ instead of b_i and then replace each measured bitstring b_j with $\bar{b}_j$, where $\bar{b}$ means to apply a NOT operator to all of the bits in b. The advantage gained from this technique is illustrated in Figure C.1C.

Figure C.1A shows exponential decay of the fraction of counts that survive the post-selection. We call the decay constant the algorithmic relaxation time, T_1^A, which, as illustrated in Figure C.1C, is about 3 cycles at half filling. If one naively estimates T_1^A at half-filling from

$$e^{-t/T_1^A} \overset{?}{=} \prod_{i=1}^{23} e^{-t/T_1^{(i)}} \Longrightarrow \frac{1}{T_1^A} \overset{?}{=} \sum_{i=1}^{23} \frac{1}{T_1^{(i)}}, \tag{C.4}$$

where $T_1^{(i)}$ is the T_1 of qubit i, measured at its idle frequency, we obtain an estimate for T_1^A of over 7 cycles, even if we pick sum to be over the 23 qubits with the shortest T_1 out of the 46 total.

There are two main mechanisms expected to cause discrepancies between the algorithmic T1 and the estimate based on single-qubit T1 values. First, when the coupling is turned on, the coupler is brought close to the qubits in frequency, allowing noise in the coupler to affect the qubit. This can also enable noise-induced transitions from the qubit to the coupler. Second, the relevant T_1 for the experiment is not the T_1 at the idle frequency, even though that is what is typically optimized for and reported. The fSim gates are implemented as in Refs. [119, 74]; pairs of qubits are tuned to their interaction frequencies in a trapezoidal coupler pulse, the amplitude and duration of which are tuned to obtain the desired SWAP and controlled-phase angles. The resulting fSim gate includes single-qubit phases, which must be calibrated to zero by applying physical Z rotations [119]. Physical Z gates are fixed-duration 10-ns gates in which the qubit frequency is detuned from the idle frequency f_0 to the frequency f_z. In the frame rotating at the idle frequency, the qubit accumulates a phase of $2\pi(f_z - f_0) \times 10\,\mathrm{ns}$. Therefore, the full range of phases from $-\pi$ to π can be obtained by the range of frequencies $|f_z - f_0| \leq 0.05\,\mathrm{GHz}$. Figure C.2A shows T_1 as a function of frequency for a typical qubit in our chain, indicating the idle frequency, the interaction frequencies with the two neighboring qubits, and the range of frequencies used for physical-Z rotations. It is readily seen that, although the idle frequency may be optimized to give a long T_1, other frequencies used during the circuit execution have T_1s that can be about a factor of 2 shorter. Figure C.2B shows how the frequencies for each of the 46 qubits in our chain vary over the course of a cycle. Clearly, T_1 during the circuit execution differs from T_1 at the idle frequency, and the factor-of-two difference between the measured and predicted algorithmic relaxation time is plausibly explained.

Post selection largely mitigates against amplitude damping, characterized by the algorithmic relaxation time, at the cost of an exponential overhead in the number of shots required. However, some errors make it past the post selection. In particular, although amplitude damping causes the number of 1s to decrease, $0 \to 1$ readout errors cause it to increase. Therefore, when both amplitude damping and $0 \to 1$ readout error occur, the measured bitstrings can have the correct number of 1s and pass the post-selection. As evident in Figure C.1A, by later cycles, the vast majority of bitstrings have had some amplitude damping, and on 46 qubits, it is likely that at least one $0 \to 1$ readout error will occur (typical readout error rates are shown in Figure C.3), so this is a non-negligable effect. It manifests as excitations appearing to jump nonlocally along the chain, moving from the side with high concentration to the side with low concentration faster than they would without noise. As described in Section C.2.3, we perform simulations including this effect and find that it explains most of the discrepancy between the noiseless simulation and the experiment. Evidently, the quality of our post-selected experimental results could be improved by changing how the readout calibration is done; the readout centers could be chosen to decrease the $0 \to 1$ error rate at the expense of the $1 \to 0$ error rate (and hence requiring more shots). We leave this modified calibration technique for future work.

There are other sources of errors that are not mitigated by post-selection. These include dephasing, leakage (occupation of the $|2\rangle$ state), and control errors (fSim angle miscalibrations). Because the coupling strengths used here are not particularly high, leakage is not expected to be a dominant source of error. We characterize single- and two-qubit dephasing, as well as control errors, and include these effects in our simulations. However, we find that most of the observed discrepancy is explained by T_1 and readout errors alone (Figure C.5).

The data included in Chapter 5 were collected over the course of several months, on different sets of qubits and two different devices. In order to ensure consistent data quality, readout error rates, single-qubit error rates, and two-qubit cross-entropy benchmarking (XEB) fidelities were measured periodically, typically after every 10 initial states; we kept only data for which the maximum error rates were below thresholds that we set.

C.1.2 Causal filter

Here we present the causal filter with which we post-select our measured bitstrings. We illustrate the algorithm with the two bitstrings mentioned in Section C.1.1. Suppose that the initial bitstring is 11011000 and we want to know whether 01011001 is a possible measurement outcome after one cycle.

To answer this question, we determine the minimum number of cycles required to obtain 01011001 in the ideal dynamics. First, we assign identities to the "1"s in the initial and final bitstring determined by their order. This is illustrated in Figure C.4, where we assign colors to the "1"s in both the initial bitstring (on the left) and the final bitstring (on the right). We then consider each layer of fSim gates and move the excitations if it is allowed by the gates and if doing so brings the excitations closer to their desired locations. For example, in the first layer, the blue and orange excitations are blocking each other from moving (since color is determined by order). The green one could move up, but doing so would bring it further from its final position, so it does not move. The red excitation moves down because that brings it closer to its final position. From the figure, it is clear that at least 1.5 cycles are needed to obtain 01011001 from 11011000. Therefore, this bitstring would be filtered out if seen after only one cycle.

C.1.3 Gate calibration

We implement fSim gates using the same trapezoidal coupler pulses used in previous works [127, 119, 74]. Here, we note some differences between how gates are calibrated here versus in previous works. (1) in Ref. [119], the fSim gates were not of uniform duration across pairs. Here, we adjust the padding (the idle time in Figure C.2B) so that the gate duration (including the added padding) is uniform across pairs, leading to the neat alignment in time acros qubits shown in Figure C.2B. (2) In Ref. [119], Floquet calibration of the fSim angles allowed them to be controlled with high precision. Because control errors (called disorder in Figure C.5) are a negligible source of error for us, we instead use unitary tomography to calibrate the fSim angles. This enables us to calibrate gates quickly and in a way that is mostly automated. We tried Floquet calibration, which allows for more precise calibrations of the angles at the cost of a higher overhead, but found that it did not improve our gate fidelities. (3) We iteratively calibrate the hold time T and coupling strength g_{max} of the trapezoidal pulse by measuring the fSim angles θ and ϕ at points in a small cross shape in the (T, g_{max}) plane centered at the previous guess, fitting the polynomials

$$\begin{aligned} \theta &= f((b_1 g_{\mathrm{max}} + b_0)(T + T_b)) \\ \phi &= (c_1 g_{\mathrm{max}} + c_0)(T + T_c), \end{aligned} \tag{C.5}$$

where f is the triangle-wave function illustrated in Figure C.6. This technique allows for fast calibration of fSim gates without relying on expensive 2d sweeps and in a way that is more robust to noise than gradient descent. Figure C.7 illustrates this calibration procedure.

An advantage of being able to quickly optimize T and $g_{\max}$ for the desired fSim angles is that we can now put this in an outer loop that adjusts the interaction frequencies. Indeed, during gate calibration, we set a minimum two-qubit XEB fidelity and optimize the interaction frequencies (re-optimizing T and $g_{\max}$ each time) until all qubit pairs achieve the desired fidelity.

C.1.4 Jackknife estimate of uncertainties

We use the "delete one" jackknife to estimate the statistical uncertainty of our skewness and kurtosis, and dynamical exponent measurements [141, 162, 47]. For finite m, where we average over initial states, define $\hat{\theta}_{(i)}$ to be the quantity of interest, for example the skewness or kurtosis, computed with initial state i removed. Define $\hat{\theta}_{(.)} = \frac{1}{N_s}\sum_i \hat{\theta}_{(i)}$, where N_s is the number of initial states. Then we use

$$\sigma_{\hat{\theta}} = \sqrt{\frac{N_s - 1}{N_s}\sum_i \left(\hat{\theta}_{(i)} - \hat{\theta}_{(.)}\right)^2} \tag{C.6}$$

as our estimate of the statistical uncertainty of the quantity $\hat{\theta}$.

The averages of skewness or skewness over cycle number are uncertainty-weighted averages, and the uncertainty of the average is computed directly using Eq. (C.6).

In the case of $m = \infty$, there is only one initial state, so we instead perform the jackknife estimate by deleting each one of the shots. This gives us the uncertainty of the skewness and kurtosis at each cycle. Because the shots, unlike the initial states, are different at each cycle number, we cannot use the jackknife to directly compute the uncertainty of the cycle-averaged skewness and kurtosis for $m = \infty$. Instead, in this case, we treat the skewness and kurtosis at each cycle as independent random variables, so the uncertainty of their weighted average is

$$\sigma_{\text{weighted avg}} = \frac{1}{\sqrt{\sum_t w_t}}, \tag{C.7}$$

where the weight, w_t, at cycle t is $1/\sigma_t^2$, where σ_t is the uncertainty of the quantity (either skewness or kurtosis) at cycle t.

C.2 Simulation techniques and numerical results

Our main quantity of interest, the transferred magnetization $\mathcal{M}$, counts the number of excitations that have moved across the center of the 1D chain. A significant simplification stems from the fact that, at early times, excitations far from the center have not had time to cross the center of the chain. As a result, we can imagine an infinitely large system and only simulate a finite number of sites in order to study it. In particular, at cycle t, it is only necessary to simulate $2t$ sites [1]. As a result, the optimal simulation technique varies depending on the cycle number. Through cycle 8 (16 qubits), we obtain exact results by simulating the full density matrix. Beyond that point, density matrix simulations become costly, so for cycles 9–18 (18–36 qubits), we instead sample random initial bitstring states, as done in the experiment, and apply exact statevector simulation to these initial states. For the pure domain wall case ($\mu = \infty$), we employ tensor-network simulations using the time-evolving block decimation (TEBD) algorithm [166] to extend the simulations to cycle 23.

---

[1] For a particular initial bitstring state, the transferred magnetization is only independent of system size when $N_Q \geq 4t - 2$, as one would expect from the causality structure of the circuit, but after averaging over initial bitstrings (or equivalently in the mixed initial state), we find the transferred magnetization is exactly independent of system size as long as $N_Q \geq 2t$ sites.

C.2.1 Analytical results

For very early times, the matrices involved are small enough that it is possible to obtain relatively simple analytical results. Some of these are tabulated here:

Cycle 1

To compute the transferred magnetization at cycle 1, it is only necessary to consider two qubits with an fSim gate between them. Therefore, the probability distribution of the transferred magnetization takes a simple form. For positive integer power k, we have:

$$\langle \hat{\mathcal{M}}^k \rangle = \begin{cases} 2^k \sin^2\theta \tanh\mu, & k \text{ odd} \\ 2^{k-1} \sin^2\theta \left(1 + \tanh^2\mu\right), & k \text{ even} \end{cases} . \tag{C.8}$$

In particular, the mean, variance, and skewness are

$$\begin{aligned} \langle \hat{\mathcal{M}} \rangle &= 2\sin^2\theta \tanh\mu \\ \text{var} &= 2\sin^2\theta \left(1 + \cos(2\theta)\tanh^2\mu\right) \\ \mathcal{S} &= A/B, \quad \text{where} \\ A &= 2\sqrt{2}\Big(\left(2\sin^4\theta \tanh^2\mu + 1\right) \left(\sinh(2\mu) + \cosh(2\mu) + 1\right)^2 \\ &\qquad - 3\left(\sinh(4\mu) + \cosh(4\mu) + 1\right)\sin^2(\theta)\Big) \tanh\mu \\ B &= \left(\sinh(2\mu) + \cosh(2\mu) + 1\right)^2 \sqrt{\left(\cos(2\theta)\tanh^2\mu + 1\right)^3} \sin\theta, \end{aligned} \tag{C.9}$$

and, to leading order in μ, we have

$$\begin{aligned} \mathcal{S} &= \mu\sqrt{2}(2\csc\theta - 3\sin\theta) + O(\mu^3) \\ \mathcal{Q} &= 2\csc^2\theta - 3 + O(\mu^2). \end{aligned} \tag{C.10}$$

Observe that, for $\mu \ll 1$, the skewness and kurtosis are both positive for small θ, i.e. the Trotter limit, and negative for large θ. The crossover happens at $\theta = \arcsin\left(\sqrt{2/3}\right) \approx 0.3\pi$. In Figure 3 of the main text, we choose $\theta = 0.4\pi$, which is why we observe negative skewness and kurtosis. In continuous-time Hamiltonian dynamics, we expect the opposite signs.

Cycle 2

It is also possible to obtain analycial expressions at cycle 2. In particular, the mean and variance of the transferred magnetization are

$$\begin{aligned} \langle\hat{\mathcal{M}}\rangle &= 2\mu\sin^2\theta\left(\cos^4\theta(3+\cos\phi) + 2\sin^2\theta\right) + O\left(\mu^3\right) \\ \mathrm{var} &= \sin^4\theta(1-\cos\phi) + \frac{1}{8}(3+\cos\phi)(7\sin^2\theta + \sin^2(3\theta)) + O(\mu^2) \end{aligned} \tag{C.11}$$

C.2.2 Simulation cost and runtime

To perform statevector simulations out to cycle 18 (36 qubits), we use NVIDIA's cuQuantum [55] and its interface with qsim [140]. cuQuantum supports multi-GPU quantum simulations, and with eight 80-GB NVIDIA A100 GPUs, available to virtual machines running in Google Cloud's compute services, we can simulate up to 36 qubits. On this platform, a noiseless 18-cycle simulation takes about 17.6 seconds per initial state. However, the memory required to store the state increases exponentially in the cycle number, as shown in Figure C.8. The cuQuantum implementation stores 2^{2t} complex numbers. The memory footprint could be reduced by taking advantage of number conservation, in which case only $\begin{pmatrix} 2t \\ t \end{pmatrix}$ complex numbers would be needed to represent the state.

C.2.3 Noisy simulations

In this subsection, we describe how we performed the noisy simulations shown in Figure C.5 and Figure 2 of the main text. The simulations with disorder were performed in a straightforward way; we simply measured the actual fSim angles, including the single-qubit phases, using unitary tomography, and used the measured angles in the simulation, simulating only the $2t$ qubits about the center, as in the noiseless statevector simulations. The simulations with dephasing were also performed in a straightforward way; we simply averaged over many circuits, adding high- and low-frequency Gaussian noise to the fSim angles, as well as Z-rotations between the fSim gates with random angles that vary both within a circuit and across shots.

