Measurement error mitigation in quantum computers through classical bit-flip correction

Lena Funcke ,[1] Tobias Hartung ,[2] Karl Jansen,[3] Stefan Kühn ,[4] Paolo Stornati ,[3,5] and Xiaoyang Wang [6]
[1] Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON, Canada N2L 2Y5
[2] Department of Mathematics, King's College London, Strand, London WC2R 2LS, United Kingdom
[3] NIC, DESY Zeuthen, Platanenallee 6, 15738 Zeuthen, Germany
[4] Computation-Based Science and Technology Research Center, The Cyprus Institute, 20 Kavafi Street, 2121 Nicosia, Cyprus
[5] Institut für Physik, Humboldt-Universität zu Berlin, Zum Großen Windkanal 6, D-12489 Berlin, Germany
[6] School of Physics, Peking University, 5 Yiheyuan Road, Haidian District, Beijing 100871, China

(Received 15 September 2021; accepted 23 February 2022; published 3 June 2022)

We develop a classical bit-flip correction method to mitigate measurement errors on quantum computers. This method can be applied to any operator, any number of qubits, and any realistic bit-flip probability. We first demonstrate the successful performance of this method by correcting the noisy measurements of the ground-state energy of the longitudinal Ising model. We then generalize our results to arbitrary operators and test our method both numerically and experimentally on IBM quantum hardware. As a result, our correction method reduces the measurement error on the quantum hardware by up to one order of magnitude. We finally discuss how to preprocess the method and extend it to other error sources beyond measurement errors. For local Hamiltonians, the overhead costs are polynomial in the number of qubits, even if multiqubit correlations are included.

DOI: 10.1103/PhysRevA.105.062404

I. INTRODUCTION

Quantum computers have the potential to outperform classical computers in a variety of tasks ranging from combinatorial optimization [1,2] over cryptography [3,4] to machine learning [5,6]. In particular, the prospect of being able to efficiently simulate quantum systems makes them a promising tool for solving quantum many-body problems in physics and chemistry. Despite recent progress, a large-scale, fault-tolerant digital quantum computer is still not available, and current intermediate scale devices suffer from a considerable level of noise. Although this limits the depth of the circuits that can be executed faithfully, these noisy intermediate-scale quantum (NISQ) devices [7] are already able to exceed the capabilities of classical computes in certain cases [8].

In the context of quantum many-body systems, a promising approach for exploiting the power of NISQ devices is variational quantum simulation (VQS), a class of hybrid quantum-classical algorithms for solving optimization problems [9,10]. These make use of a feedback loop between a classical computer and a quantum coprocessor; the latter is used to efficiently evaluate the cost function for a given set of variational parameters, which are optimized on a classical computer based on the measurement outcome obtained from the quantum coprocessor. In particular, it has been experimentally demonstrated that VQS allows for finding both the ground state and low-lying excitations of systems relevant for condensed matter and particle physics as well as quantum chemistry [11–20].

NISQ devices are susceptible to errors, which can only be partially mitigated using error-mitigation procedures (see, e.g., Refs. [12,21–41]). In particular, the qubit measurement is among the most error-prone operations on NISQ devices, with error rates ranging from 8% to 30% for current hardware [38]. These errors arise from bit flips, i.e., from erroneously recording an outcome as 0 given it was actually 1, and vice versa.

The goal of this paper is to mitigate these types of measurement errors, in principle, for any operator, any number of qubits, and any bit-flip probability. We develop an efficient mitigation method that relies on cancellations of different erroneous measurement outcomes. This cancellation results from relative minus signs stemming from the default measurement basis of current hardware, $Z = \mathrm{diag}(1, -1)$. The only input requirement for this approach is the knowledge of the different bit-flip probabilities during readout for each qubit. Our method mainly focuses on measurement bit flips that are uncorrelated between the qubits for multiqubit measurements, which is true in good approximation in many cases (see, e.g., Refs. [42–44]). However, our method can also be extended to multiqubit correlations and different error sources beyond measurement errors, as we discuss in the end of the paper.

Our paper is organized as follows. In Sec. II, we demonstrate the performance of our mitigation method by correcting the noisy energy histograms for the longitudinal Ising (LI) model [the transversal Ising (TI) model is discussed in Appendix A 5]. For simplicity, we assume all bit-flip probabilities to be equal. In Sec. III, we generalize our method to different bit-flip probabilities and arbitrary operators. We now correct each bit flip directly at the measurement step, which allows us to mitigate the measurement errors of any expectation value of any operator. In Sec. IV, we demonstrate the experimental applicability of our method on IBM quantum hardware. In Sec. V, we discuss our results and compare them to previous work. Moreover, we comment on the inclusion of multiqubit correlations, provide an extension of our method to mitigate relaxation errors, work out a probabilistic implementation of our method, and finally discuss preprocessing and overhead costs. In Sec. VI, we summarize our results.

II. MITIGATION OF MEASUREMENT ERRORS FOR ENERGY HISTOGRAMS

Throughout this article, we focus on classical bit-flip errors (referred to as measurement or readout errors) and neglect any other sources of error, such as gate errors and decoherence. Thus, we assume that the quantum device prepares a pure state $|\psi\rangle$ for N qubits, which we measure in the computational basis

$$|\psi\rangle = \sum_{i=0}^{2^N-1} c_i|i\rangle. \tag{1}$$

Here, $|i\rangle$ is a shorthand notation for the computational-basis state corresponding to a bit string for the binary representation of i (e.g., for $N = 4$ the state $|5\rangle$ corresponds to $|0101\rangle$). A perfect, noise-free projective measurement would thus yield the bit string q with probability $|c_i|^2$; however, bit flips during readout can lead to erroneously recording $j \neq i$ instead. Throughout the main body of this article, we make the assumption that each bit flips independently of the others, which is a good approximation on current quantum hardware (see, e.g., Ref. [43]). Eventually, we will discuss in Sec. V how to relax this assumption and include multiqubit correlations into our method.

Our goal is to obtain the expectation value $\langle\psi|\mathcal{H}|\psi\rangle$ for a given Hamiltonian $\mathcal{H}$ from a quantum device. Without loss of generality, we assume that $\mathcal{H}$ is of the form

$$\mathcal{H} = \sum_k h_k U_k^* O_k U_k, \tag{2}$$

where O_k is a string of the Pauli matrices $\mathbb{1}$ and Z acting on N qubits, and the unitary U_k transforms this string to $U_k^* O_k U_k \in \{\mathbb{1}, X, Y, Z\}^{\otimes N}$. Note that throughout the paper, we denote the adjoint of operators with asterisks. Since in an experiment we can only measure the final state in the Z basis, we cannot directly obtain $\langle\psi|\mathcal{H}|\psi\rangle$. We have to determine instead the expectation values of individual Pauli strings O_k by applying the postrotation U_k to $|\psi\rangle$. Subsequently, we can correlate O_k against the distribution of bit strings obtained from the measurement. Thus, we focus throughout the paper on Pauli strings of the form $\{\mathbb{1}, Z\}^{\otimes N}$. Moreover, in the following we assume that each summand $U_k^* O_k U_k$ in Eq. (2) is measured separately. For efficient implementations, multiple summands can also be measured simultaneously, which will be considered later (see Sec. III E).