The simulation of amplitude damping and readout error is slightly more involved. For each initial 46-qubit bitstring, we consider separately the $2t$ qubits about the center and the remaining $46 - 2t$ outer qubits. For the center qubits, we perform a noisy simulation using cirq/qsim [140] that includes the measured amplitude damping as gates applied between the layers of fSim gates. The outer qubits are treated as if no two-qubit gates are applied; qubits prepared in $|1\rangle$ are stochastically flipped to $|0\rangle$ with a probability $1 - e^{-t/T_1}$. The resulting bitstrings from the center and outer qubits are concatenated back together, and then bits are randomly flipped according to the $0 \to 1$ readout error rate e_0 and the $1 \to 0$ readout error rate e_1. Finally, the same post-selection that is used in the experiment is applied to the simulated bitstrings, so that only those conserving the number of 1s and satisfying the causality constraints survive. The readout error rates used here are those measured on the device at the time the experiment was run, including the qubit-by-qubit variations. The amplitude damping rate, T_1, is obtained as in Figure C.1A and is approximated as being the same across all qubits.

C.2.4 Length independence

As demonstrated in Table C.1, the transferred magnetization is independent of the length of the chain as long as the chain consists of at least $2t$ qubits, where t is the cycle number.

		Absolute difference from reference kurtosis when computed with N_Q qubits.							
Cycle	Reference kurtosis	$N_Q = 2$	$N_Q = 4$	$N_Q = 6$	$N_Q = 8$	$N_Q = 10$	$N_Q = 12$	$N_Q = 14$	$N_Q = 16$
1	-0.7888543819998315	0	10^{-15}	10^{-15}	10^{-15}	10^{-15}	10^{-14}	10^{-14}	NC
2	-0.3236513411118609		0	10^{-15}	0	10^{-15}	10^{-15}	10^{-15}	NC
3	-0.19032877952363814			0	0	10^{-15}	10^{-15}	10^{-15}	NC
4	-0.13634880999637522				0	10^{-15}	10^{-15}	0	NC
5	-0.11064712866845028					0	10^{-15}	0	NC
6	-0.0964617033065931						0	10^{-15}	NC
7	-0.08726032489412416							0	NC
8	-0.08047534387997635								0

Table C.1: Illustration of the length-independence of the transferred magnetization, as long as $N_Q \geq 2t$. Deviations are consistent with double-precision floating-point arithmetic. The values are computed using exact density matrix simulations for $(\theta, \phi) = (0.4\pi, 0.8\pi)$ and $\mu = 0$, but we observe similar agreement across system sizes for all choices of these parameters and for all of the moments. Here "NC" means "not computed." The kurtosis at $t = 1, N_Q = 2$, is given by Eq. (C.9): $\mathcal{Q} = 2\csc^2\theta - 3$.

C.2.5 Crossing time

In Figure 3 of the main text (panels C and D), we plot the skewness and kurtosis as functions of $\mu t^{2/3}$. The data collapse observed in Fig. 3C is seen even more clearly in the numerics, shown in Figure C.9. In the inset to panel A, we see that the time at which the skewness becomes positive scales like $\mu^{-3/2}$, and, in panel B, we see a collapse of the numerical data when plotted as a function of $\mu t^{2/3}$. The power law scaling sensibly predicts that the crossing time becomes infinite as $\mu \to 0$, which makes sense because the skewness is always 0 at $\mu = 0$. The kurtosis, however, cannot be a function of $\mu t^{2/3}$ because it is not constant when $\mu = 0$.

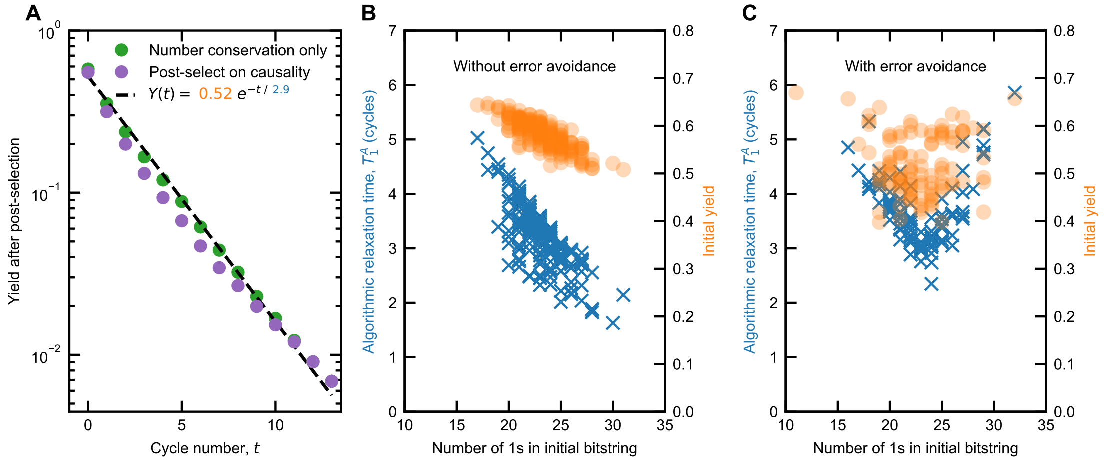


Figure C.1: **A.** The fraction of the total counts that pass the post-selection, as a function of cycle number, for a single initial bitstring. The green markers (higher points) indicate the yield when post-selecting only on number conservation. The purple markers (lower points) indicate the yield when post selecting on whether the observed bitstring is causally possible given the initial bitstring in the ideal circuit, a more stringent criteria than number conservation. Past cycle 11, the causality filter has no effect beyond enforcing number conservation (for this particular bitstring and set of measurements). A decaying exponential is fit to the yield when post-selected on number conservation only. We call the yield at $t = 0$ the initial yield, which is less than one because of readout error. We call the decay constant the algorithmic relaxation time, T_1^A, which is a measure of the effective amplitude damping rate in the experiment. **B.** The initial yield (orange circles) and algorithmic relaxation time (blue ×s) as a function of the number of 1s in the initial bitstring. As expected, the algorithmic relaxation time decreases as the number of 1s increases. The initial yield also decreases with the number of 1s in ithe initial bitstring because $0 \to 1$ readout error rates tend to be lower than $1 \to 0$ readout error rates. **C.** Employing the error avoidance technique in which initial bitstrings b_i with more than $N_Q/2 = 23$ 1s are replaced with $\bar{b}_i$ and the resulting measured bitstrings b_f are replaced with $\bar{b}_f$, so that the number of 1s physically implemented is never more than 23. (Or equivalently, we relabel the $|1\rangle$ and $|0\rangle$ states when the number of 1s is more than 23.) As expected, this results in a V shape for T_1^A as a function of the number of 1s in the initial bitstring, with the minimum at 23 1s, corresponding to half-filling. At cycle 23, the yield at half-filling is about 2×10^{-4}, and we use 6×10^4 shots per initial state, resulting in about 12 shots per initial state after post-selection. We average over approximately 100 initial states for each value of m, θ, and ϕ that we study.

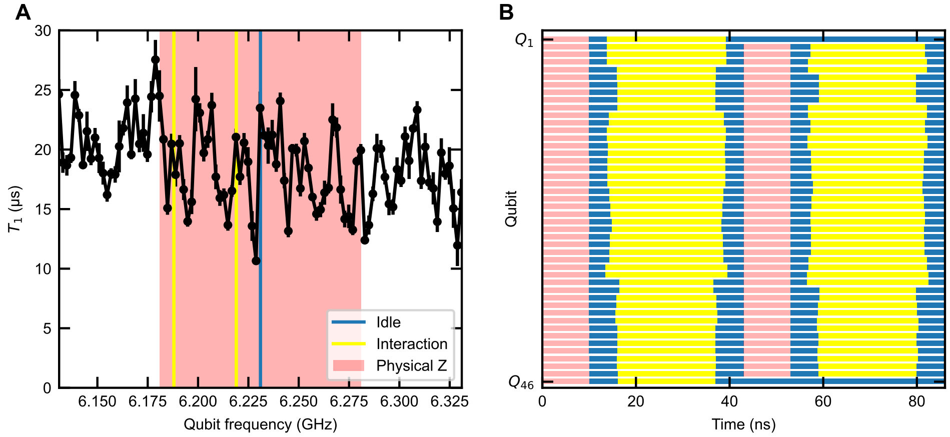


Figure C.2: **A** T_1 as a function of frequency for a typical qubit in our chain of 46 qubits. Vertical lines indicate the idle frequency and the interaction frequencies with each of the two neighboring qubits. The shaded region indicates the range of frequencies used to obtain physical Z rotation angles from $-\pi$ to π. **B** Schematic showing the qubit frequency as a function of time during one cycle for all 46 qubits in a typical calibration used during this experiment.

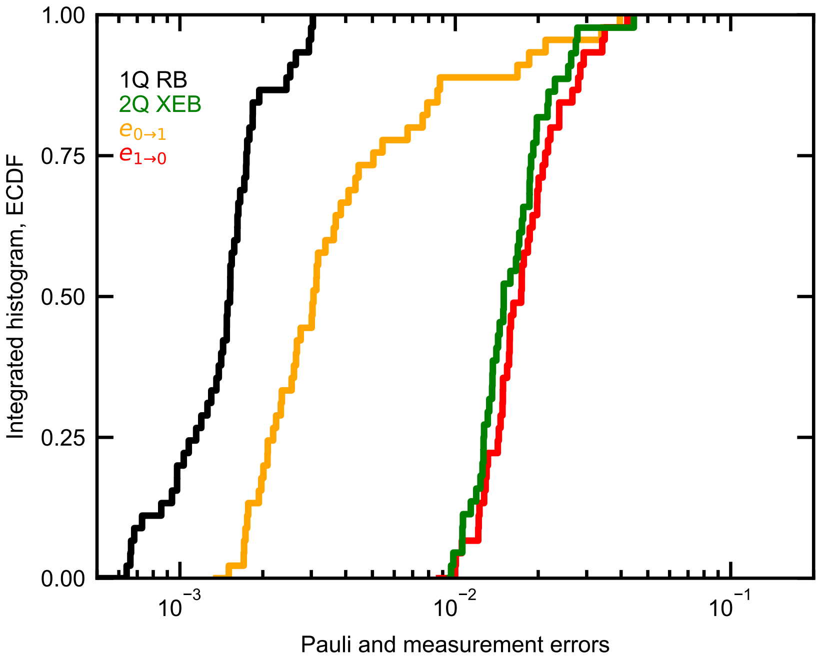


Figure C.3: Typical error rates achieved on our 46-qubit chain, showing readout error, single-qubit randomized benchmarking (RB), and two-qubit cross-entropy-benchmarking (XEB) cycle Pauli errors. The slightly lower error rates reported in Ref. [119] are due to the different sets of angles used.

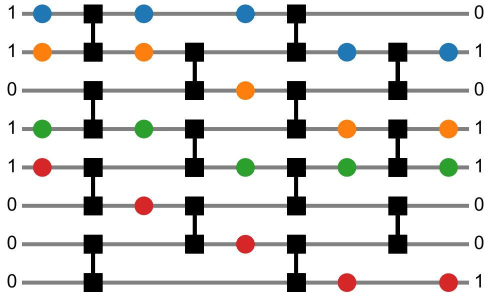


Figure C.4: An illustration of how we determine the minimum number of cycles needed to obtain a given final bitstring from a given initial bitstring. In this case, at least 1.5 cycles are needed to obtain 01011001 from an initial bitstring of 11011000 under the ideal dynamics.

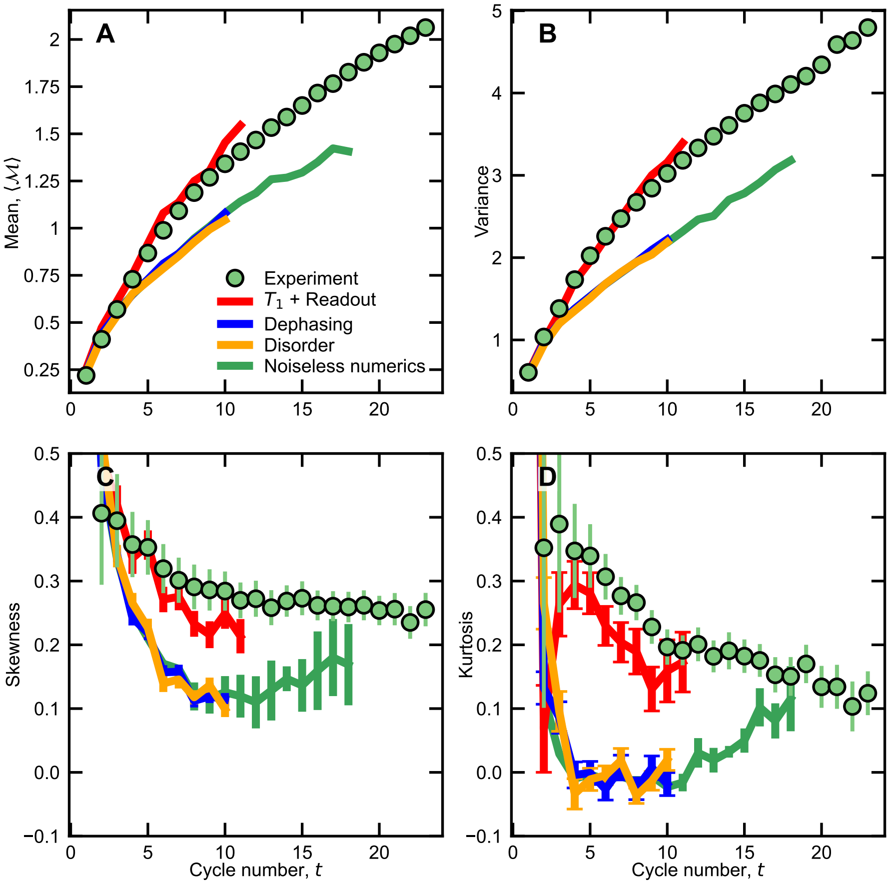


Figure C.5: Effects of various sources of error on the first four moments of the transferred magnetization, shown for $(\theta, \phi) = (0.17\pi, 0.6\pi)$ and $m = 0.5$, the same as the green curve in Figure 2 of the main paper. The red curve shows the effects of amplitude damping and readout error. The amplitude damping rate, T_1, is extracted from the experimental data as in Figure C.1A and is assumed to be the same for all qubits. Readout error rates are measured by sampling random bitstrings on all 46 qubits. Dephasing is characterized using a floquet protocol. Disorder refers to miscalibrations of the angles θ and ϕ as well as the single-qubit Z phases (see Eq. S2 of [119]). Miscalibrations of these angles are measured using unitary tomography. Error bars are estimated using the delete-one jackknife method described in Section C.1.4. It is clear that amplitude damping and readout constitute the dominant source of error. This is encouraging for future experiments because the overall error rate is controlled by the $0 \rightarrow 1$ readout error rate, which can be decreased at the expense of the $1 \rightarrow 0$ error rate.

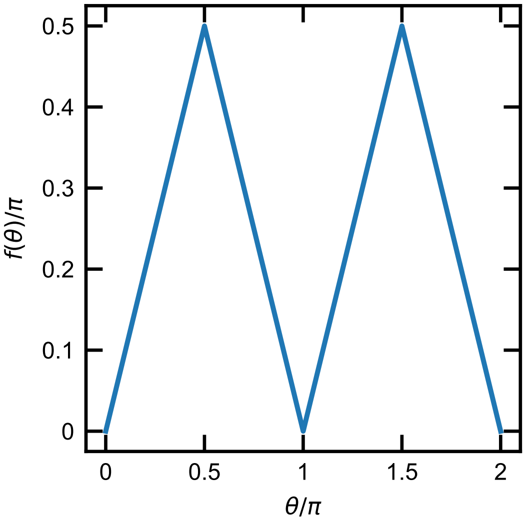


Figure C.6: The periodic triangle wave function f appearing in Eq. (C.5).