To obtain the distribution of bit strings, we have to execute the quantum circuit preparing $U_k|\psi\rangle$ a number of times and record the measurement outcome for each run. Throughout the paper, we refer to this number of repetitions as the number of shots s.

A. Prediction for the longitudinal Ising model

As a pedagogical introductory example that illustrates the basic idea of our method, let us briefly analyze the noisy energy histograms of the LI model with periodic boundary conditions. For this, we assume for simplicity that all bit-flip probabilities are equal, $p(|0\rangle \to |1\rangle) = p(|1\rangle \to |0\rangle) =: p$, for all qubits. We will explain all technical details of this example in Appendix A, and we will also discuss the TI model in Appendix A 5. We will turn to the more general case in Sec. III, where we will discuss different bit-flip probabilities, arbitrary operators, and arbitrary (pure or mixed) states.

The Hamiltonian of the LI model reads

$$\mathcal{H}_{\mathrm{LI}} = J\sum_{q=1}^{N} Z_q Z_{q+1} + h\sum_{q=1}^{N} Z_q, \tag{3}$$

where we assume $J < 0$ and $h > 0$ and we identify $N + 1$ with 1. The true ground-state energy of the model is

$$E_0 = E_{ZZ} + E_Z = NJ - Nh, \tag{4}$$

which is the sum of the individual ground-state energies for $h = 0$ and $J = 0$, which we call E_{ZZ} and E_Z, respectively.

Now we wish to determine the expectation $\mathbb{E}$ of the noisy ground-state energy $\tilde{E}_0$ measured on a quantum computer, where the tilde denotes a noisy outcome. We note that "expectation" here means the expectation with respect to the bit-flip probability p, which should not be confused with the quantum-mechanical expectation value of the Hamiltonian, $\langle\psi|\mathcal{H}|\psi\rangle = E$. Thus, the expectation $\mathbb{E}\tilde{\mathcal{H}}$ is the expected value (as an operator to be measured subject to bit flips; see also Sec. III) for the noisy Hamiltonian $\tilde{\mathcal{H}}$, while $\mathbb{E}\langle\psi|\tilde{\mathcal{H}}|\psi\rangle = \mathbb{E}\tilde{E}$ is the expected value for the noisy (quantum-mechanical) expectation value $\langle\psi|\tilde{\mathcal{H}}|\psi\rangle = \tilde{E}$.

To determine the noisy expectation of E_0 in Eq. (4), we will first discuss a single Z_q operator, then a single $Z_q Z_{q+1}$ operator, and finally we will take the sum over all qubits to recover the LI model. Starting with a single Z_q operator, we notice the following:

(i) If there are no bit flips and both possible measurement outcomes for the qubit are recorded correctly, i.e., $|0\rangle \xrightarrow{1-p} |0\rangle$, $|1\rangle \xrightarrow{1-p} |1\rangle$, we measure the true expectation value $\langle\psi|Z|\psi\rangle$ with probability $(1-p)^2$.

(ii) If there are two bit flips and both measurement outcomes are recorded incorrectly, i.e., $|0\rangle \xrightarrow{p} |1\rangle$, $|1\rangle \xrightarrow{p} |0\rangle$, we measure the negative expectation value $-\langle\psi|Z|\psi\rangle$ (due to $\langle 1|Z|1\rangle = -\langle 0|Z|0\rangle$) with probability p^2.

(iii) If there are single bit flips and one possible measurement outcome is recorded correctly, while the other one is recorded incorrectly, i.e., $|0\rangle \xrightarrow{p} |1\rangle$, $|1\rangle \xrightarrow{1-p} |1\rangle$ or $|0\rangle \xrightarrow{1-p} |0\rangle$, $|1\rangle \xrightarrow{p} |0\rangle$, we measure outcomes with opposite signs that cancel identically.

Thus, in total we get the expectation

$$\begin{aligned}\mathbb{E}\langle\psi|\tilde{Z}|\psi\rangle &= (1-p)^2\langle\psi|Z|\psi\rangle + p^2(-\langle\psi|Z|\psi\rangle)\\ &= (1-2p)\langle\psi|Z|\psi\rangle.\end{aligned} \tag{5}$$

For a single $Z_q Z_{q+1}$ operator, we get three different nonzero outcomes:

(i) The absence of any bit flip gives the true expectation value $\langle\psi|Z_qZ_{q+1}|\psi\rangle$ with probability $(1-p)^2$, just as before.

(ii) Total bit flips, $|0\rangle \xrightarrow{p} |1\rangle$ and $|1\rangle \xrightarrow{p} |0\rangle$ for both qubits, also give $\langle\psi|Z_qZ_{q+1}|\psi\rangle$ (due to $\langle 00|Z_1Z_2|00\rangle = \langle 11|Z_1Z_2|11\rangle$) with probability p^2, unlike before.

(iii) Total bit flips for *one* qubit but no bit flip for the *other* qubit gives the negative expectation value $-\langle\psi|Z_qZ_{q+1}|\psi\rangle$ with a combined probability of $p(1-p)+(1-p)p = 2p(1-p)$.

All other possible outcomes cancel identically, similar to the third case discussed previously for the $\langle\psi|Z|\psi\rangle$ case. In total, this yields

$$\begin{aligned}
&\mathbb{E}\langle\psi|\tilde{Z}_q\tilde{Z}_{q+1}|\psi\rangle \\
&\quad= (1-p)^2\langle\psi|Z_qZ_{q+1}|\psi\rangle + p^2\langle\psi|Z_qZ_{q+1}|\psi\rangle \\
&\qquad+ 2p(1-p)(-\langle\psi|Z_qZ_{q+1}|\psi\rangle) \\
&\quad= (1-2p)^2\langle\psi|Z_qZ_{q+1}|\psi\rangle. \qquad (6)
\end{aligned}$$

A more detailed derivation of these results can be found in Appendix A and Sec. III C.

Finally, to derive the noisy expectation of the full ground-state energy E_0 in Eq. (4), we can sum Eqs. (5) and (6) over the N different qubits. Thus, the final result for the LI model reads

$$\mathbb{E}\tilde{E}_0 = (1-2p)E_Z + (1-2p)^2E_{ZZ}. \qquad (7)$$

Our method allows us to predict the variance of the noisy energy histograms as well, as we will explain in detail in Sec. III E and Appendix B. Based on these results, Fig. 1 shows the resulting energy histograms for the ground state of $\mathcal{H}_{\rm LI}$ with different choices of the parameters N, J, h, s, and p, where we measure the ground state 2048 times for each parameter combination. The noise model, with the mean energy from Eq. (7) and the variance from Eq. (B17), agrees with the data for all the parameters. Indeed, our prediction (solid orange line in Fig. 1) perfectly matches the fitted data of the histogram (dashed black line). This allows us to retrieve the true ground-state energy E_0 (dashed green line) using Eq. (7).