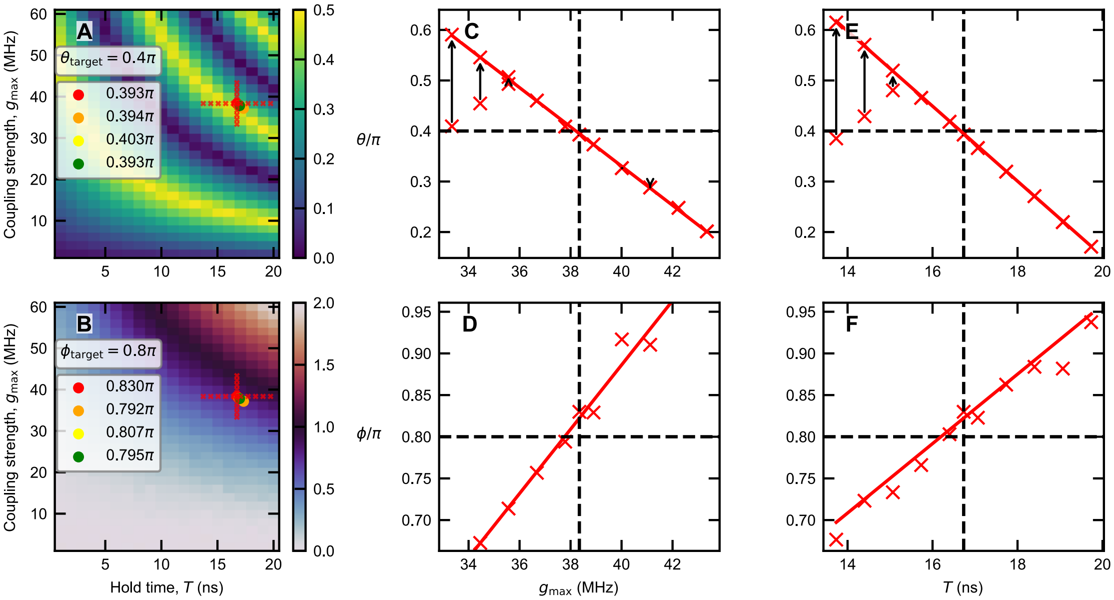


Figure C.7: Illustration of the iterative calibration of fSim gates. Panels **A** and **B** show, in the background, the results of a 2D grid sweep, characterizing the fSim angles θ and ϕ, respectively, as functions of the hold time, T, and coupling strength, $g_{\max}$, of the trapezoidal pulse, for a particular qubit pair. (The rise time of the trapezoid is fixed at 5 ns.) The red colored marker indicates an initial guess for T and $g_{\max}$. In this case, the initial guess comes from the 2D grid sweep, but an advantage of the iterative technique is that the grid sweep is not needed. More often, we provide the initial guess without ever doing a 2D sweep, allowing faster calibration of gates. The measured fSim angles for this guess are shown in the boxes on the left. We then measure the fSim angles along the cross shape indicated by the small red xs in panels **A** and **B**. The results of these measurements are shown in panels **C**-**F**, with panels **C** and **D** showing the constant-T part of the cross, and panels **E** and **F** showing the constant-$g_{\max}$ part. For the θ fits, the triangle-wave (f in Eq. (C.5)) is unwound as indicated by the arrows. This fitting procedure gives us the six parameters in Eq. (C.5). We then invert the fitted polynomial to obtain the next guess for T and $g_{\max}$, indicated by the orange point in panels **A** and **B**. Sometimes, no real solutions exist within the physical bounds that we place on T and $g_{\max}$, in which case we find the closest approach, minimizing the cost function $C = (\phi - \phi_0)^2/\phi_0^2 + (\theta - \theta_0)^2/\theta_0^2$, where θ_0 and ϕ_0 are the target angles, subject to the bounds that we place on T and $g_{\max}$. The cross sweep is iterated several times (in this case 3 times, resulting in 4 successive guesses), and the best guess (the one minimizing the cost function) is used.

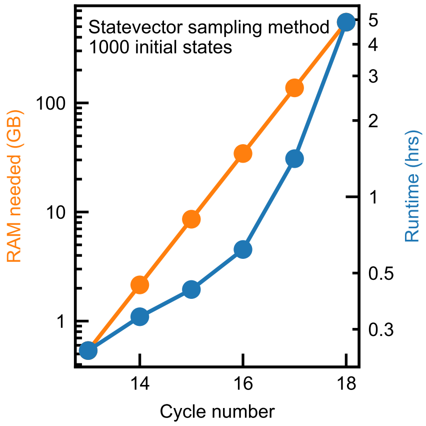


Figure C.8: The resources needed to generate 1000 initial states using our implementation of the statevector sampling method on up to eight 80-GB NVIDIA A100 GPUs using NVIDIA cuQuantum [55].

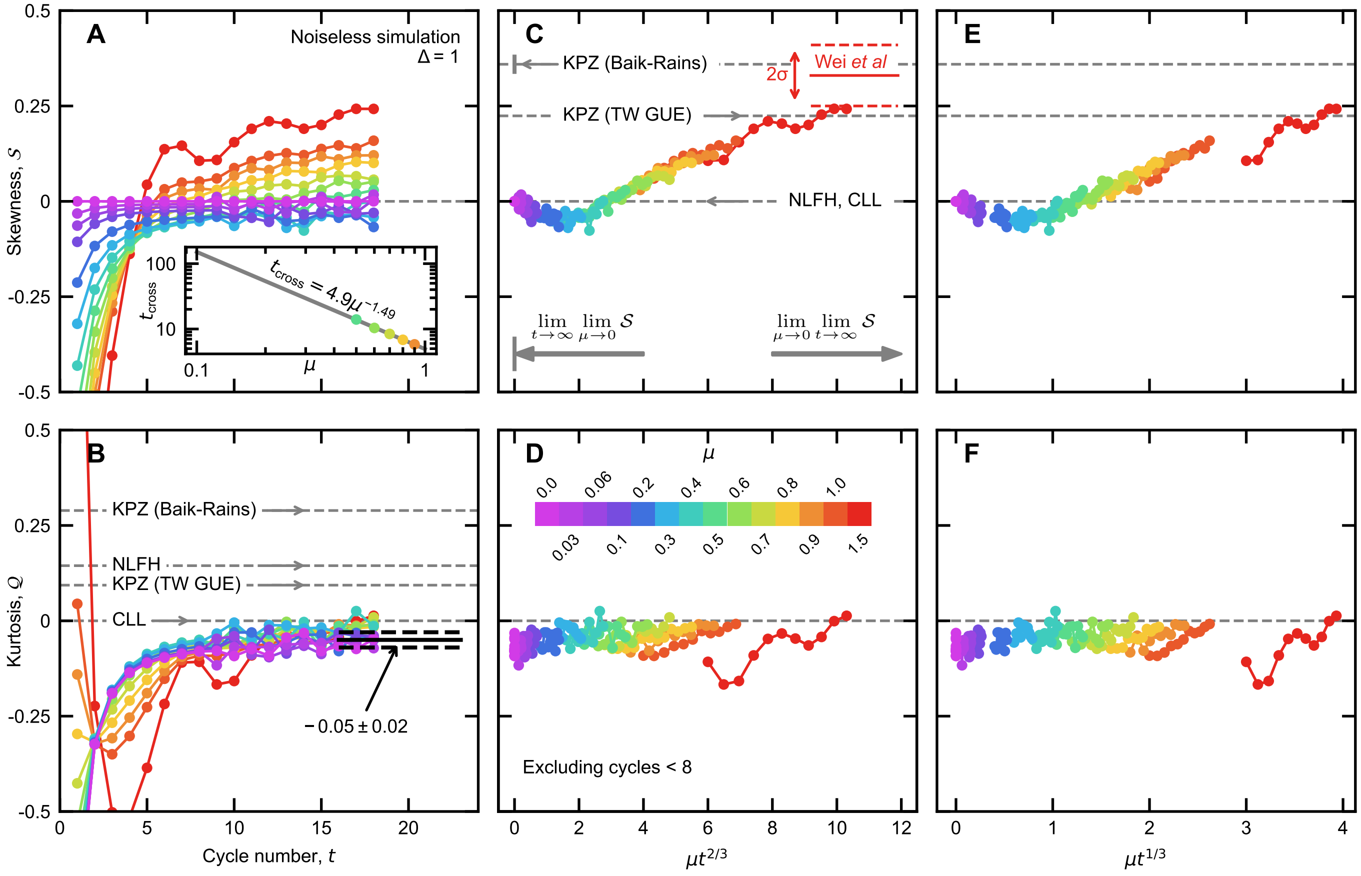


Figure C.9: Data collapse. Panels **A**–**D** are the same as Figure 3 of the main text, but showing only data from noiseless numerical simulations. The added inset in panel **A** shows the time at which the skewness becomes positive as a function of μ, with a fitted power law. The fact that the fitted exponent is close to $-3/2$ justifies the scaling in panels **C-D**. We further justify this scaling by noting that, among all scaling variables of the form μt^{γ}, $\gamma \approx 0.65$ leads to the best collapse of the skewness, as measured by the sum of the residuals. In Panels **E**–**F**, we try a scaling exponent of $1/3$, which is motivated by the theoretical finding that the dynamics become diffusive at a time that scales as μ^{-3} [64]. Although the skewness collapses reasonably well in both panels **C** and **E**, a careful visual comparison shows that the collapse is better in panel **C**. The kurtosis does not collapse with either scaling, as expected; the kurtosis cannot be a function of μt^{γ} for any positive γ because, unlike the skewness, it has time dependence even when $\mu = 0$.

Figure C.9 shows that the skewness also appears to collapse reasonably well when plotted as a function of $\mu t^{1/3}$, thus conveying the difficulty of estimating the exponent of t by eye. In order to do so in an unbiased manner, we define a quantitative measure of data collapse based on the appropriately normalized sum of fit residuals (where the fit describes the purportedly universal scaling function). Using μt^{γ} as the scaling variable, for varying γ, we find that this metric is minimized for $\gamma \sim 0.65$. This is consistent with our initial observation of $\mu t^{2/3}$ seemingly yielding the best data collapse.

C.2.6 Sweeps of anisotropy and imbalance

The transport characteristics of the XXZ model are strongly dependent on the anisotropy parameter, Δ. In particular, ballistic, superdiffusive and diffusive behaviors are expected in the regimes where $\Delta < 1$, $\Delta = 1$ and $\Delta > 1$, respectively. Moreover, the KPZ conjecture was only proposed for the isotropic point ($\Delta = 1$). In order to get a better sense of the parameter space in which we are operating, we performed 2D numerical sweeps of the anisotropy Δ and the initial imbalance μ. The results are shown in Figure C.10. They illustrate a sign change in the skewness and kurtosis close to the Heisenberg point, $\Delta = 1$, as well as a clear change in the dynamical exponent.

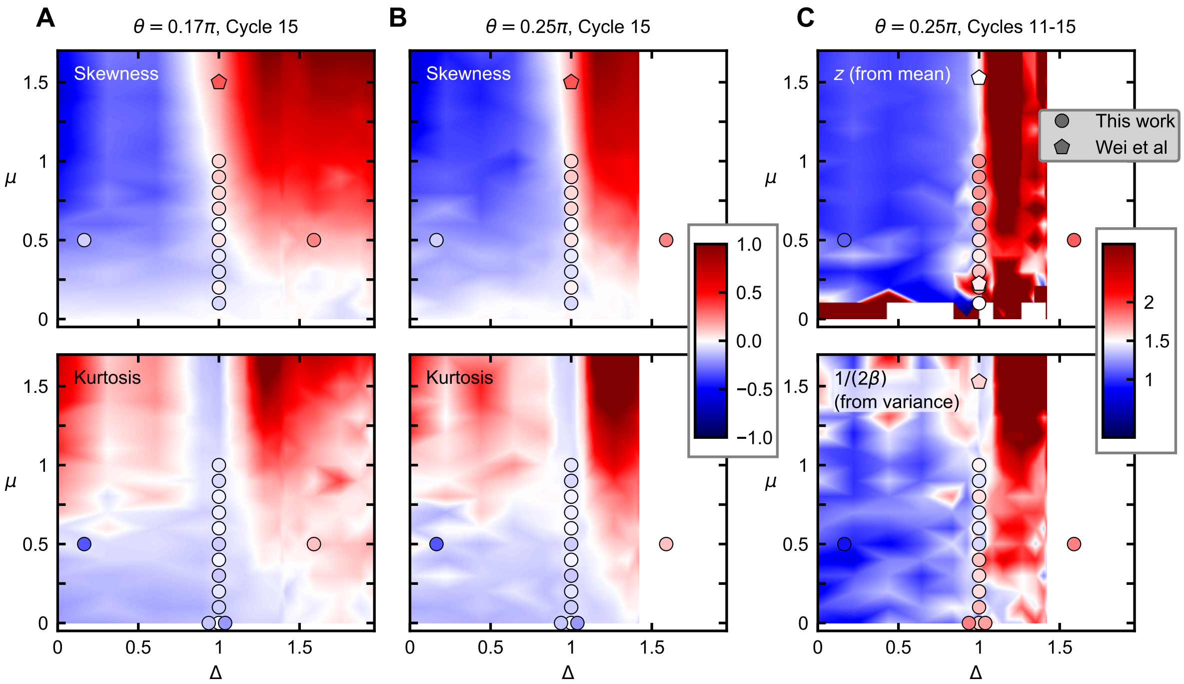


Figure C.10: Skewness, kurtosis, and dynamical exponents as a function of initial imbalance μ and anisotropy Δ. The heatmaps show interpolated numerical results, computed using 500 random initial states. The markers plotted on top of the background indicate our experimental results and those of Ref. [171]. Our experimental values of skewness and kurtosis are averaged over cycles 15 through 23, and the experimental dynamical exponents are fit to cycles 11 through 23. The background heatmaps are computed for the indicated values of θ, which, except for $\theta = 0.17\pi$ and the $\Delta = 1.59$ experimental point, are not the same θ as used in the experiments; the other experimental points use $\theta = 0.4\pi$.

C.3 Further experimental data

C.3.1 Dynamical exponent

In Fig. C.11, we plot both the experimentally observed and numerically simulated dynamical exponents of the mean and the variance, as a function of the initial imbalance, μ. At large μ, we find that the dynamical exponent is higher than the superdiffusive value of 3/2. The observed values are consistent with Ref. [104], where it was found that the dynamical exponent drifts from 3/2 at small initial imbalance m to approximately 5/3 when m is about 1.

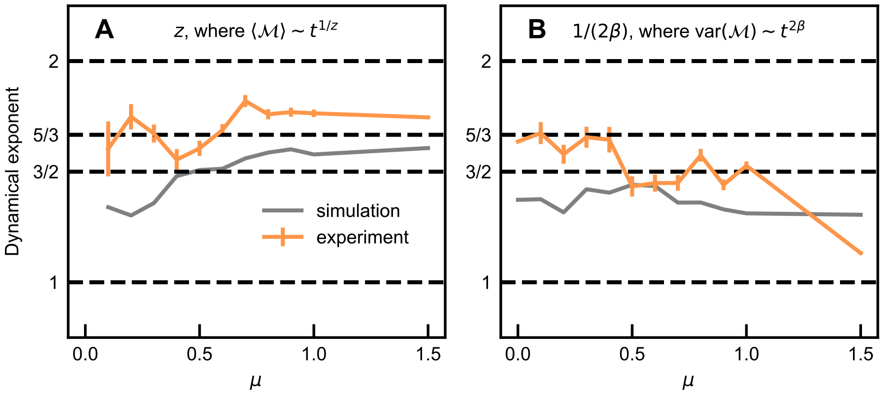


Figure C.11: The dynamical exponent (**A** of the mean, and **B** of the variance) as a function of the initial imbalance. The KPZ prediction in both cases is 3/2. The dynamical exponents are extracted from fits over cycles 11-23 for the experiment and 11-18 for the simulation.

C.3.2 Pure domain walls

In addition to the results presented in the main paper, we also studied the pure domain wall, $\mu = \infty$. This is a simpler experiment because it does not require any averaging over initial states. It is also easier for classical simulations; TEBD simulations converge at least to cycle 23, allowing us to check our experimental results at later cycles than is possible at finite μ with the simulation techniques used here. Experiment and simulation results are shown in Figure C.12. Our findings are largely consistent with expectations (see Section 6.1 of [20] for a review). We observe an absence of transport in the easy-axis regime ($\Delta > 1$), with the observed transport in the experiment consistent with amplitude damping and readout errors (see Figure C.14). In the easy-plane regime ($\Delta < 1$), we observe ballistic transport. In the isotropic ($\Delta = 1$) case, we observe transport with a dynamical exponent of about 5/3, consistent with the finding of Ref. [104] and differing from the 3/2 dynamical exponent of KPZ. Figure C.13 compares the skewness and kurtosis at the isotropic point with the KPZ predictions. While the skewness is close to the KPZ prediction, it can be seen in the simulation results that it continues to increase above the KPZ value. The kurtosis approaches the KPZ prediction, but it is not clear that it has stopped increasing.

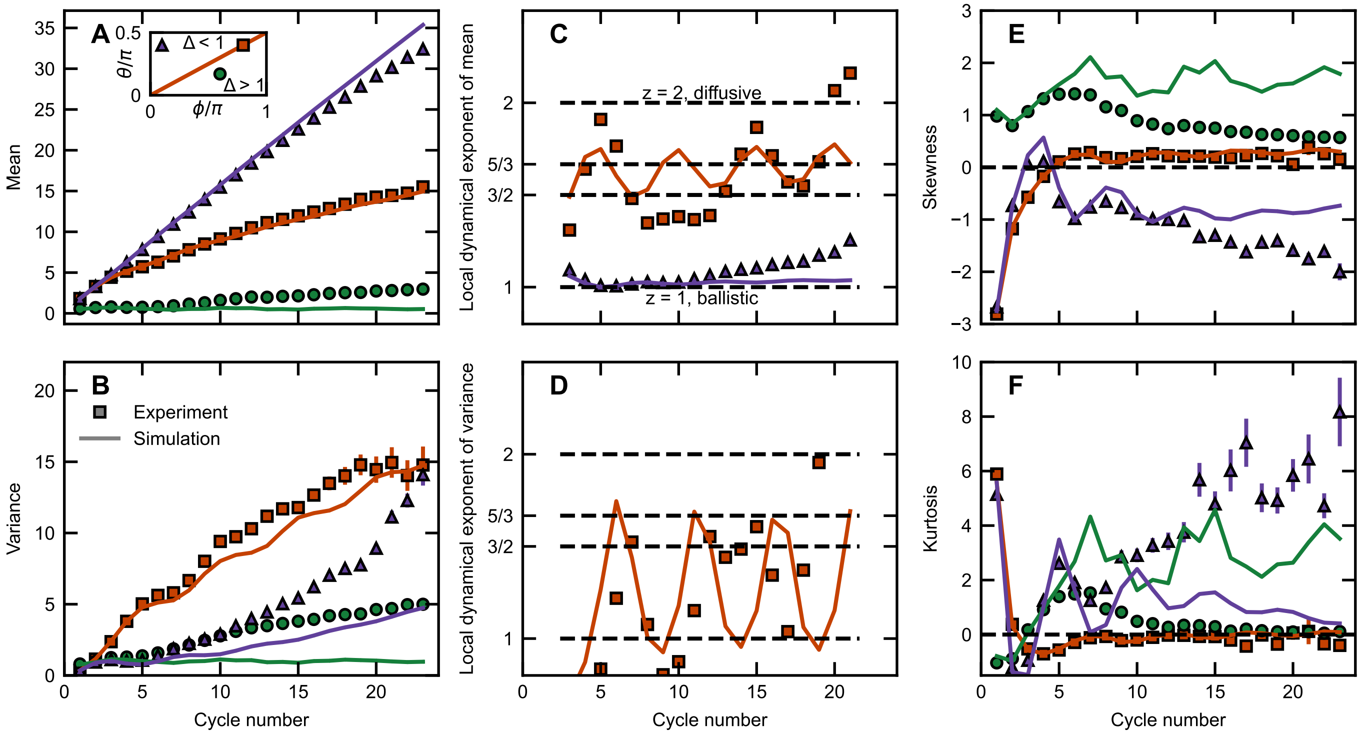


Figure C.12: Experiment and simulation results for the pure domain wall initial state, $\mu = \infty$. The three pairs of fSim angles shown are $(\theta, \phi) = (0.4\pi, 0.8\pi)$ (orange squares), $(\theta, \phi) = (0.4\pi, 0.1\pi)$ (purple triangles), and $(\theta, \phi) = (0.17\pi, 0.6\pi)$ (green circles). The simulation results for cycles 19-23 are computed using TEBD.

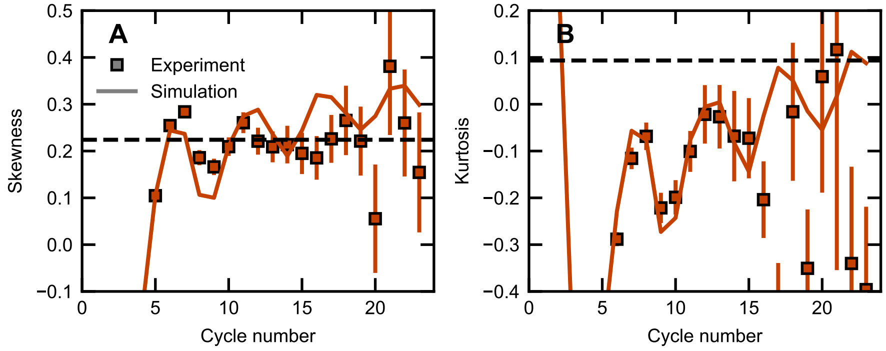


Figure C.13: Experiment and simulation results for the pure domain wall initial state, $\mu = \infty$ with $(\theta, \phi) = (0.4\pi, 0.8\pi)$, i.e. $\Delta = 1$. Horizontal dashed lines indicate the KPZ prediction for a wedge-shaped initial state (see Table C.2). We show these values for comparison only. The fact that the dynamical exponent is larger than 3/2 (see Fig. C.12C) already indicates the lack of KPZ scaling. Further, the skewness continues to increase above the KPZ prediction.

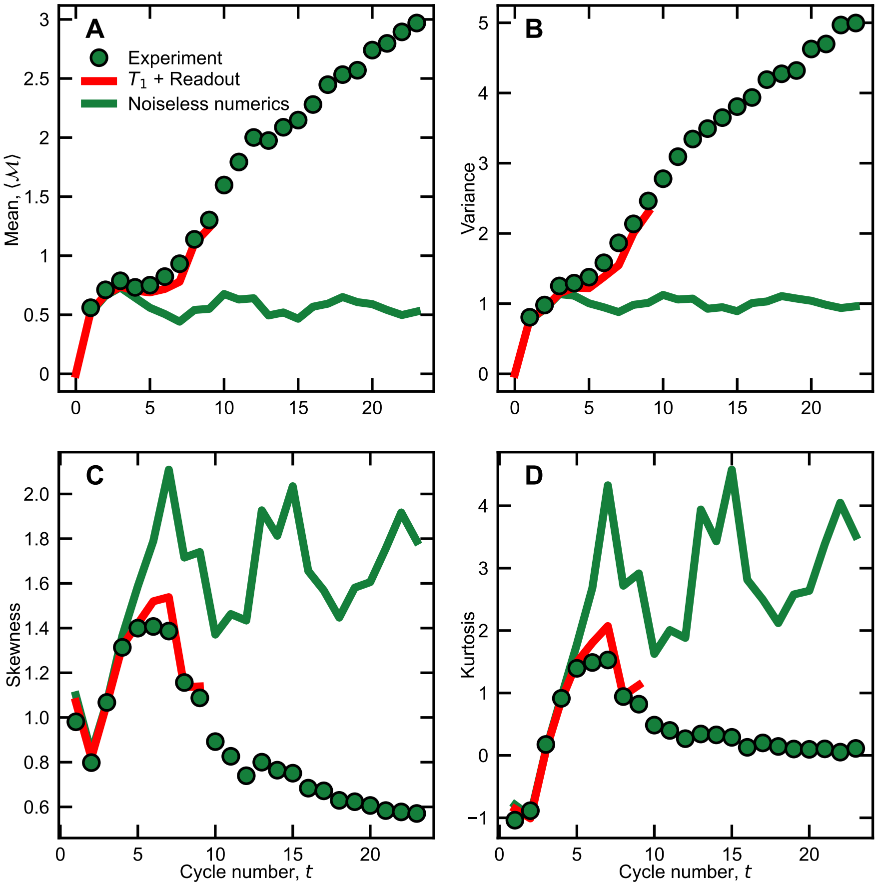


Figure C.14: Comparison of experimental results for the pure domain wall simulation ($\mu = \infty$), $(\theta, \phi) = (0.17\pi, 0.6\pi)$, $\Delta = 1.6$, with noiseless and noisy simulations. The noisy simulations include only the effects of amplitude damping (T_1) and readout error, with T_1 extracted from the measured algorithmic relaxation time and the readout error rates benchmarked on the device. Compare with Figure C.5, which shows similar results for $\mu = 0.5$.

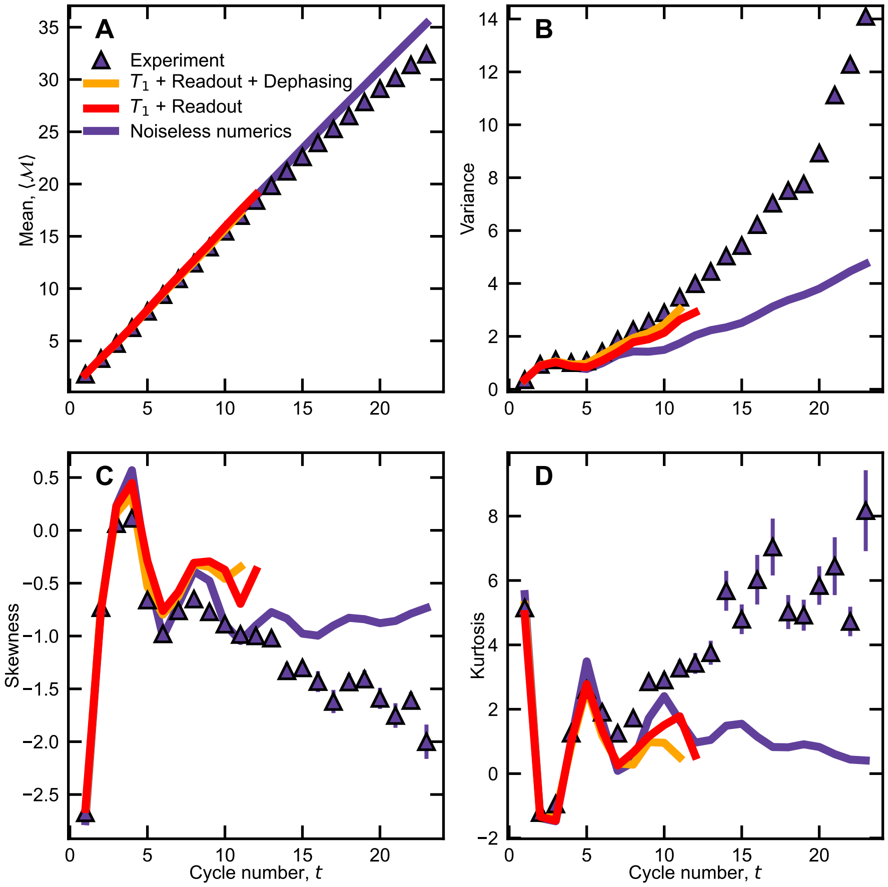


Figure C.15: Sources of noise in the $\Delta = 0.16$, $\mu = \infty$ experiment. Similar to Figure C.14 but for $(\theta, \phi) = (0.4\pi, 0.1\pi)$. In this case, dephasing does have a noticeable effect on the variance.

C.4 The Kardar-Parisi-Zhang universality class

In 1985, Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang [86] set out to study the general properties of stochastically growing interfaces, examples of which include flame fronts and tumors. They abstracted these situations to a height function $h(\vec{x}, t)$ that obeys

$$\frac{\partial h}{\partial t} = \nu \nabla^2 h + \frac{\lambda}{2} (\nabla h)^2 + \eta(\vec{x}, t), \tag{C.12}$$

where $\eta(\vec{x}, t)$, at each position and time, is an independent zero-mean Gaussian random variable with variance proportional to a parameter D. This is the stochastic Burgers equation for the slope ∇h. The $\nabla^2 h$ is a diffusive term, and the $(\nabla h)^2$ is a nonlinearity. In general, an equation describing a growing interface may have higher powers of the slope, such as $(\nabla h)^4$, but these are irrelevant to the large-scale physics in the renormalization group (RG) sense. Indeed, in 1+1 dimensions ($d = 1$), the KPZ equation [Eq. (C.12)] has divergences that must be regularized, e.g., by putting it on a spatial lattice or imposing a maximum cutoff wavenumber Λ in Fourier space. Kardar Parisi, and Zhang considered an RG flow in which (1) high-wavenumber modes are integrated out ($\Lambda \to e^{-l}\Lambda$), (2) space and time are rescaled so that the new smallest length scale is labeled by the same numerical value as the old smallest length scale ($\vec{x} \to e^{-l}\vec{x}$, $t \to e^{-zl}t$), and (3) the height function is rescaled ($h \to e^{-(d+\chi)l}h$). This procedure can be thought of as coarse graining and zooming out, while changing units so that it appears as though the zoom-out did not occur. Kardar, Parisi and Zhang showed that Eq. (C.12) is a fixed-point of this RG flow, for specific choices of the scaling exponents and parameters. In 1+1 dimensions in particular, if $z = 3/2$, $\chi = 1/2$, and $\lambda^2 D/\nu^3 = 2$, then Eq. (C.12) is invariant under the rescaling and coarse graining procedure. Further, this fixed point is stable; if $\lambda^2 D/\nu^3 > 2$, then it will flow down to 2 under the coarse graining. Conversely, if $\lambda^2 D/\nu^3 < 2$, then it will flow up to 2.

Kardar, Parisi, and Zhang therefore proposed a new universality class (the KPZ universality class) into which their equation falls. They conjectured that a variety of other systems, such as ballistic deposition (e.g., snow falling and sticking together), and the Eden model (a random growth model), are in this universality class, based on scaling exponents observed in earlier numerical experiments [134, 54].