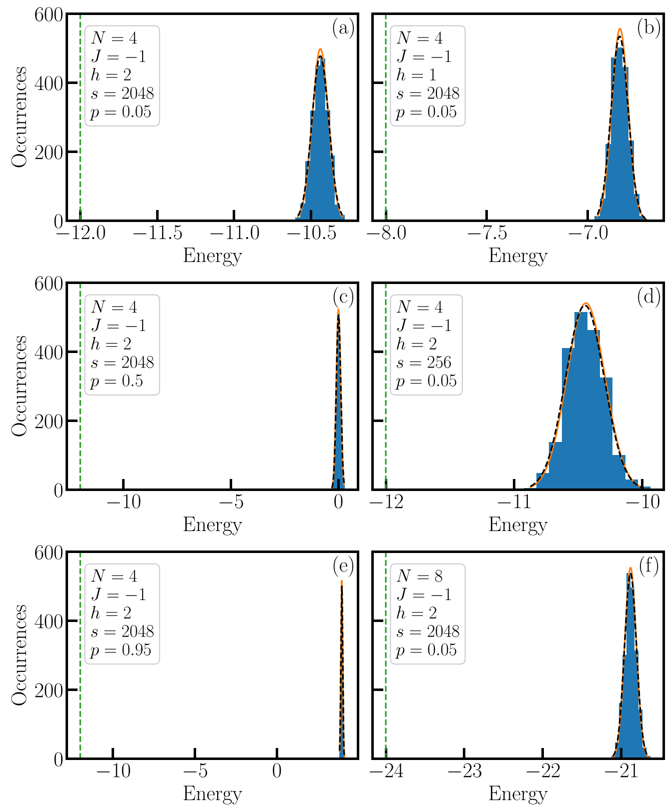


FIG. 1. Energy histograms for the LI model. The vertical dashed green line indicates the true ground-state energy, the solid orange line the prediction from Eqs. (7) and (B17), and the dashed black line a fit to the data. The left column corresponds to $N=4$, $J=-1$, $h=2$, and $s=2048$ with (a) $p=0.05$, (c) $p=0.50$, and (e) $p=0.95$. The right column shows varied N, h, and s: (b) $h=1$, (d) $s=256$, and (f) $N=8$.

III. MITIGATION OF MEASUREMENT ERRORS FOR ARBITRARY OPERATORS

In this section, we generalize our previous results to *arbitrary* operators acting on Q different qubits $q=1,\dots,Q \leqslant N$, where N is the total number of qubits in the system (including the ones the operators are not acting on). We also generalize our previous results to allow for different bit-flip probabilities, $p(|0\rangle \to |1\rangle) \neq p(|1\rangle \to |0\rangle)$, which can also differ among the qubits.

These generalizations are greatly aided by a change in point of view. Whereas previously we treated the bit-flip error as part of the measurement process, i.e., we projectively measured the state $|\psi\rangle$ onto a basis bit string and randomly flipped the bits of this bit string, we now consider the bit flip as part of the operator. In other words, the measurement process no longer includes the bit flips and instead we consider *random* operators to be measured. While this point of view is conceptually very different, we will demonstrate that these random operators yield a distribution of measurements that precisely coincides with the distribution of measurements for a nonrandom operator subject to bit flips.

Our analysis will be split into four parts. First, we will consider a single Z operator acting on a single qubit while allowing for different bit-flip probabilities, $p(|0\rangle \to |1\rangle) \neq p(|1\rangle \to |0\rangle)$, in Sec. III A. In particular, we will compute the operator's expectation as a random operator subject to classical bit flips during measurement. This computation will be the stepping stone to subsequently construct the expectations for noisy measurements of $Z_Q \otimes \cdots \otimes Z_1$ operators with $Q > 1$ in Sec. III B. This construction is inductive with respect to Q and will allow us to construct a classical bit-flip correction procedure for the noisy measurement of $Z_Q \otimes \cdots \otimes Z_1$. It is important to note that the classical bit-flip correction procedure can be preprocessed (replacing the operator to be measured; see Sec. V E) as well as postprocessed (measuring the necessary information first and then extracting the bit-flip corrected expectation values from the measured data).

In Sec. III C, we will consider the special case of equal bit-flip probabilities for all qubits to compare the results directly to Sec. II. In Sec. III D, we will generalize the classical bit-flip correction procedure to arbitrary operators that are measured from bit-string distributions of the state $|\psi\rangle$.

We note that Sec. III D denotes a change in measurement paradigm compared to the previous sections, which affects the variance of the histogram means. We will discuss the different measurement paradigms in detail in Sec. III E and return to the TI model for an explicit illustration. The derivation of the corresponding variances is provided in Appendix B.

A. Measurement of a single Z operator

1. Prediction for the noisy expectation value

For $Q = 1$ and arbitrary N, the noise-free operator Z_q gets replaced by the random noisy operator $\tilde{Z}_q$, which can take the following values:

(i) Z_q with probability $(1 - p_{q,0})(1 - p_{q,1})$.
(ii) $-\mathbb{1}_q$ with probability $p_{q,0}(1 - p_{q,1})$.
(iii) $\mathbb{1}_q$ with probability $(1 - p_{q,0})p_{q,1}$.
(iv) $-Z_q$ with probability $p_{q,0}p_{q,1}$.

Here, $p_{q,b}$ is the probability of flipping the qubit q given that it is in the state $b = |0\rangle$ or $|1\rangle$. For example, $p_{3,0}$ is the probability of flipping $|0\rangle \to |1\rangle$ for qubit 3.

Then, we obtain the noisy expectation $\mathbb{E}\tilde{Z}_q$ for the random operator $\tilde{Z}_q$,

$$\mathbb{E}\tilde{Z}_q = (1 - p_{q,0} - p_{q,1})Z_q + (p_{q,1} - p_{q,0})\mathbb{1}_q, \tag{8}$$

which reduces to Eq. (5) for $p_{q,0} = p_{q,1} =: p$. As before, "expectation" here means the expectation with respect to the bit-flip probabilities, which should not be confused with the quantum-mechanical expectation value $\langle\psi|O|\psi\rangle$ of the operator O. The expectation $\mathbb{E}\tilde{O}$ is the expected value (as an operator) for the noisy operator $\tilde{O}$, while $\mathbb{E}\langle\psi|\tilde{O}|\psi\rangle$ is the expected value for the noisy (quantum-mechanical) expectation value $\langle\psi|\tilde{O}|\psi\rangle$ of the operator $\tilde{O}$.

2. Density matrix description and visualization of measurement noise

For the single-qubit case, it is instructive to express our results in terms of density matrices. Starting from an arbitrary single-qubit density operator,

$$\rho = (\mathbb{1} + \vec{r}\cdot\vec{\sigma})/2, \tag{9}$$

where $\vec{r}$ is a real vector with $\|\vec{r}\| \leqslant 1$, and $\vec{\sigma}$ is the vector containing the Pauli matrices, any quantum channel acting on the state ρ is an affine linear map,

$$\vec{r} \mapsto \vec{r}\,' = M\vec{r} + \vec{c}, \tag{10}$$

where M is a 3×3 real matrix and $\vec{c}$ is a constant real vector [45]. In particular, a noise-free projective measurement in the computational basis corresponds to a unital channel with $M = \mathrm{diag}(0, 0, 1)$ and $\vec{c} = 0$. For an arbitrary pure single-qubit state, $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, with density operator

$$\rho = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \beta\alpha^* & |\beta|^2 \end{pmatrix}, \tag{11}$$

such a projective measurement yields the classical mixture $\rho_c = \mathrm{diag}(|\alpha|^2, |\beta|^2)$.