The KPZ equation and universality class have since been studied in great detail. For example, Kurt Johansson [80] studied the asymmetric exclusion process, a model in this universality class, for an initial state corresponding to a wedge-shaped $h(x, t = 0)$, and found that a quantity corresponding to the regularized height function, $2h(0, t) - h(\infty, t) - h(-\infty, t)$, follows the Tracy-Widom (TW) distribution for the largest eigenvalue in the Gaussian Unitary Ensemble (GUE) at late times. Prähofer and Spohn [138] generalized this result, using a mapping from the polynuclear growth model, which is in the KPZ universality class, to random permutations and from there to random Gaussian matrices to again identify the asymptotic probability distribution of the regularized height function. The precise distribution depends on the initial conditions, and there are three cases that are relevant for us: (1) *flat*, meaning $h(x, 0) = 0$, (2) *stationary*, meaning that $h(x_i, 0) = h(x_i - 1, 0) + \eta_i$, where η_i is randomly ±1 with equal probabilities, and (3) *wedge-shaped*, meaning that $h(x,0) = -|x|/\delta$. The wedge-shaped initial condition leads to the same distribution as the curved initial condition studied by Prähofer and Spohn [69] and is unaffected by nonzero variance in $h(x_i, 0)$, unlike the flat initial condition, which becomes the stationary initial condition with the addition of fluctuations. Prähofer and Spohn found that the asymptotic probability distributions of the regularized height function in these three cases are (1) GOE Tracy-Widom, (2) Baik-Rains, and (3) GUE Tracy-Widom, in agreement with [80]. The corresponding values of skewness and kurtosis are listed in Table C.2.

Spin chain initial state	KPZ initial condition	Asymptotic probability distribution	Skewness	Kurtosis
	Flat	GOE TW	0.294	0.165
$\mu = 0$	Stationary	Baik-Rains	0.359	0.289
$\mu > 0$	Wedge	GUE TW	0.224	0.093

Table C.2: Asymptotic probability distribution of the regularized KPZ height function [138], for the three initial conditions described in the text. The conjecture that the Heisenberg spin chain is in the KPZ universality class implies that the asymptotic transferred magnetization has the same distribution as the regularized KPZ height function. Therefore, the conjecture predicts an asymptotic skewness of 0.359 and an asymptotic kurtosis of 0.289 for the $\mu = 0$ state. Note that although we included $\mu > 0$ in the first column, the infinite-time dynamics of the Heisenberg chain in this case have been shown to be diffusive rather than KPZ [99, 64].

Prähofer and Spohn [137] further used the polynuclear growth model to solve for the two point correlation correlation function,

$$C(x,t) = \langle (h(x,t) - h(0,0) - t\langle \partial_t h \rangle) \rangle \,, \tag{C.13}$$

assuming stationary initial conditions, which imply that $C(x,0) = \frac{D}{\nu}|x|$. The slope-slope correlation function can then be obtained as

$$\langle \partial_x h(0,0) \partial_x h(x,t) \rangle = \frac{1}{2} \partial_x^2 C(x,t). \tag{C.14}$$

$C(x,t)$ takes the form

$$C(x,t) = t^{2/3} g\left(\text{const} \cdot x/t^{2/3}\right), \tag{C.15}$$

where $g(y)$ is a universal scaling function. Defining the scaling function $f(y) = \frac{1}{4}g''(y)$, which is proportional to the slope-slope correlation function, Prähofer and Spohn obtained exact numerical solutions for $f(y)$, which they found to behave as $f(y) \sim e^{-0.295|y|^3}$ for large y, falling off faster than a Gaussian.

Evidence for anomalous transport in the Heisenberg spin chain at nonzero temperature was first found in the late 1990s. Sachdev and Damle [145, 39] explained diffusive ($z = 2$) nonzero-temperature transport in the easy-axis ($\Delta > 1$) XXZ model even though quasiparticles propagate ballistically, whereas other works [183, 139] found ballistic ($z = 1$) behavior at finite temperature in the easy-plane ($\Delta < 1$) regime, suggesting anomalous transport at $\Delta = 1$.

The first numerical evidence for anomalous transport in the infinite-temperature $\Delta = 1$ Heisenberg spin chain was provided in 2011 by Ref. [169]. The $z = 3/2$ exponent was demonstrated numerically in 2017 [104], in partially polarized domain wall initial states similar to those studied in our work, which approach the infinite-temperature state as $\mu \to 0$. The 3/2 exponent alone was not enough for the authors to propose that the Heisenberg spin chain is in the KPZ universality class, as it could have other explanations. In 2019 [106], however, they found numerically, for both the continuous-time Heisenberg model and the Floquet version studied here, that the two-point spin-spin correlation function at infinite temperature precisely matched the KPZ prediction for the slope-slope correlation function, Eq. (C.14), including the deviations from Gaussian at the tails. They therefore proposed that the infinite-temperature spin-1/2 Heisenberg model is in the KPZ universality class, with $\sigma_i^z \leftrightarrow \partial_x h(x_i)$, and that the infinite-temperature initial condition on the spin chain side corresponds to the stationary initial state (see Table C.1) on the KPZ side. In their work, they used the finite-μ domain wall states studied here as a computational tool for obtaining the two-point correlation function at infinite temperature.

A number of works have proposed theoretical explanations for the observed $z = 3/2$ dynamical exponent in the Heisenberg spin chain (e.g., [77, 62, 41, 63, 21, 43], see [20] for a review). The picture that emerges from these theoretical explanations is that the $z = 3/2$ dynamical exponent and the two-point correlation function are universal for 1D integrable quantum systems with a non-Abelian global symmetry [77, 43, 178], which, in the spin-1/2 Heisenberg model, is SU(2).

The $z = 3/2$ dynamical exponent has also been observed in several experiments [171, 150, 88]. Ref. [171], for example, studied the transferred magnetization, which, assuming $\sigma_z \leftrightarrow \partial_x h$, corresponds to $2h(0,t) - h(\infty,t) - h(-\infty,t)$ for domain-wall initial states with several initial imbalances μ. They confirmed the KPZ predictions that the mean and variance of the transferred magnetization both grow as $t^{2/3}$. They also measured the skewness which, in a nonzero-μ domain wall initial state, is expected to asymptote to 0.2241 (see Table C.2). They measured 0.33 ± 0.08, where the uncertainty is one standard deviation, a result consistent with the KPZ prediction. They also confirmed that breaking either integrability or SU(2) symmetry causes the dynamics to become either ballistic ($z = 1$) or diffusive ($z = 2$). However, they measured the skewness for a domain wall with a very large initial imbalance ($\mu = 1.5$), whereas the KPZ dynamics are expected to emerge at small μ.

However, there is a problem with the conjecture that the infinite-temperature ($\mu = 0$) Heisenberg model is in the KPZ universality class. As Refs. [99, 123] point out, the probability distribution of the transferred magentization in this state must be symmetric; excitations are just as likely to move from the right side of the chain to the left as from the left to the right. Therefore, all of the odd moments of this distribution must be zero. This differs from the Baik-Rains distribution, which has a nonzero skewness of 0.359 [138]; see Table C.2.

If one reversed the order of limits, first taking $t \to \infty$ and then $\mu \to 0$, the resulting transferred magnetization distribution may be skewed because even an infinitesimal domain wall breaks the mirror symmetry of the $\mu = 0$ state. However, this does not resolve the issue because the late-time behavior at nonzero μ has been shown to be diffusive rather than KPZ [99, 64]. Ref. [21] would suggest otherwise, namely, that KPZ dynamics emerge even for $\mu > 0$. In that work, the Heisenberg spin chain is coarse-grained and the global SU(2) symmetry is promoted to a gauge symmetry, with a dynamical gauge field specifying the direction of the local Bethe vacuum in each lattice cell. A long-wavelength torsional mode of the gauge field is shown to obey a stochastic Burgers equation when the quasiparticle occupancy is uniform across cells, a condition that also holds for the $\mu > 0$ domain-wall states. However, this work does not connect the dynamics of the torsional mode with the dynamics of σ_i^z, the variable that is has been observed to play the role of $\partial_x h$. A summary of the different regimes discussed here and ways in which they differ from the KPZ universality class is shown in Table C.3.

There are other ways of taking the limit, illustrated in Fig. C.16. So far, we have considered taking $\mu \to 0$ first (purple line) or $t \to \infty$ first (red line), neither of which can result in KPZ dynamics. Taking a simultaneous limit can avoid these theoretical arguments. In particular, diffusive dynamics are expected to emerge at a time that scales as $1/\mu^3$ [64]. Therefore, if t is scaled with μ in such a way that it remains less than $1/\mu^3$, for infinitesimal μ, but still approaches infinity, the theoretical arguments against KPZ are avoided and the experiment has a chance of providing new information. In Fig. C.16, the orange curve, $\mu \sim t^{-1/3}$ indicates the threshold for diffusive dynamics; any curve approaching the origin that remains above the orange curve will result in diffusion. The purple curve results in symmetrically-distributed transferred magnetization, unlike the KPZ prediction, which is skewed. The green curve is an example of a way of taking the limit that is not ruled out theoretically. Although there is no prior theoretical motivation for taking the limit in any order other than $\mu \to 0$ first, the collapsing behavior observed in Fig. C.9 suggests that the skewness may be a function of $\mu t^{2/3}$, which is constant along the $\mu \sim t^{-2/3}$ curve. Further, from Fig. C.9 and Fig. 3 of the main text, if $\mu t^{2/3}$ is fixed to a large number (at least about 10), the skewness appears to be consistent with that of the TW GUE probability distribution. However, we experimentally find a kurtosis of about -0.05 ± 0.02 (Fig. 3 of the main text and Fig. C.9), inconsistent with the TW GUE kurtosis of 0.09. This rules out KPZ dynamics on the timescale of the experiment. It does not rule out KPZ dynamics at much later times, but we do not see evidence for these dynamics either.

Therefore, the challenge is to explain why the dynamical exponent and two-point correlation function, at infinite temperature, look like KPZ, universally across integrable 1D quantum systems with a global non-Abelian symmetry, and yet other observables, such as the skewness of transferred magnetization when $\mu = 0$, differ. In response to this challenge, Refs. [99, 98] proposed and studied a classical Landau-Lifshitz (CLL) model and Refs. [21, 123] a non-linear fluctuating hydrodynamics (NLFH) model. Both systems predict a symmetric distribution for the transferred magnetization. CLL predicts an excess kurtosis close to 0, whereas NLFH model predicts 0.14. Although CLL agrees nicely with the findings reported in this work, it is an example of another system with similar behavior rather than an explanation of why these features should appear in the Heisenberg spin chain.

<table>
<tr><th rowspan="2"></th><th rowspan="2">SU(2)?</th><th rowspan="2">Dynamical exponent</th><th rowspan="2">Skewness in Heisenberg</th><th colspan="3">Skewness in KPZ</th><th rowspan="2">Universality of Heisenberg</th></tr>
<tr><th>$h(x, t = 0)$</th><th>$p(\mathcal{M})$</th><th>$\mathcal{S}$</th></tr>
<tr><td>$\mu = \infty$ (pure DW),
t finite</td><td>No</td><td><table><tr><th>Source</th><th>z</th></tr><tr><td>Numerics [104, 105]</td><td>5/3</td></tr><tr><td>Fig. C.12C</td><td>5/3</td></tr></table></td><td>~ 0.3 (Fig. C.13A)</td><td>Wedge</td><td>TW GUE</td><td>0.224</td><td>KPZ not conjectured:
t finite, SU(2) broken</td></tr>
<tr><td>$\mu = \infty$, $t \to \infty$</td><td>No</td><td>log-enhanced diffusion
($z = 2$) [117, 156]</td><td></td><td>Wedge</td><td>TW GUE</td><td>0.224</td><td>KPZ not conjectured:
SU(2) broken</td></tr>
<tr><td>μ finite, t finite</td><td>No</td><td><table><tr><th>Source</th><th>μ</th><th>z</th></tr><tr><td>Ref. [171]</td><td>0.22</td><td>1.54</td></tr><tr><td>Ref. [171]</td><td>1.5</td><td>1.49</td></tr><tr><td>Ref. [171]</td><td>1.8</td><td>1.45</td></tr><tr><td>Fig. C.11, sim</td><td>0.5</td><td>1.49</td></tr><tr><td>Fig. C.11, exp</td><td>0.5</td><td>1.59</td></tr><tr><td>Fig. C.11, sim</td><td>1.0</td><td>1.59</td></tr><tr><td>Fig. C.11, exp</td><td>1.0</td><td>1.75</td></tr></table></td><td>Ref. [171] ($\mu = 1.5$): 0.3
This work: function of $\mu t^{2/3}$
Consistent with Ref. [171].</td><td>Wedge</td><td>TW GUE</td><td>0.224</td><td>KPZ not conjectured:
t finite, SU(2) broken</td></tr>
<tr><td>$t \to \infty$, then $\mu \to 0$</td><td>Subtle*</td><td><table><tr><th>Source</th><th>z</th></tr><tr><td>Numerics [99]</td><td>2</td></tr><tr><td>Theory [64]</td><td>2</td></tr></table></td><td>Consistent with 0.224
(Figs 3, C.9 with $\mu t^{2/3} \to \infty$)</td><td>Wedge</td><td>TW GUE</td><td>0.224</td><td>Not KPZ: Wrong z</td></tr>
<tr><td>$\mu = 0$ (∞-temp state),
$t \to \infty$
($\mu \to 0$, then $t \to \infty$)</td><td>Yes</td><td><table><tr><th>Source</th><th>$1/(2\beta)$</th></tr><tr><td>Numerics [104, 106]</td><td>3/2</td></tr><tr><td>Theory [77, 62, 41, 63, 21, 43]</td><td>3/2</td></tr><tr><td>Fig. C.11B, sim</td><td>1.4</td></tr><tr><td>Fig. C.11B, exp</td><td>1.6</td></tr></table></td><td>Zero by symmetry</td><td>Stationary</td><td>Baik-Rains</td><td>0.359</td><td>Not KPZ: Wrong $\mathcal{S}$</td></tr>
</table>

Table C.3: Regimes of the 1D spin-1/2 Heisenberg model, starting in μ-parametrized domain wall states (Eq. 3 of the main text). From the left, the first column indicates the regime, i.e., the initial imbalance μ and the time t. The second column indicates whether the initial state is SU(2)-symmetric. The third column lists the dynamical exponent in this regime. The fourth column indicates the skewness of the transferred magnetization. Column 5 shows the corresponding initial condition of the KPZ height function, assuming that $\sigma_z \sim \partial_x h$. Column 6 states the asymptotic probability distribution of the regularized KPZ height function for the given initial condition (see Table C.2); the regularized height function corresponds to the transferred magnetization if $\sigma_z \sim \partial_x h$. The seventh column states the skewness of the probability distribution in the sixth column (see Table C.2). Finally, the eighth column lists reasons why the particular case is not in the KPZ universality class. In the second-to-last row, "subtle" refers to the fact that a state with infinitesimal μ breaks the SU(2) symmetry by an infinitesimal amount. Such states would be appropriate for studying spontaneous symmetry breaking, where the symmetry might be broken by an infinitesimal external field.