In the case of a noisy measurement, the bit flips change the classical state that one obtains after the measurement. As discussed above, (i) with probability $(1 - p_0)(1 - p_1)$ we obtain the original state, (ii) with probability $p_0(1 - p_1)$ the $|0\rangle$ flips to a $|1\rangle$, (iii) with probability $(1 - p_0)p_1$ the $|1\rangle$ flips to a $|0\rangle$, and (iv) with probability $p_0 p_1$ both measurement outcomes flip. The resulting classical state can be expressed as a convex linear combination of the different outcomes,

$$\begin{aligned}\tilde{\rho}_c &= \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}(1 - p_0)(1 - p_1) + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} p_0(1 - p_1) + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} p_1(1 - p_0) + \begin{pmatrix} |\beta|^2 & 0 \\ 0 & |\alpha|^2 \end{pmatrix} p_1\, p_0 \\ &= \begin{pmatrix} (1 - p_0 - p_1)|\alpha|^2 + p_1 & 0 \\ 0 & (1 - p_0 - p_1)|\beta|^2 + p_0 \end{pmatrix}. \end{aligned} \tag{12}$$

The expectation value of the Z_q operator then reads

$$\begin{aligned}\langle\tilde{Z}_q\rangle &= \mathrm{Tr}(\tilde{\rho}_c Z_q) \\ &= (1 - p_0 - p_1)(|\alpha|^2 - |\beta|^2) + p_1 - p_0, \end{aligned} \tag{13}$$

which is equivalent to computing the quantum expectation value of Eq. (8), $\mathrm{Tr}(\rho\, \mathbb{E}\tilde{Z}_q)$.

Moreover, we see that Eq. (12) arises from the original density operator ρ by applying the quantum channel

$$\tilde{M} = \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 - p_0 - p_1 \end{pmatrix}, \quad \tilde{\vec{c}} = \begin{pmatrix} 0 \\ 0 \\ p_1 - p_0 \end{pmatrix}. \tag{14}$$

From the equation above, it is apparent that the channel is no longer unital. For $p_0 = p_1$, all quantum states ρ in the equatorial plane of the Bloch sphere, corresponding to $r_z = 0$, are unaffected. The closer the state is to the polar region of the sphere, the more pronounced is the effect of the measurement errors. Compared to the classical state ρ_c obtained from a noise-free projective measurement, the Bloch vector corresponding to $\tilde{\rho}_c$ is shortened because of $\tilde{M}$, and translated along the z axis by $\tilde{\vec{c}}$ (see Fig. 2). Moreover, for $p_0 + p_1 = 1$, the channel maps any state to the same point inside the Bloch sphere. As a result, our mitigation method is not applicable to that special case, which will be further discussed in the next section.

B. Measurement of $Z_Q \otimes \cdots \otimes Z_1$ operators

Going beyond $Q = 1$, we can now compute the noisy expectations for arbitrary operators $Z_Q \otimes \cdots \otimes Z_1$ with $Q > 1$ and arbitrary N. For this, we assume that the expectations of the individual operators can be measured *independently* of

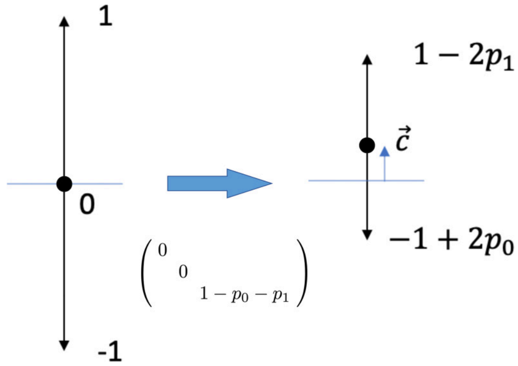


FIG. 2. Left panel: possible range of Bloch vectors of the classical states ρ_c obtained from a noise-free projective measurement in the computational basis. Right panel: deformed range of Bloch vectors corresponding to the classical state $\tilde{\rho}_c$ resulting from a measurement in the presence of measurement noise.

each other. In this case, the noisy expectation of the tensor product $\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1$ equals the tensor product of the individual noisy expectations,

$$\mathbb{E}(\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1) = \mathbb{E}\tilde{Z}_Q \otimes \cdots \otimes \mathbb{E}\tilde{Z}_1. \tag{15}$$

Equation (15) can be proven by considering two different noisy operators $\tilde{O}_1$ and $\tilde{O}_2$ acting on different qubits, and defining their conditional expectations $\mathbb{E}^{\tilde{O}_1}\tilde{O}_1 =: \Omega_1$ and $\mathbb{E}^{\tilde{O}_2}\tilde{O}_2 =: \Omega_2$. The term "conditional" here means that the expectations are only taken with respect to the qubits on which the operators are acting, leaving the other qubits untouched. Now, if we assume $\tilde{O}_1$ takes the values χ_α with probabilities p_α, for example $\tilde{O}_1 = \tilde{Z}_q$ could take $\chi_\alpha \in \{Z_q, -\mathbb{1}_q, \mathbb{1}_q, -Z_q\}$ as above, then we observe

$$\begin{aligned}\mathbb{E}(\tilde{O}_1 \otimes \tilde{O}_2) &= \sum_\alpha p_\alpha\, \mathbb{E}^{\tilde{O}_2}(\chi_\alpha \otimes \tilde{O}_2) \\ &= \sum_\alpha p_\alpha\, \chi_\alpha \otimes \Omega_2 = \Omega_1 \otimes \Omega_2,\end{aligned} \tag{16}$$

which directly yields Eq. (15).

Our final goal is to *reconstruct* the noise-free quantum-mechanical expectation value $\langle\psi|O|\psi\rangle$ of an arbitrary operator $O = O_Q \otimes \cdots \otimes O_1 \in \{\mathbb{1}, Z\}^{\otimes Q}$ from its noisy measurement. To this end, we need to find a matrix ω^{-1} that multiplies the noisy expectations $\mathbb{E}\langle\psi|\tilde{O}|\psi\rangle$ and yields the noise-free expectation values $\langle\psi|O|\psi\rangle$,

$$\langle\psi|O|\psi\rangle = \sum_{\tilde{O}\in\{\mathbb{1},Z\}^{\otimes Q}} \omega^{-1}_{O,\tilde{O}}\mathbb{E}\langle\psi|\tilde{O}|\psi\rangle. \tag{17}$$

For this, we first express the noisy expectation of $\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1$ in Eq. (15) in terms of the noise-free operators $O_Q \otimes \cdots \otimes O_1$. Using Eq. (8), we find

$$\begin{aligned}&\mathbb{E}(\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1) \\ &\quad = \sum_{O\in\{\mathbb{1},Z\}^{\otimes Q}} \gamma(O_Q)O_Q \otimes \cdots \otimes \gamma(O_1)O_1,\end{aligned} \tag{18}$$

where the coefficients γ in front of the noise-free operators are defined as

$$\gamma(O_q) := \begin{cases} 1 - p_{q,0} - p_{q,1} & \text{for } O_q = Z_q, \\ p_{q,1} - p_{q,0} & \text{for } O_q = \mathbb{1}_q. \end{cases} \tag{19}$$

To construct the value of $\mathbb{E}(\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1)$ in Eq. (18) inductively, it is advantageous to choose the "lexicographic order" $\preceq$ for both the noise-free operators $O \in \{\mathbb{1}, Z\}^{\otimes Q}$ and the noisy operators $\tilde{O} \in \{\mathbb{1}, Z\}^{\otimes Q}$,

$$\begin{aligned}&\mathbb{1}_3 \otimes \mathbb{1}_2 \otimes \mathbb{1}_1 \preceq \mathbb{1}_3 \otimes \mathbb{1}_2 \otimes Z_1 \\ &\preceq \mathbb{1}_3 \otimes Z_2 \otimes \mathbb{1}_1 \preceq \mathbb{1}_3 \otimes Z_2 \otimes Z_1 \\ &\preceq Z_3 \otimes \mathbb{1}_2 \otimes \mathbb{1}_1 \preceq Z_3 \otimes \mathbb{1}_2 \otimes Z_1 \\ &\preceq Z_3 \otimes Z_2 \otimes \mathbb{1}_1 \preceq Z_3 \otimes Z_2 \otimes Z_1 \preceq \cdots .\end{aligned} \tag{20}$$

This choice implies $O_Q \otimes \cdots \otimes O_1 \preceq Z_Q \otimes \cdots \otimes Z_1$ and will later ensure that the matrix ω in Eq. (17) is a lower triangular matrix, which is invertible as long as none of its diagonal entries vanish. To determine the matrix ω, we need to generalize Eq. (18) to arbitrary noisy operators,

$$\begin{aligned}&\mathbb{E}(\tilde{O}_Q \otimes \cdots \otimes \tilde{O}_1) \\ &\quad = \sum_{O\in\{\mathbb{1},Z\}^{\otimes Q}} \Gamma(O_Q|\tilde{O}_Q)O_Q \otimes \cdots \otimes \Gamma(O_1|\tilde{O}_1)O_1,\end{aligned} \tag{21}$$

where the coefficients Γ in front of the noise-free operators are now defined as

$$\Gamma(O_q|\tilde{O}_q) = \begin{cases} \gamma(O_q) & \text{for } \tilde{O}_q = \tilde{Z}_q, \\ 1 & \text{for } O_q = \mathbb{1}_q \wedge \tilde{O}_q = \tilde{\mathbb{1}}_q, \\ 0 & \text{for } O_q = Z_q \wedge \tilde{O}_q = \tilde{\mathbb{1}}_q. \end{cases} \tag{22}$$

Using this definition, we can now define the matrix ω as

$$\begin{aligned}\omega(O|\tilde{O}) &:= \prod_{q=1}^{Q} \Gamma(O_q|\tilde{O}_q), \\ \omega &:= [\omega(O|\tilde{O})]_{\tilde{O},O\in\{\mathbb{1},Z\}^{\otimes Q}}.\end{aligned} \tag{23}$$

It is important to note that $\tilde{O} \prec O$ implies $\omega(O|\tilde{O}) = 0$. In other words, ω is a lower triangular matrix and therefore is invertible as long as none of its diagonal entries vanish. The diagonal entries are $\prod_{q=1}^{Q} \Gamma(O_q|\tilde{O}_q)$ and thus can only vanish if one of the $\gamma(Z_q)$ vanishes, i.e., ω is invertible as long as $\forall\, q:\ p_{q,0} + p_{q,1} \neq 1$. If that is the case, then we obtain the bit-flip corrected operators

$$(O)_{O\in\{\mathbb{1},Z\}^{\otimes Q}} = \omega^{-1}(\mathbb{E}\tilde{O})_{\tilde{O}\in\{\mathbb{1},Z\}^{\otimes Q}}. \tag{24}$$

In particular, for $O = Z_2 \otimes Z_1$, we obtain

$$\begin{aligned}Z_2 \otimes Z_1 = &\frac{1}{\gamma(Z_2)\gamma(Z_1)}\mathbb{E}(\tilde{Z}_2 \otimes \tilde{Z}_1) - \frac{\gamma(\mathbb{1}_1)}{\gamma(Z_2)\gamma(Z_1)}\mathbb{E}(\tilde{Z}_2) \otimes \mathbb{1}_1 \\ &- \frac{\gamma(\mathbb{1}_2)}{\gamma(Z_2)\gamma(Z_1)}\mathbb{1}_2 \otimes \mathbb{E}(\tilde{Z}_1) + \frac{\gamma(\mathbb{1}_2)\gamma(\mathbb{1}_1)}{\gamma(Z_2)\gamma(Z_1)}\mathbb{1}_2 \otimes \mathbb{1}_1.\end{aligned} \tag{25}$$

In Fig. 3, we show the relative error for the bit-flip corrected expectation value of $\langle\psi|\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1|\psi\rangle$, as retrieved from histogram data using Eq. (17), compared to the bit-flip

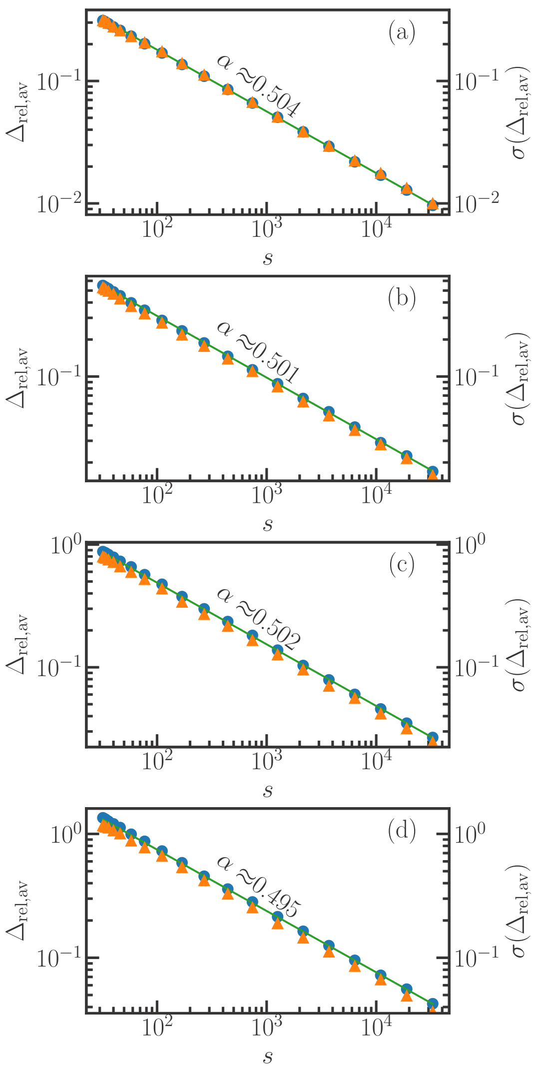


FIG. 3. Mean value $\Delta_{\mathrm{rel,av}}$ (blue dots, left y-axis) and standard deviations $\sigma(\Delta_{\mathrm{rel,av}})$ (orange triangles, right y-axis) of the relative error for the bit-flip corrected expectation values of $\langle\psi|\tilde{Z}_Q\otimes\cdots\otimes\tilde{Z}_1|\psi\rangle$, as retrieved from histogram data using Eq. (17), compared to the "true" bit-flip free expectation values of $\langle\psi|Z_Q\otimes\cdots\otimes Z_1|\psi\rangle$; see Eq. (26). Shown are the four different operators (a) Z_1, (b) $Z_2\otimes Z_1$, (c) $Z_3\otimes Z_2\otimes Z_1$, and (d) $Z_4\otimes Z_3\otimes Z_2\otimes Z_1$. The average relative errors are fitted with a power law in the number of shots s, $y(s)\propto s^{\alpha}$ (green lines); the slopes obtained are indicated in the different panels. The standard deviations of the relative errors are extracted from $2^{12}=4096$ random states $|\psi\rangle$ and random bit-flip probabilities $p_{q,b}$.