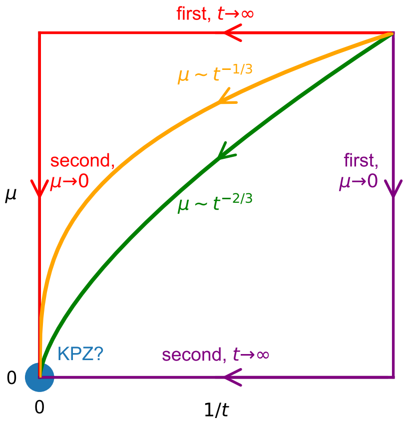


Figure C.16: Different ways of taking the limits $t \to \infty$ and $\mu \to 0$, i.e. of approaching the point where KPZ is conjectured. Taking the limit along any curve above the orange curve results in diffusive dynamics [64]. This includes the red curve, which corresponds to taking time to infinity first. The purple curve results in a symmetrically-distributed transferred magnetization, differing from the KPZ prediction, which is skewed [123, 98, 99]. The green curve shows an example of a way to take the limits simultaneously. Although this way could exhibit a skewness and a dynamical exponent consistent with KPZ, its predicted kurtosis disagrees with our observations for the times considered in our experiment.

C.5 Unifying fSim conventions

C.5.1 Placement of the phase angle

In our work, we use the following definition of the fSim gate:

$$U_{\text{fSim}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) & i\sin(\theta) & 0 \\ 0 & i\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 & e^{-i\phi} \end{pmatrix}, \tag{C.16}$$

which is a fully general number-conserving two-qubit gate up to single-qubit Z rotations. Another natural choice would be to split the phase between the $|00\rangle$ and $|11\rangle$ states:

$$U'_{\text{fSim}} = \begin{pmatrix} e^{-i\phi/2} & 0 & 0 & 0 \\ 0 & \cos(\theta) & i\sin(\theta) & 0 \\ 0 & i\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 & e^{-i\phi/2} \end{pmatrix}. \tag{C.17}$$

This latter definition, which is more directly related to the trotterized Heisenberg Hamiltonian, is related to ours by

$$U'_{\text{fSim}} = e^{-i\sigma_1^z\phi/4}e^{-i\sigma_2^z\phi/4}U_{\text{fSim}} = U_{\text{fSim}}e^{-i\sigma_1^z\phi/4}e^{-i\sigma_2^z\phi/4}. \tag{C.18}$$

It is not immediately obvious that our experiment is insensitive to whether we use U_{fSim} or U'_{fSim}, but this turns out to be the case. The transferred magnetization is independent of whether one uses U'_{fSim} or U_{fSim} as long as the number of cycles t

is at most $N_Q/2$. We verified this similarly to how we verified that the transferred magnetization is independent of N_Q as long as $t \leq N_Q/2$ in Table C.1. For example, with $\mu = 0.5$, and $(\theta, \phi) = (0.4\pi, 0.8\pi)$, a density matrix simulation on four qubits gives a kurtosis after two cycles of -0.30867052, to this many digits, regardless of whether one uses U_{fSim} or U'_{fSim}, and an 8-qubit simulation gives a kurtosis of -0.12588028 at cycle 4 regardless of which fSim gate one uses. The choice matters starting at cycle $N_Q/2+1$, but we are not interested in times past $N_Q/2$ because we have already seen (Table C.1) that finite size effects appear there.

Similarly, the transferred magnetization is unchanged, even past time $N_Q/2$, under either $\theta \to -\theta$ or $\phi \to -\phi$.

C.5.2 Comparison to the η, λ parameterization

Following Ref. [107], it can be shown that Floquet application of the gate unitary,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{\sin\eta}{\sin(\lambda+\eta)} & \frac{\sin\lambda}{\sin(\lambda+\eta)} & 0 \\ 0 & \frac{\sin\lambda}{\sin(\lambda+\eta)} & \frac{\sin\eta}{\sin(\lambda+\eta)} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{C.19}$$

gives the desired Heisenberg Hamiltonian evolution with $\Delta = \cos(\eta)$ in the limit $\lambda \to 0$. Here, λ is imaginary and η is real in the gapless ($\Delta < 1$) regime, while λ is real and η is imaginary in the gapped ($\Delta > 1$) regime. We will here consider the latter case; however, a similar derivation applies for $\Delta < 1$ as well.

Setting this equal to $e^{i\phi/2}U'_{\mathrm{fSim}}$ (Eq. C.17), we require that:

$$ie^{i\phi/2}\sin(\theta) = \frac{\sin\lambda}{\sin(\lambda+\eta)} \tag{C.20}$$

Comparing the magnitudes and phases of the two sides of this equation, one finds, respectively:

$$\begin{aligned} \tan^2(\theta) &= -\frac{\sin^2(\lambda)}{\sin^2(\eta)}, \\ \tan\left(\frac{\phi}{2}\right) &= \frac{i\tan(\lambda)}{\tan(\eta)}. \end{aligned} \tag{C.21}$$

Eliminating λ and using $\Delta = \cos(\eta)$, we have:

$$\Delta^2\tan^2(\theta) = \tan^2(\phi/2)\left(1 + (1-\Delta^2)\tan^2(\theta)\right), \tag{C.22}$$

$$\Delta = \frac{\sin(\phi/2)}{\sin(\theta)} \tag{C.23}$$

BIBLIOGRAPHY

[1] We use Qiskit's implementation of SPSA, which is documented at `https://qiskit.org/documentation/stubs/qiskit.algorithms.optimizers.SPSA.html`.

[2] See Google's error correction roadmap at `https://quantumai.google/qecmilestone` (scroll down to the bottom of the page).

[3] See `https://quantum-computing.ibm.com/lab/docs/iql/manage/systems/properties`.

[4] Rajeev Acharya et al. "Suppressing quantum errors by scaling a surface code logical qubit". In: *Nature* 614.7949 (2023), pp. 676–681. DOI: `10.1038/s41586-022-05434-1`.

[5] F. Arute et al. "Quantum supremacy using a programmable superconducting processor". In: *Nature* 574 (2019), pp. 505–510. DOI: `10.1038/s41586-019-1666-5`.

[6] Abraham Asfaw et al. *Learn Quantum Computation Using Qiskit.* 2020. URL: `http://community.qiskit.org/textbook`.

[7] Ryan Babbush, Dominic W. Berry, and Hartmut Neven. "Quantum simulation of the Sachdev-Ye-Kitaev model by asymmetric qubitization". In: *Physical Review A* 99 (4 2019), p. 040301. DOI: `10.1103/PhysRevA.99.040301`.

[8] Bela Bauer et al. “Quantum Algorithms for Quantum Chemistry and Quantum Materials Science”. In: *Chemical Reviews* 120.22 (2020), 12685–12717. DOI: `10.1021/acs.chemrev.9b00829`.

[9] Matthew J. S. Beach et al. “Making trotters sprint: A variational imaginary time ansatz for quantum many-body systems”. In: *Physical Review B* 100.9 (2019). DOI: `10.1103/physrevb.100.094434`.

[10] Tomislav Begušić and Garnet Kin-Lic Chan. *Fast classical simulation of evidence for the utility of quantum computing before fault tolerance.* 2023. arXiv: `2306.16372 [quant-ph]`.

[11] Ewout van den Berg, Zlatko K. Minev, and Kristan Temme. “Model-free readout-error mitigation for quantum expectation values”. In: *Physical Review A* 105.3 (2022). DOI: `10.1103/physreva.105.032620`.

[12] Ewout van den Berg et al. “Probabilistic error cancellation with sparse Pauli–Lindblad models on noisy quantum processors”. In: *Nature Physics* (2023). DOI: `10.1038/s41567-023-02042-2`.

[13] B. Bertini et al. “Transport in Out-of-Equilibrium XXZ Chains: Exact Profiles of Charges and Currents”. In: *Physical Review Letters* 117 (20 2016), p. 207201. DOI: `10.1103/PhysRevLett.117.207201`.

[14] H. Bethe. “Zur Theorie der Metalle”. In: *Zeitschrift für Physik* 71.3 (1931), pp. 205–226. DOI: `10.1007/BF01341708`.

[15] Sergio Boixo et al. “Characterizing quantum supremacy in near-term devices”. In: *Nature Physics* 14.6 (2018), pp. 595–600. DOI: `10.1038/s41567-018-0124-x`.

[16] Sergey Bravyi et al. “Mitigating measurement errors in multiqubit experiments”. In: *Physical Review A* 103 (4 2021), p. 042605. DOI: `10.1103/PhysRevA.103.042605`.

[17] Sergey B. Bravyi and Alexei Yu. Kitaev. “Fermionic Quantum Computation”. In: *Annals of Physics* 298.1 (2002), 210–226. DOI: `10.1006/aphy.2002.6254`.

[18] Adam R. Brown et al. “Quantum Gravity in the Lab. I. Teleportation by Size and Traversable Wormholes”. In: *PRX Quantum* 4 (1 2023), p. 010320. DOI: `10.1103/PRXQuantum.4.010320`.

[19] Tiff Brydges et al. “Probing Rényi entanglement entropy via randomized measurements”. In: *Science* 364.6437 (2019), pp. 260–263. DOI: `10.1126/science.aau4963`.

[20] V. Bulchandani et al. “Superdiffusion in spin chains”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2021.8 (2021), p. 084001. DOI: `10.1088/1742-5468/ac12c7`.

[21] Vir B. Bulchandani. “Kardar-Parisi-Zhang universality from soft gauge modes”. In: *Physical Review B* 101 (4 2020), p. 041411. DOI: `10.1103/PhysRevB.101.041411`.

[22] Vir B. Bulchandani, David A. Huse, and Sarang Gopalakrishnan. "Onset of many-body quantum chaos due to breaking integrability". In: *Physical Review B* 105 (21 2022), p. 214308. DOI: `10.1103/PhysRevB.105.214308`.

[23] Vir B. Bulchandani et al. "Bethe-Boltzmann hydrodynamics and spin transport in the XXZ chain". In: *Physical Review B* 97 (4 2018), p. 045407. DOI: `10.1103/PhysRevB.97.045407`.

[24] Vir B. Bulchandani et al. "Solvable Hydrodynamics of Quantum Integrable Systems". In: *Physical Review Letters* 119 (22 2017), p. 220604. DOI: `10.1103/PhysRevLett.119.220604`.

[25] Elena Cáceres, Anderson Misobuchi, and Rafael Pimentel. "Sparse SYK and traversable wormholes". In: *Journal of High Energy Physics* 2021.11 (2021). DOI: `10.1007/jhep11(2021)015`.

[26] Zhenyu Cai and Simon C. Benjamin. "Constructing Smaller Pauli Twirling Sets for Arbitrary Error Channels". In: *Scientific Reports* 9.1 (2019). DOI: `10.1038/s41598-019-46722-7`.

[27] Zhenyu Cai, Xiaosi Xu, and Simon C. Benjamin. "Mitigating coherent noise using Pauli conjugation". In: *npj Quantum Information* 6.1 (2020). DOI: `10.1038/s41534-019-0233-0`.

[28] Olalla A. Castro-Alvaredo, Benjamin Doyon, and Takato Yoshimura. "Emergent Hydrodynamics in Integrable Quantum Systems Out of Equilibrium". In: *Physical Review X* 6 (4 Dec. 2016), p. 041065. DOI: `10.1103/PhysRevX.6.041065`.

[29] Jean-Sébastien Caux and Fabian H. L. Essler. “Time Evolution of Local Observables After Quenching to an Integrable Model”. In: *Physical Review Letters* 110 (25 2013), p. 257203. DOI: `10.1103/PhysRevLett.110.257203`.

[30] M. Cerezo et al. “Cost function dependent barren plateaus in shallow parametrized quantum circuits”. In: *Nature Communications* 12.1 (2021). DOI: `10.1038/s41467-021-21728-w`.

[31] Jianxin Chen et al. *Classical Simulation of Intermediate-Size Quantum Circuits*. 2018. arXiv: `1805.01450 [quant-ph]`.

[32] Yu Chen et al. “Qubit Architecture with High Coherence and Fast Tunable Coupling”. In: *Physical Review Letters* 113 (22 2014), p. 220502. DOI: `10.1103/PhysRevLett.113.220502`.

[33] Zijun Chen. “Metrology of Quantum Control and Measurement in Superconducting Qubits”. PhD thesis. UC Santa Barbara, 2018. URL: `https://www.proquest.com/dissertations-theses/metrology-quantum-control-measurement/docview/2023572021/se-2`.

[34] Jerry M. Chow et al. “Simple All-Microwave Entangling Gate for Fixed-Frequency Superconducting Qubits”. In: *Physical Review Letters* 107 (8 2011), p. 080502. DOI: `10.1103/PhysRevLett.107.080502`.

[35] Anirban Narayan Chowdhury and Rolando D. Somma. “Quantum Algorithms for Gibbs Sampling and Hitting-Time Estimation”. In: *Quantum Info. Comput.* 17.1–2 (2017), 41–64. DOI: `10.5555/3179483.3179486`.

[36] Pieter W. Claeys, Jonah Herzog-Arbeitman, and Austen Lamacraft. "Correlations and commuting transfer matrices in integrable unitary circuits". In: *SciPost Physics* 12 (2022), p. 007. DOI: `10.21468/SciPostPhys.12.1.007`.

[37] William Cottrell et al. "How to build the thermofield double state". In: *Journal of High Energy Physics* 2019.2 (2019). DOI: `10.1007/jhep02(2019)058`.

[38] Piotr Czarnik et al. "Error mitigation with Clifford quantum-circuit data". In: *Quantum* 5 (Nov. 2021), p. 592. DOI: `10.22331/q-2021-11-26-592`.

[39] Kedar Damle and Subir Sachdev. "Spin dynamics and transport in gapped one-dimensional Heisenberg antiferromagnets at nonzero temperatures". In: *Physical Review B* 57 (14 1998), pp. 8307–8339. DOI: `10.1103/PhysRevB.57.8307`.

[40] Christoph Dankert et al. "Exact and approximate unitary 2-designs and their application to fidelity estimation". In: *Physical Review A* 80.1 (2009). DOI: `10.1103/physreva.80.012304`.

[41] J. De Nardis et al. "Anomalous Spin Diffusion in One-Dimensional Antiferromagnets". In: *Physical Review Letters* 123 (18 2019), p. 186601. DOI: `10.1103/PhysRevLett.123.186601`.

[42] J. De Nardis et al. "Hydrodynamic diffusion in integrable systems". In: *Physical Review Letters* 121.16 (2018), p. 160603. DOI: `10.1103/PhysRevLett.121.160603`.

[43] Jacopo De Nardis et al. "Superdiffusion from emergent classical solitons in quantum spin chains". In: *Physical Review Letters* 125.7 (2020), p. 070601. DOI: `10.1103/PhysRevLett.125.070601`.

[44] B. Doyon and T. Yoshimura. "A note on generalized hydrodynamics: inhomogeneous fields and other concepts". In: *SciPost Physics* 2 (2017), p. 014. DOI: `10.21468/SciPostPhys.2.2.014`.

[45] Maxime Dupont and Joel E. Moore. "Universal spin dynamics in infinite-temperature one-dimensional quantum magnets". In: *Physical Review B* 101 (12 2020), p. 121106. DOI: `10.1103/PhysRevB.101.121106`.

[46] Maxime Dupont, Nicholas E. Sherman, and Joel E. Moore. "Spatiotemporal Crossover between Low- and High-Temperature Dynamical Regimes in the Quantum Heisenberg Magnet". In: *Physical Review Letters* 127 (10 2021), p. 107201. DOI: `10.1103/PhysRevLett.127.107201`.

[47] Bradley Efron. *The Jackknife, the Bootstrap and Other Resampling Plans*. Society for Industrial and Applied Mathematics, 1982. DOI: `10.1137/1.9781611970319`.