free expectation value $\langle\psi|Z_Q\otimes\cdots\otimes Z_1|\psi\rangle$:

$$\Delta_{\mathrm{rel}}=\frac{|\langle\psi|\tilde{Z}_Q\otimes\cdots\otimes\tilde{Z}_1|\psi\rangle-\langle\psi|Z_Q\otimes\cdots\otimes Z_1|\psi\rangle|}{|\langle\psi|Z_Q\otimes\cdots\otimes Z_1|\psi\rangle|}. \tag{26}$$

We also plot the standard deviation of this relative error, alternatively to plotting the error bars. Figure 3 also contains a fit $y(s)=Cs^{-\alpha}$ of the relative error in Eq. (26), where s is again the number of shots, i.e., the number of $\langle\psi|\tilde{Z}_Q\otimes\cdots\otimes\tilde{Z}_1|\psi\rangle$ evaluations needed to produce the histogram. In particular, the fit indicates Monte-Carlo-type convergence $\alpha\approx 1/2$ for $Q\in\{1,2,3,4\}$. Figure 3 has been generated using $2^{12}=4096$ random states $|\psi\rangle$ satisfying $|\langle\psi|Z_Q\otimes\cdots\otimes Z_1|\psi\rangle|\geqslant 0.25$ to avoid dividing by small numbers when computing relative errors. For each $|\psi\rangle$ we randomly chose the bit-flip probabilities $p_{q,b}$ uniformly in (0.05,0.25).

C. Measurement of $Z_Q\otimes\cdots\otimes Z_1$ operators assuming equal bit-flip probabilities

To compare the results of the previous two subsections with the results obtained in Sec. II, we now set all bit-flip probabilities $p_{q,b}=p$ to be equal. For the case $Q=1$, the expectation $\mathbb{E}\tilde{Z}_q$ in Eq. (8) reduces to

$$\mathbb{E}\tilde{Z}_q=(1-2p)Z_q, \tag{27}$$

in agreement with Eq. (5). For $Q>1$, the expectation in Eq. (15) reduces to

$$\mathbb{E}(\tilde{Z}_Q\otimes\cdots\otimes\tilde{Z}_1)=(1-2p)^{Q}\,Z_Q\otimes\cdots\otimes Z_1, \tag{28}$$

which yields Eq. (6) for $Q=2$. This implies that the matrix ω in Eq. (23) becomes diagonal with

$$\mathbb{E}(\tilde{O}_Q\otimes\cdots\otimes\tilde{O}_1)=(1-2p)^{\#Z(O)}O_Q\otimes\cdots\otimes O_1, \tag{29}$$

where $\#Z(O)$ is the number of terms $O_q=Z_q$ in the tensor product $O=O_N\otimes\cdots\otimes O_1$. In particular, ω is invertible as long as $p\neq 1/2$. We again observe in Eqs. (28) and (29) that the noisy expectations of arbitrary operators can be related to the true operators in a surprisingly simple way, which requires no knowledge of the quantum hardware apart from the different bit-flip probabilities of the qubits.

D. Measurement of general operators $\mathcal{H}$ from bit-string distributions of $|\psi\rangle$

1. Prediction for the noisy expectation value

Our analysis of the bit-flip error above assumed that we measure general operators $\mathcal{H}$ by expressing them as linear combinations of operators U^*OU with $O\in\{\mathbb{1},Z\}^{\otimes N}$ on an N-qubit machine, and by measuring each O independently (U being the transformation into the Z basis). For example, if we are interested in measuring $\mathcal{H}_{ZZ}=J\sum_{i=1}^{N}Z_iZ_{i+1}$ with $N=3$ qubits, then we generate independent histograms for $\langle\psi|\mathbb{1}_3\otimes Z_2\otimes Z_1|\psi\rangle$, $\langle\psi|Z_3\otimes Z_2\otimes\mathbb{1}_1|\psi\rangle$, and $\langle\psi|Z_3\otimes\mathbb{1}_2\otimes Z_1|\psi\rangle$, we extract their expectation values, and we recover $\langle\psi|\mathcal{H}_{ZZ}|\psi\rangle$ accordingly. Alternatively, we can measure the *distribution* of $|\psi\rangle$ and obtain a single histogram in terms of the computational basis $\{|j\rangle;\ j\in\mathbb{N}_{0,<2^N}\}$. Hence, if the probability of measuring $|j\rangle$ is $\mathfrak{p}_j$, then we can recover $\langle\psi|\mathcal{H}_{ZZ}|\psi\rangle$ from $\sum_j\mathfrak{p}_j\langle j|\mathcal{H}|j\rangle$. While both approaches yield the same expectation value, the variance obtained for both approaches will in general be different, as we will discuss further below.

Moreover, for a general Hamiltonian $\mathcal{H}$, the full expectation value $\langle\psi|\mathcal{H}|\psi\rangle$ cannot always be recovered from a single histogram via $\sum_j\mathfrak{p}_j\langle j|\mathcal{H}|j\rangle$. For example, if we are interested in measuring the TI Hamiltonian $\mathcal{H}_{\mathrm{TI}}=J\sum_{i=1}^{N}Z_iZ_{i+1}+h\sum_{i=1}^{N}X_i$, we cannot directly recover the full expectation value $\langle\psi|\mathcal{H}_{\mathrm{TI}}|\psi\rangle$ from measuring the distribution of $|\psi\rangle$, because the terms in the Hamiltonian do not all commute.

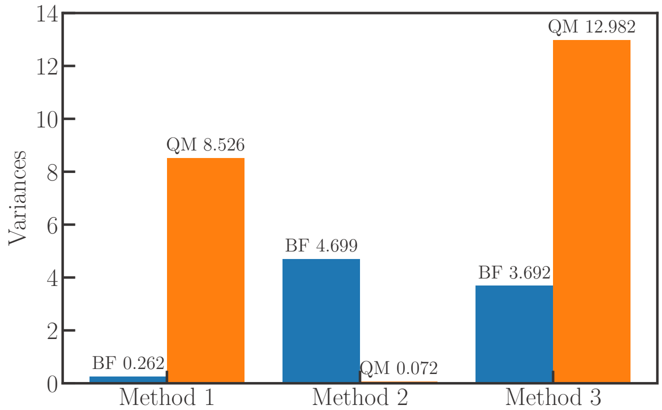


FIG. 4. Contributions to the variance of histogram means for the bit-flip corrected TI Hamiltonian in Eq. (37) evaluated on the ground state of the "true" TI Hamiltonian in Eq. (A12). The different bars correspond to the bit-flip (BF, blue) and quantum-mechanical (QM, orange) variance contributions for the three different measurement methods. We used the parameters $N = 4$, $J = -1$, $h = 2$, and $p_{q,b} = p = 0.05$. All values are normalized by setting $s = 1$.

However, as we discussed below Eq. (2), an efficient implementation on the quantum hardware can be achieved by splitting the Hamiltonian into two sums of Pauli strings $U_k^* O_k U_k \in \{\mathbb{1}, X, Y, Z\}^{\otimes N}$, where multiple summands of the Hamiltonian are measured simultaneously. For example, both $\mathcal{H}_{ZZ} = J\sum_{i=1}^{N} Z_i Z_{i+1}$ and $\mathcal{H}_X = h\sum_{i=1}^{N} X_i$ can be measured using bit-string distributions. Here, $\mathcal{H}_{ZZ}$ can be measured directly by using the bit-string distribution of the state $|\psi\rangle$, and $\mathcal{H}_X$ can be measured by using $h\sum_{i=1}^{N} Z_i$ and the bit-string distribution of the state $H^{\otimes N}|\psi\rangle$, i.e., after applying a Hadamard gate H on each qubit. Hence, using the bit-string distribution, we can measure all the ZZ terms and all the X terms in the TI Hamiltonian simultaneously. In other words, we are only required to measure two bit-string distributions instead of measuring each of the $2N$ Pauli terms separately. This allows for an efficient implementation on the quantum hardware.

If we measure the distribution of $|\psi\rangle$, the measurements of $\langle\psi|U^* O U|\psi\rangle$ comprising $\langle\psi|\mathcal{H}|\psi\rangle$ are no longer independent. This has an impact on the variance of measurement histograms, as we will discuss in Sec. III E. However, it has no impact on the expectation subject to bit flips, since linearity of the expectation value implies

$$\begin{aligned}\mathbb{E}\langle\psi|\tilde{\mathcal{H}}|\psi\rangle &= \mathbb{E}\langle\psi|\sum_\alpha \lambda_\alpha U_\alpha^* \tilde{O}_\alpha U_\alpha|\psi\rangle \\ &= \langle\psi|\sum_\alpha \lambda_\alpha U_\alpha^* (\mathbb{E}\tilde{O}_\alpha) U_\alpha|\psi\rangle, \qquad (30)\end{aligned}$$

which is precisely the expression we would obtain from summing the independently measured operators $\tilde{O}_\alpha$.

2. Prediction for the bit-flip corrected operator

To correct for bit flips in this setting, we need to keep in mind that the general case requires measurements of all operators $O \preceq O_\alpha$ (with respect to the lexicographic order $\preceq$ on $\{\mathbb{1}, Z\}^{\otimes N}$) for all operators O_α in $\mathcal{H} = \sum_\alpha \lambda_\alpha U_\alpha^* O_\alpha U_\alpha$. Hence, the histogram for $\langle\psi|\tilde{\mathcal{H}}|\psi\rangle$ does not contain sufficient information. However, we can use the classical bit-flip correction method as discussed above to find coefficients $\omega_{\alpha,O}$ such that

$$O_\alpha = \sum_{O\preceq O_\alpha} \omega_{\alpha,O} \mathbb{E}\tilde{O} \qquad (31)$$

holds. Inserting this into $\mathcal{H}$, we can express $\mathcal{H}$ as

$$\mathcal{H} = \sum_\alpha \lambda_\alpha U_\alpha^* \sum_{O\preceq O_\alpha} \omega_{\alpha,O} \mathbb{E}\tilde{O} U_\alpha. \qquad (32)$$

In other words, we can replace the operator $\mathcal{H}$ by the bit-flip corrected noisy operator

$$\tilde{\mathcal{H}}_{\rm bfc} := \sum_\alpha \lambda_\alpha U_\alpha^* \sum_{O\preceq O_\alpha} \omega_{\alpha,O} \tilde{O} U_\alpha \qquad (33)$$

and obtain

$$\mathbb{E}\langle\psi|\tilde{\mathcal{H}}_{\rm bfc}|\psi\rangle = \langle\psi|\mathcal{H}|\psi\rangle. \qquad (34)$$

3. Prediction for equal bit-flip probabilities

To compare our results to Secs. II and III C, let us assume that the bit-flip probabilities $p_{q,b}$ satisfy $p_{q,0} = p_{q,1} = p_q$, i.e., there is no difference between $p(|0\rangle \to |1\rangle)$ and $p(|1\rangle \to |0\rangle)$ for each qubit, but this value might depend on the individual qubit. Then we obtain $\omega_{\alpha,O} = 0$ unless $O = O_\alpha = O_{\alpha,N} \otimes \cdots \otimes O_{\alpha,1}$, for which we find

$$\omega_{\alpha,O_\alpha} =: \omega_\alpha = \prod_q \frac{1}{(1-2p_q)}, \qquad (35)$$

where q ranges over all qubits satisfying $O_{\alpha,q} = Z_q$. For $p_{q,b} = p$, this result agrees with Eqs. (5), (6), and (28).

Thus, the bit-flip corrected noisy operator

$$\tilde{\mathcal{H}}_{\rm bfc} := \sum_\alpha \lambda_\alpha \omega_\alpha U_\alpha^* \tilde{O}_\alpha U_\alpha \qquad (36)$$

has the same Pauli-sum structure as the original operator $\mathcal{H}$, changing only the coefficients. This is completely analogous to the independent measurement case. In both cases, if we have $p_{q,0} = p_{q,1}$, then we can correct for bit flips without additional cost to the quantum device.

E. Impact of measurement choices

In general, we will extract the quantum-mechanical expectation of an operator by running the circuit preparing $|\psi\rangle$ followed by a projective measurement in the computational basis a number of times. As before, we refer to these repetitions as the number of shots, s. Of course, these shots are still subject to statistical fluctuations. Hence, if we generate $N_{\rm hist}$ histograms with s shots each, we can generate a histogram from the means extracted from each histogram. This will yield results as in Figs. 1 and 7. Using bit-flip corrected operators as in Eq. (33), we can shift the expected mean to coincide with the quantum-mechanical expectation of the operator we wish to measure. However, the variance of histogram means is then highly dependent on the measurement paradigm.

For illustration, let us consider the TI model $\mathcal{H}_{\rm TI} = J\sum_{j=1}^{N} Z_j Z_{j+1} + h\sum_{j=1}^{N} X_j$, which we will measure on the ground state $|\psi\rangle$. The first step is to compute the bit-flip corrected noisy Hamiltonian $\tilde{\mathcal{H}}_{\rm TI,bfc}$. For simplicity, we will