[48] A. Elben et al. "Rényi Entropies from Random Quenches in Atomic Hubbard and Spin Models". In: *Physical Review Letters* 120 (5 2018), p. 050406. DOI: `10.1103/PhysRevLett.120.050406`.

[49] Joseph Emerson, Robert Alicki, and Karol Życzkowski. "Scalable noise estimation with random unitary operators". In: *Journal of Optics B: Quan-*

tum and Semiclassical Optics 7.10 (2005), S347–S352. DOI: `10.1088/1464-4266/7/10/021`.

[50] Suguru Endo, Simon C Benjamin, and Ying Li. "Practical quantum error mitigation for near-future applications". In: *Physical Review X* 8.3 (2018), p. 031027. DOI: `10.1103/physrevx.8.031027`.

[51] S. J. van Enk and C. W. J. Beenakker. "Measuring Trρ^n on Single Copies of ρ Using Random Measurements". In: *Physical Review Letters* 108 (11 2012), p. 110503. DOI: `10.1103/PhysRevLett.108.110503`.

[52] S. Erne et al. "Universal dynamics in an isolated one-dimensional Bose gas far from equilibrium". In: *Nature* 563.7730 (2018), pp. 225–229. DOI: `10.1038/s41586-018-0667-0`.

[53] Glen Evenbly and Guifré Vidal. "Algorithms for entanglement renormalization". In: *Physical Review B* 79.14 (2009), p. 144108. DOI: `10.1103/PhysRevB.79.144108`.

[54] F Family and T Vicsek. "Scaling of the active zone in the Eden process on percolation networks and the ballistic deposition model". In: *Journal of Physics A: Mathematical and General* 18.2 (1985), p. L75. DOI: `10.1088/0305-4470/18/2/005`.

[55] Leo Fang et al. *NVIDIA/cuQuantum: cuQuantum Python v22.11.0.1*. Version v22.11.0.1. Jan. 2023. DOI: `10.5281/zenodo.7523366`.

[56] Q. Fontaine et al. “Observation of KPZ universal scaling in a one-dimensional polariton condensate”. In: *Nature* 608 (2022), pp. 687–691. DOI: `10.1038/s41586-022-05001-8`.

[57] B. Foxen et al. “Demonstrating a Continuous Set of Two-Qubit Gates for Near-Term Quantum Algorithms”. In: *Physical Review Letters* 125 (12 2020), p. 120504. DOI: `10.1103/PhysRevLett.125.120504`.

[58] Lena Funcke et al. “Measurement error mitigation in quantum computers through classical bit-flip correction”. In: *Physical Review A* 105 (6 2022), p. 062404. DOI: `10.1103/PhysRevA.105.062404`.

[59] Thierry Giamarchi. *Quantum Physics in One Dimension*. Oxford University Press, Dec. 2003. ISBN: 9780198525004. DOI: `10.1093/acprof:oso/9780198525004.001.0001`.

[60] Christian Gogolin and Jens Eisert. “Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems”. In: *Reports on Progress in Physics* 79.5 (2016), p. 056001. DOI: `10.1088/0034-4885/79/5/056001`.

[61] Google AI Quantum and collaborators. *Observation of separated dynamics of charge and spin in the Fermi-Hubbard model*. 2020. arXiv: `2010.07965 [quant-ph]`.

[62] S. Gopalakrishnan and R. Vasseur. “Kinetic Theory of Spin Diffusion and Superdiffusion in XXZ Spin Chains”. In: *Physical Review Letters* 122 (12 2019), p. 127202. DOI: `10.1103/PhysRevLett.122.127202`.

[63] S. Gopalakrishnan, R. Vasseur, and B. Ware. “Anomalous relaxation and the high-temperature structure factor of XXZ spin chains”. In: *Proceedings of the National Academy of Sciences* 116.33 (2019), pp. 16250–16255. DOI: `10.1073/pnas.1906914116`.

[64] Sarang Gopalakrishnan and Romain Vasseur. “Anomalous transport from hot quasiparticles in interacting spin chains”. In: *Reports on Progress in Physics* 86.3 (2023), p. 036502. DOI: `10.1088/1361-6633/acb36e`.

[65] Sarang Gopalakrishnan et al. *Distinct universality classes of diffusive transport from full counting statistics*. 2023. arXiv: `2203.09526 [cond-mat.stat-mech]`.

[66] Lov K. Grover. “A Fast Quantum Mechanical Algorithm for Database Search”. In: *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*. STOC ’96. Philadelphia, Pennsylvania, USA: Association for Computing Machinery, 1996, 212–219. DOI: `10.1145/237814.237866`.

[67] Jakob M Günther et al. “Improving readout in quantum simulations with repetition codes”. In: *Quantum Science and Technology* 7.1 (2021), p. 015009. DOI: `10.1088/2058-9565/ac3386`.

[68] Kathleen E. Hamilton et al. “Scalable quantum processor noise characterization”. In: *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*. 2020, pp. 430–440. DOI: `10.1109/QCE49297.2020.00060`.

[69] Alexander K. Hartmann et al. "High-precision simulation of the height distribution for the KPZ equation". In: *Europhysics Letters* 121.6 (2018), p. 67004. DOI: `10.1209/0295-5075/121/67004`.

[70] Andre He et al. "Zero-noise extrapolation for quantum-gate error mitigation with identity insertions". In: *Physical Review A* 102 (1 2020), p. 012426. DOI: `10.1103/PhysRevA.102.012426`.

[71] W. Heisenberg. "Zur Theorie des Ferromagnetismus". In: *Z. Phys.* 49.9 (1928), pp. 619–636. DOI: `10.1007/BF01328601`.

[72] Rebecca Hicks et al. "Active readout-error mitigation". In: *Physical Review A* 105 (1 2022), p. 012419. DOI: `10.1103/PhysRevA.105.012419`.

[73] Wen Wei Ho and Timothy H. Hsieh. "Efficient variational simulation of non-trivial quantum states". In: *SciPost Physics* 6.3 (2019). DOI: `10.21468/scipostphys.6.3.029`.

[74] Jesse C. Hoke et al. *Quantum information phases in space-time: measurement-induced entanglement and teleportation on a noisy quantum processor.* 2023. arXiv: `2303.04792 [quant-ph]`.

[75] William J. Huggins et al. "Virtual Distillation for Quantum Error Mitigation". In: *Phys. Rev. X* 11 (4 2021), p. 041036. DOI: `10.1103/PhysRevX.11.041036`.

[76] E. Ilievski et al. "Superuniversality of Superdiffusion". In: *Physical Review X* 11 (3 2021), p. 031023. DOI: `10.1103/PhysRevX.11.031023`.

[77] Enej Ilievski et al. "Superdiffusion in One-Dimensional Quantum Lattice Models". In: *Physical Review Letters* 121 (23 2018), p. 230602. DOI: `10.1103/PhysRevLett.121.230602`.

[78] Daniel Jafferis et al. "Traversable wormhole dynamics on a quantum processor". In: *Nature* 612.7938 (2022), pp. 51–55. DOI: `10.1038/s41586-022-05424-3`.

[79] P. N. Jepsen et al. "Spin transport in a tunable Heisenberg model realized with ultracold atoms". In: *Nature* 588.7838 (2020), pp. 403–407. DOI: `10.1038/s41586-020-3033-y`.

[80] Kurt Johansson. "Shape Fluctuations and Random Matrices". In: *Communications in Mathematical Physics* 209.2 (2000), pp. 437–476. DOI: `10.1007/s002200050027`.

[81] Sonika Johri, Damian S. Steiger, and Matthias Troyer. "Entanglement spectroscopy on a quantum computer". In: *Physical Review B* 96 (19 2017), p. 195136. DOI: `10.1103/PhysRevB.96.195136`.

[82] M. K. Joshi et al. "Observing emergent hydrodynamics in a long-range quantum magnet". In: *Science* 376.6594 (2022), pp. 720–724. DOI: `10.1126/science.abk2400`.

[83] Leo P Kadanoff. "Scaling and universality in statistical physics". In: *Physica A: Statistical Mechanics and its Applications* 163.1 (1990), pp. 1–14. DOI: `10.1016/0378-4371(90)90309-G`.

[84] Abhinav Kandala et al. “Error mitigation extends the computational reach of a noisy quantum processor”. In: *Nature* 567.7749 (2019), pp. 491–495. DOI: `10.1038/s41586-019-1040-7`.

[85] Abhinav Kandala et al. “Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets”. In: *Nature* 549.7671 (2017), 242–246. DOI: `10.1038/nature23879`.

[86] M. Kardar, G. Parisi, and Y.-C. Zhang. “Dynamic Scaling of Growing Interfaces”. In: *Physical Review Letters* 56 (9 1986), pp. 889–892. DOI: `10.1103/PhysRevLett.56.889`.

[87] K. Kechedzhi et al. *Effective quantum volume, fidelity and computational cost of noisy quantum processing experiments.* 2023. arXiv: `2306.15970 [quant-ph]`.

[88] Nathan Keenan et al. “Evidence of Kardar-Parisi-Zhang scaling on a digital quantum simulator”. In: *npj Quantum Information* 9.1 (2023). DOI: `10.1038/s41534-023-00742-4`.

[89] Hyungwon Kim and David A. Huse. “Ballistic Spreading of Entanglement in a Diffusive Nonintegrable System”. In: *Physical Review Letters* 111 (12 2013), p. 127205. DOI: `10.1103/PhysRevLett.111.127205`.

[90] Hyungwon Kim, Tatsuhiko N. Ikeda, and David A. Huse. “Testing whether all eigenstates obey the eigenstate thermalization hypothesis”. In: *Physical Review E* 90 (5 2014), p. 052105. DOI: `10.1103/PhysRevE.90.052105`.

[91] Youngseok Kim et al. "Evidence for the utility of quantum computing before fault tolerance". In: *Nature* 618.7965 (2023), pp. 500–505. DOI: `10.1038/s41586-023-06096-3`.

[92] A. Yu. Kitaev. *Quantum measurements and the Abelian Stabilizer Problem.* 1995. arXiv: `quant-ph/9511026 [quant-ph]`.

[93] Alexei Kitaev. "A simple model of quantum holography, part 1". In: (2015). URL: `http://online.kitp.ucsb.edu/online/entangled15/kitaev/`.

[94] Alexei Kitaev. "A simple model of quantum holography, part 2". In: (2015). URL: `http://online.kitp.ucsb.edu/online/entangled15/kitaev2/`.

[95] Alexei Kitaev. "Hidden correlations in the Hawking radiation and thermal noise". In: (2015). URL: `http://online.kitp.ucsb.edu/online/joint98/kitaev/`.

[96] Jens Koch et al. "Charge-insensitive qubit design derived from the Cooper pair box". In: *Physical Review A* 76 (4 2007), p. 042319. DOI: `10.1103/PhysRevA.76.042319`.

[97] Bálint Koczor. "Exponential Error Suppression for Near-Term Quantum Devices". In: *Physical Review X* 11.3 (2021). DOI: `10.1103/physrevx.11.031057`.

[98] Ž. Krajnik, Enej Ilievski, and Tomaž Prosen. *Universal distributions of magnetization transfer in integrable spin chains.* 2023. arXiv: `2303.16691 [cond-mat.stat-mech]`.

[99] Ziga Krajnik, E. Ilievski, and T. Prosen. “Absence of Normal Fluctuations in an Integrable Magnet”. In: *Physical Review Letters* 128 (9 2022), p. 090604. DOI: `10.1103/PhysRevLett.128.090604`.

[100] Ziga Krajnik, Enej Ilievski, and Tomaž Prosen. “Integrable matrix models in discrete space-time”. In: *SciPost Physics* 9 (2020), p. 038. DOI: `10.21468/SciPostPhys.9.3.038`.

[101] Haggai Landa et al. “Experimental Bayesian estimation of quantum state preparation, measurement, and gate errors in multiqubit devices”. In: *Phys. Rev. Res.* 4 (1 2022), p. 013199. DOI: `10.1103/PhysRevResearch.4.013199`.

[102] Ying Li and Simon C. Benjamin. “Efficient Variational Quantum Simulator Incorporating Active Error Minimization”. In: *Physical Review X* 7 (2 2017), p. 021050. DOI: `10.1103/PhysRevX.7.021050`.

[103] Jürgen Lisenfeld, Alexander Bilmes, and Alexey V. Ustinov. “Enhancing the coherence of superconducting quantum bits with electric fields”. In: *npj Quantum Information* 9.1 (2023). DOI: `10.1038/s41534-023-00678-9`.

[104] M. Ljubotina, M. Žnidarič, and Tomaž Prosen. “Spin diffusion from an inhomogeneous quench in an integrable system”. In: *Nature Communications* 8.1 (2017), pp. 1–6. DOI: `10.1038/ncomms16117`.

[105] Marko Ljubotina, Marko Žnidarič, and Tomaž Prosen. “A class of states supporting diffusive spin dynamics in the isotropic Heisenberg model”. In: *Journal of Physics A: Mathematical and Theoretical* 50.47 (2017), p. 475002. DOI: `10.1088/1751-8121/aa8bdc`.

[106] Marko Ljubotina, Marko Žnidarič, and Tomaž Prosen. “Kardar-Parisi-Zhang Physics in the Quantum Heisenberg Magnet”. In: *Physical Review Letters* 122 (21 2019), p. 210602. DOI: `10.1103/PhysRevLett.122.210602`.

[107] Marko Ljubotina, Lenart Zadnik, and Tomaž Prosen. “Ballistic Spin Transport in a Periodically Driven Integrable Quantum System”. In: *Physical Review Letters* 122 (15 2019), p. 150605. DOI: `10.1103/PhysRevLett.122.150605`.

[108] Zhihuang Luo et al. “Quantum Simulation of the Non-Fermi-Liquid State of Sachdev-Ye-Kitaev Model”. In: *npj Quantum Inf.* 5 (2019), p. 53. DOI: `10.1038/s41534-019-0166-7`.

[109] J. Lux et al. “Hydrodynamic long-time tails after a quantum quench”. In: *Physical Review A* 89 (5 2014), p. 053608. DOI: `10.1103/PhysRevA.89.053608`.

[110] Filip B. Maciejewski et al. “Modeling and mitigation of cross-talk effects in readout noise with applications to the Quantum Approximate Optimization Algorithm”. In: *Quantum* 5 (2021), p. 464. DOI: `10.22331/q-2021-06-01-464`.

[111] Juan Maldacena, Stephen H. Shenker, and Douglas Stanford. “A bound on chaos”. In: *Journal of High Energy Physics* 2016.8 (2016). DOI: `10.1007/jhep08(2016)106`.

[112] Juan Maldacena and Douglas Stanford. “Remarks on the Sachdev-Ye-Kitaev model”. In: *Phys. Rev. D* 94.10 (2016), p. 106002. DOI: `10.1103/PhysRevD.94.106002`.

[113] Juan Maldacena, Douglas Stanford, and Zhenbin Yang. “Conformal symmetry and its breaking in two-dimensional nearly anti-de Sitter space”. In: *Progress of Theoretical and Experimental Physics* 2016.12 (Nov. 2016). 12C104. DOI: `10.1093/ptep/ptw124`.

[114] John Martyn and Brian Swingle. “Product spectrum ansatz and the simplicity of thermal states”. In: *Physical Review A* 100 (3 2019), p. 032107. DOI: `10.1103/PhysRevA.100.032107`.

[115] X. Mi et al. *Stable Quantum-Correlated Many Body States via Engineered Dissipation*. 2023. arXiv: `2304.13878 [quant-ph]`.

[116] Xiao Mi et al. “Information scrambling in quantum circuits”. In: *Science* 374.6574 (2021), pp. 1479–1483. DOI: `10.1126/science.abg5029`.

[117] Grégoire Misguich, Kirone Mallick, and P. L. Krapivsky. “Dynamics of the spin-$\frac{1}{2}$ Heisenberg chain initialized in a domain-wall state”. In: *Physical Review B* 96 (19 2017), p. 195151. DOI: `10.1103/PhysRevB.96.195151`.

[118] Ashley Montanaro and Stasja Stanisic. *Error mitigation by training with fermionic linear optics*. 2021. arXiv: `2102.02120 [quant-ph]`.

[119] A. Morvan et al. “Formation of robust bound states of interacting microwave photons”. In: *Nature* 612.7939 (2022), pp. 240–245. DOI: `10.1038/s41586-022-05348-y`.

[120] A. Morvan et al. *Phase transition in Random Circuit Sampling*. 2023. arXiv: `2304.11119 [quant-ph]`.

[121] F. Motzoi et al. "Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits". In: *Physical Review Letters* 103 (11 2009), p. 110501. DOI: `10.1103/PhysRevLett.103.110501`.

[122] Kouhei Nakaji and Naoki Yamamoto. "Expressibility of the alternating layered ansatz for quantum computation". In: *Quantum* 5 (2021), p. 434. DOI: `10.22331/q-2021-04-19-434`.

[123] Jacopo De Nardis, Sarang Gopalakrishnan, and Romain Vasseur. *Non-linear fluctuating hydrodynamics for KPZ scaling in isotropic spin chains*. 2023. arXiv: `2212.03696 [cond-mat.quant-gas]`.

[124] Paul Nation. *Excited State Promotion (ESP) Readout*. 2021. URL: `https://nonhermitian.org/posts/2021/2021-12-20-esp_readout.html`.

[125] Paul Nation. *Improving state prep errors on IBM Quantum systems*. 2021. URL: `https://nonhermitian.org/posts/2021/2021-11-07-rep_delay.html`.

[126] Paul D. Nation et al. "Scalable Mitigation of Measurement Errors on Quantum Computers". In: *PRX Quantum* 2 (4 2021), p. 040326. DOI: `10.1103/PRXQuantum.2.040326`.

[127] C. Neill et al. "Accurately computing the electronic properties of a quantum ring". In: *Nature* 594 (2021), 508—512. DOI: `10.1038/s41586-021-03576-2`.

[128] Charles J. Neill. “A Path Towards Quantum Supremacy with Superconducting Qubits”. PhD thesis. 2017, p. 309. URL: `https://www.proquest.com/dissertations-theses/path-towards-quantum-supremacy-with/docview/2024216315/se-2`.

[129] Sepehr Nezami et al. “Quantum Gravity in the Lab. II. Teleportation by Size and Traversable Wormholes”. In: *PRX Quantum* 4 (1 2023), p. 010321. DOI: `10.1103/PRXQuantum.4.010321`.

[130] G. Odor. *Universality In Nonequilibrium Lattice Systems: Theoretical Foundations*. World Scientific publishing Co., 2008. DOI: `10.1142/6813`.

[131] Matthew Otten and Stephen K. Gray. “Recovering noise-free quantum observables”. In: *Physical Review A* 99 (1 2019), p. 012338. DOI: `10.1103/PhysRevA.99.012338`.

[132] Evan Peters, Andy C. Y. Li, and Gabriel N. Perdue. “Perturbative readout-error mitigation for near-term quantum computers”. In: *Phys. Rev. A* 107 (6 2023), p. 062426. DOI: `10.1103/PhysRevA.107.062426`.

[133] Pierre Pfeuty. “The one-dimensional Ising model with a transverse field”. In: *Annals of Physics* 57.1 (1970), pp. 79–90. DOI: `10.1016/0003-4916(70)90270-8`.

[134] Michael Plischke and Zoltán Rácz. “Dynamic scaling and the surface structure of Eden clusters”. In: *Physical Review A* 32 (6 1985), pp. 3825–3828. DOI: `10.1103/PhysRevA.32.3825`.

[135] A. Polkovnikov et al. "Colloquium: Nonequilibrium dynamics of closed interacting quantum systems". In: *Reviews of Modern Physics* 83 (3 2011), pp. 863–883. DOI: `10.1103/RevModPhys.83.863`.

[136] K. Popper. *Conjectures and Refutations: The Growth of Scientific Knowledge.* Routledge Classics, 2002.

[137] Michael Prähofer and Herbert Spohn. "Exact Scaling Functions for One-Dimensional Stationary KPZ Growth". In: *Journal of Statistical Physics* 115.1 (2004), pp. 255–279. DOI: `10.1023/B:JOSS.0000019810.21828.fc`.

[138] Michael Prähofer and Herbert Spohn. "Universal Distributions for Growth Processes in $1 + 1$ Dimensions and Random Matrices". In: *Physical Review Letters* 84 (21 2000), pp. 4882–4885. DOI: `10.1103/PhysRevLett.84.4882`.

[139] Tomaž Prosen. "Open XXZ Spin Chain: Nonequilibrium Steady State and a Strict Bound on Ballistic Transport". In: *Physical Review Letters* 106 (21 2011), p. 217206. DOI: `10.1103/PhysRevLett.106.217206`.

[140] Quantum AI team and collaborators. *qsim.* Sept. 2020. DOI: `10.5281/zenodo.4023103`.

[141] M. H. Quenouille. "Approximate Tests of Correlation in Time-Series". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 11.1 (1949), pp. 68–84. DOI: `10.1111/j.2517-6161.1949.tb00023.x`.

[142] Eliott Rosenberg, Paul Ginsparg, and Peter L McMahon. "Experimental error mitigation using linear rescaling for variational quantum eigensolving with up

to 20 qubits". In: *Quantum Science and Technology* 7.1 (2022), p. 015024. DOI: `10.1088/2058-9565/ac3b37`.

[143] Eliott Rosenberg et al. *Dynamics of magnetization at infinite temperature in a Heisenberg spin chain*. 2023. arXiv: `2306.09333 [quant-ph]`.

[144] S. Sachdev. *Quantum Phase Transitions*. Cambridge University Press, 2011. DOI: `10.1017/CBO9780511973765`.

[145] Subir Sachdev and Kedar Damle. "Low Temperature Spin Diffusion in the One-Dimensional Quantum $O(3)$ Nonlinear σ Model". In: *Physical Review Letters* 78 (5 1997), pp. 943–946. DOI: `10.1103/PhysRevLett.78.943`.

[146] Rhine Samajdar et al. *Quantum turnstiles for robust measurement of full counting statistics*. 2023. arXiv: `2305.15464 [quant-ph]`.

[147] Ramsés J Sánchez, Vipin Kerala Varma, and V. Oganesyan. "Anomalous and regular transport in spin-$\frac{1}{2}$ chains: AC conductivity". In: *Physical Review B* 98.5 (2018), p. 054415. DOI: `10.1103/PhysRevB.98.054415`.

[148] Daniel Sank. "Fast, Accurate State Measurement in Superconducting Qubits". PhD thesis. UC Santa Barbara, 2014. URL: `https://www.proquest.com/dissertations-theses/fast-accurate-state-measurement-superconducting/docview/1629024078/se-2`.

[149] K. J. Satzinger et al. "Realizing topologically ordered states on a quantum processor". In: *Science* 374.6572 (2021), pp. 1237–1241. DOI: `10.1126/science.abi8378`.

[150] A Scheie et al. “Detection of Kardar–Parisi–Zhang hydrodynamics in a quantum Heisenberg spin-1/2 chain”. In: *Nature Physics* 17.6 (2021), pp. 726–730. DOI: `10.1038/s41567-021-01191-6`.

[151] Illya Shapoval et al. *Towards Quantum Gravity in the Lab on Quantum Processors*. May 2022. arXiv: `2205.14081 [quant-ph]`.

[152] Andrew Shaw. *Classical-Quantum Noise Mitigation for NISQ Hardware*. 2021. arXiv: `2105.08701 [quant-ph]`.

[153] P.W. Shor. “Algorithms for quantum computation: discrete logarithms and factoring”. In: *Proceedings 35th Annual Symposium on Foundations of Computer Science*. 1994, pp. 124–134. DOI: `10.1109/SFCS.1994.365700`.

[154] Alistair W. R. Smith et al. “Qubit readout error mitigation with bit-flip averaging”. In: *Science Advances* 7.47 (2021), eabi8009. DOI: `10.1126/sciadv.abi8009`.

[155] A. Sommer et al. “Universal spin transport in a strongly interacting Fermi gas”. In: *Nature* 472.7342 (2011), pp. 201–204. DOI: `10.1038/nature09989`.

[156] Jean-Marie Stéphan. “Return probability after a quench from a domain wall initial state in the spin-1/2 XXZ chain”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2017.10 (2017), p. 103108. DOI: `10.1088/1742-5468/aa8c19`.

[157] Vincent Paul Su. “Variational preparation of the thermofield double state of the Sachdev-Ye-Kitaev model”. In: *Physical Review A* 104.1 (2021). DOI: `10.1103/physreva.104.012427`.

[158] K. Temme et al. “Quantum Metropolis sampling”. In: *Nature* 471.7336 (2011), 87–90. DOI: `10.1038/nature09770`.

[159] Kristan Temme, Sergey Bravyi, and Jay M Gambetta. “Error mitigation for short-depth quantum circuits”. In: *Physical Review Letters* 119.18 (2017), p. 180509. DOI: `10.1103/physrevlett.119.180509`.

[160] Joseph Tindall et al. *Efficient tensor network simulation of IBM’s kicked Ising experiment.* 2023. arXiv: `2306.14887 [quant-ph]`.

[161] Matthew Treinish et al. *Qiskit: An Open-source Framework for Quantum Computing.* 2021. DOI: `10.5281/zenodo.2573505`.

[162] John W. Tukey. “Bias and confidence in not-quite large samples”. In: *The Annals of Mathematical Statistics* 29.2 (1958), p. 614. DOI: `10.1214/aoms/1177706647`.

[163] Miroslav Urbanek et al. “Mitigating Depolarizing Noise on Quantum Computers with Noise-Estimation Circuits”. In: *Phys. Rev. Lett.* 127 (27 2021), p. 270502. DOI: `10.1103/PhysRevLett.127.270502`.

[164] Matthieu Vanicat, Lenart Zadnik, and Tomaž Prosen. “Integrable Trotterization: Local Conservation Laws and Boundary Driving”. In: *Physical Review Letters* 121 (3 2018), p. 030606. DOI: `10.1103/PhysRevLett.121.030606`.

[165] G. Vidal. “Class of Quantum Many-Body States That Can Be Efficiently Simulated”. In: *Physical Review Letters* 101 (11 2008), p. 110501. DOI: `10.1103/PhysRevLett.101.110501`.

[166] Guifré Vidal. "Efficient Classical Simulation of Slightly Entangled Quantum Computations". In: *Physical Review Letters* 91.14 (2003). DOI: `10.1103/PhysRevLett.91.147902`.

[167] Jean-Loup Ville et al. "Leveraging randomized compiling for the quantum imaginary-time-evolution algorithm". In: *Phys. Rev. Res.* 4 (3 2022), p. 033140. DOI: `10.1103/PhysRevResearch.4.033140`.

[168] Joseph Vovrosh et al. "Simple mitigation of global depolarizing errors in quantum simulations". In: *Physical Review E* 104.3 (2021). DOI: `10.1103/physreve.104.035309`.

[169] M. Žnidarič. "Spin Transport in a One-Dimensional Anisotropic Heisenberg Model". In: *Physical Review Letters* 106 (22 2011), p. 220601. DOI: `10.1103/PhysRevLett.106.220601`.

[170] Joel J. Wallman and Joseph Emerson. "Noise tailoring for scalable quantum computation via randomized compiling". In: *Physical Review A* 94.5 (2016). DOI: `10.1103/physreva.94.052325`.

[171] D. Wei et al. "Quantum gas microscopy of Kardar-Parisi-Zhang superdiffusion". In: *Science* 376.6594 (2022), pp. 716–720. DOI: `10.1126/science.abk2397`.

[172] Jingxiang Wu and Timothy H. Hsieh. "Variational Thermal Quantum Simulation via Thermofield Double States". In: *Physical Review Letters* 123.22 (2019). DOI: `10.1103/physrevlett.123.220502`.

[173] Y. Wu et al. "Strong Quantum Computational Advantage Using a Superconducting Quantum Processor". In: *Physical Review Letters* 127 (18 2021), p. 180501. DOI: `10.1103/PhysRevLett.127.180501`.

[174] Yulin Wu et al. "Strong Quantum Computational Advantage Using a Superconducting Quantum Processor". In: *Phys. Rev. Lett.* 127 (18 2021), p. 180501. DOI: `10.1103/PhysRevLett.127.180501`.

[175] Jonathan Wurtz and Anatoli Polkovnikov. "Emergent conservation laws and nonthermal states in the mixed-field Ising model". In: *Physical Review B* 101 (19 2020), p. 195138. DOI: `10.1103/PhysRevB.101.195138`.

[176] Shenglong Xu et al. *A Sparse Model of Quantum Holography*. 2020. arXiv: `2008.02303 [cond-mat.str-el]`.

[177] Fei Yan et al. "Tunable Coupling Scheme for Implementing High-Fidelity Two-Qubit Gates". In: *Phys. Rev. Appl.* 10 (5 2018), p. 054062. DOI: `10.1103/PhysRevApplied.10.054062`.

[178] Bingtian Ye et al. "Universal Kardar-Parisi-Zhang Dynamics in Integrable Quantum Systems". In: *Physical Review Letters* 129 (23 2022), p. 230602. DOI: `10.1103/PhysRevLett.129.230602`.

[179] Xinyuan You et al. "Stabilizing and Improving Qubit Coherence by Engineering the Noise Spectrum of Two-Level Systems". In: *Physical Review Appl.* 18 (4 2022), p. 044026. DOI: `10.1103/PhysRevApplied.18.044026`.

[180] M.-H. Yung and A. Aspuru-Guzik. “A quantum-quantum Metropolis algorithm”. In: *Proceedings of the National Academy of Sciences* 109.3 (2012), 754–759. DOI: `10.1073/pnas.1111758109`.

[181] D. Zhu et al. “Generation of thermofield double states and critical ground states with a quantum computer”. In: *Proceedings of the National Academy of Sciences* 117.41 (2020), pp. 25402–25406. DOI: `10.1073/pnas.2006337117`.

[182] Qingling Zhu et al. “Quantum computational advantage via 60-qubit 24-cycle random circuit sampling”. In: *Science Bulletin* 67.3 (2022), pp. 240–245. DOI: `10.1016/j.scib.2021.10.017`.

[183] X. Zotos. “Finite Temperature Drude Weight of the One-Dimensional Spin-1/2 Heisenberg Model”. In: *Physical Review Letters* 82 (8 1999), pp. 1764–1767. DOI: `10.1103/PhysRevLett.82.1764`.

[184] C. Zu et al. “Emergent hydrodynamics in a strongly interacting dipolar spin ensemble”. In: *Nature* 597.7874 (2021), pp. 45–50. DOI: `10.1038/s41586-021-03763-1`.