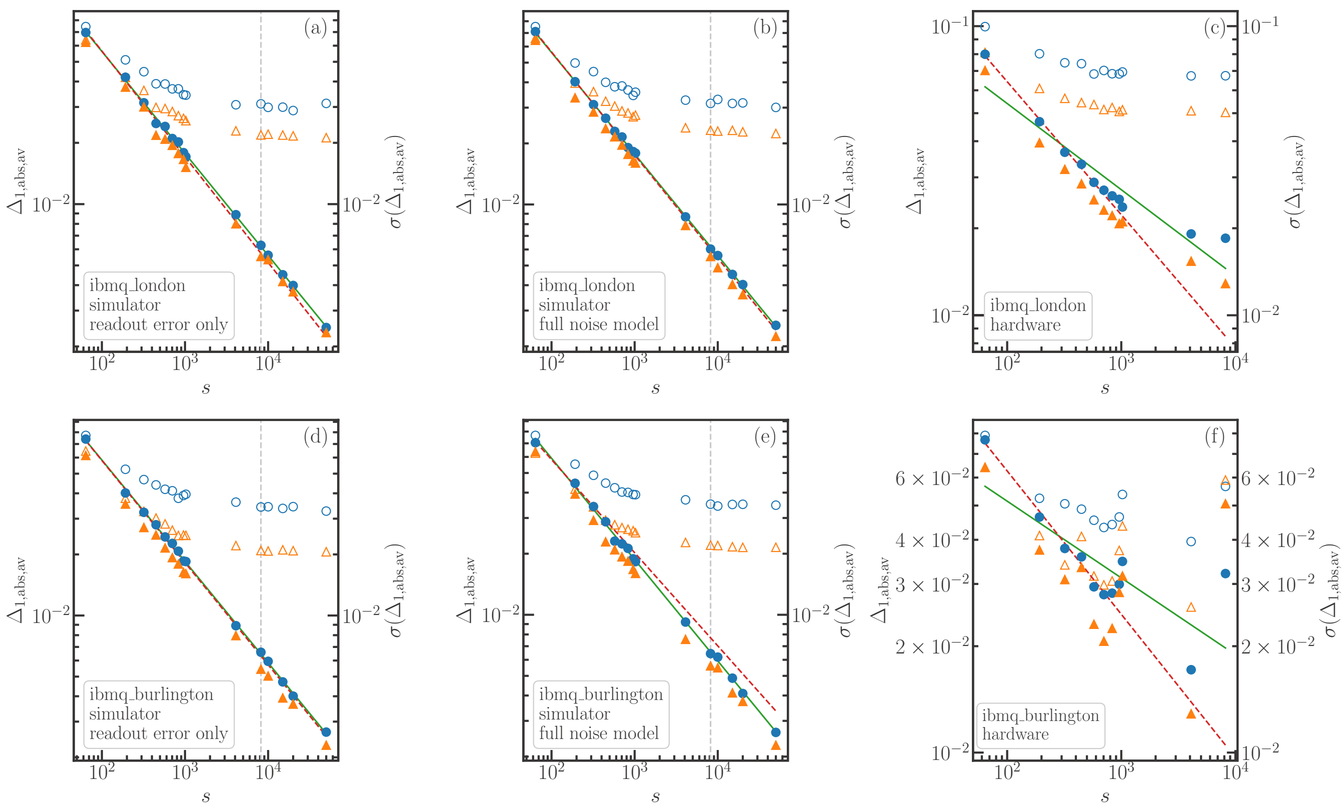


FIG. 5. Mean value $\Delta_{1,\mathrm{abs,av}}$ (blue dots, left y-axes) and standard deviation $\sigma(\Delta_{1,\mathrm{abs,av}})$ (orange triangles, right y-axes) of the absolute error in Eq. (38) after applying the correction procedure (filled symbols) and without it (open symbols) as a function of the number of shots s. The panels in the upper row correspond to data obtained for ibmq_london, the panels in the lower row to data obtained for ibmq_burlington. The different columns correspond to a classical simulation of the quantum device taking into account only readout error (first column), the full hardware noise model (second column), and data obtained on actual quantum hardware (third column). The solid green lines correspond to a power-law fit to all our data points for the mean absolute error, while the red dashed lines correspond to a fit including the lowest four number of shots. The vertical gray dashed lines in panels (a),(b) and (d),(e) indicate the maximum number of shots that can be executed on the actual hardware.

assume that all bit-flip probabilities $p_{q,b}$ coincide with some value p. This yields

$$\tilde{\mathcal{H}}_{\mathrm{TI,bfc}} = J_p \sum_{j=1}^{N} \tilde{Z}_j \tilde{Z}_{j+1} + h_p \sum_{j=1}^{N} \tilde{X}_j \tag{37}$$

with $J_p := J(1-2p)^{-2}$ and $h_p := h(1-2p)^{-1}$. Of course, this process changes the variances. In particular, since Figs. 1 and 7 show histograms without the bit-flip correction, the prediction of variances in Fig. 1 (and Fig. 7 in Appendix A) uses J and h instead of J_p and h_p.

At this point, we need to decide upon the precise way of measuring the Hamiltonian. Essentially, we have a spectrum of possibilities that contains three interesting cases:

(i) *Method 1:* measure each $\tilde{Z}_j\tilde{Z}_{j+1}$ and $\tilde{X}_j$ in Eq. (37) independently.

(ii) *Method 2:* measure the entire Hamiltonian $\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}$ in Eq. (37) from distributions of $|\psi\rangle$ measurements.

(iii) *Method 3:* measure $\tilde{\mathcal{H}}_{ZZ} := J_p \sum_{j=1}^{N} \tilde{Z}_j\tilde{Z}_{j+1}$ and $\tilde{\mathcal{H}}_X := h_p \sum_{j=1}^{N} \tilde{X}_j$ independently from distributions of $|\psi\rangle$ measurements.

Methods 1 and 2 are the two extremes discussed in Secs. III A–III C and Sec. III D, respectively. Note that Method 2 would require us to perform global projective measurements in the eigenbasis of the Hamiltonian, and therefore it is in general not applicable on real hardware devices. Nevertheless, the results allow us to quantify the effect of the bit-flip variance for the idealized setting where the quantum-mechanical contribution to the variance vanishes (up to statistical fluctuations due to a finite number of shots). Method 3 is a reasonable compromise, and it is precisely the method we used for Figs. 1 and 7. Method 3 is also an example that is closely related to implementations of quantum algorithms that are optimized for the number of calls to the quantum device, i.e., implementations in which only parts of an operator can be measured simultaneously, and both Methods 1 and 2 are impractical to various degrees.

The variance of histogram means has two contributions: bit-flip variance and quantum-mechanical variance. These contributions for each of the three methods are shown in Fig. 4. The derivation of these variances can be found in Appendix B; in particular, Fig. 4 shows Eqs. (B19), (B21), and (B24). To remove the dependence on the number of shots per histogram, all variances are multiplied by the number of shots

s, i.e., all values in Fig. 4 correspond to the normalization $s = 1$.

It is interesting to note that not only does the full variance vary in magnitude, but also the relative contribution from bit flips and quantum mechanics is vastly different between the three methods.

If we compare the two extremes—Method 1 and Method 2—we notice that for Method 1 the bit-flip induced variance is small compared to the quantum-mechanical variance, whereas for Method 2 the situation is reversed. Generically, this pattern is to be expected. Method 1 is likely to produce a much smaller bit-flip contribution since all summands are measured independently. Meanwhile, measuring with Method 2 introduces $O(4^N)$ covariance terms, which vanish in Method 1 due to independent measurements of summands. Moreover, concerning Method 2, we note that the quantum-mechanical variance vanishes upon evaluation on an eigenstate of the operator. In Fig. 4, we evaluated the bit-flip corrected TI Hamiltonian $\tilde{\mathcal{H}}_{\mathrm{TI,bfc}} = J_p \sum_{j=1}^{N} \tilde{Z}_j \tilde{Z}_{j+1} + h_p \sum_{j=1}^{N} \tilde{X}_j$ with equal bit-flip probabilities $p_{q,b} = p = 0.05$ on the ground state of the "true" TI Hamiltonian $\mathcal{H}_{\mathrm{TI}} = J \sum_{j=1}^{N} Z_j Z_{j+1} + h \sum_{j=1}^{N} X_j$. For small values of p, we can interpret the bit-flip correction as a small perturbation to the original operator. Hence, the ground state of $\mathcal{H}_{\mathrm{TI}}$ is close to an eigenstate of $\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}$, and thus the quantum-mechanical contribution to the variance is small.

For intermediate methods, such as Method 3, it is generally difficult to predict the different contributions to the variance using similar arguments as above. Depending on the practical limitation of any given implementation, it will be imperative to balance the different contributions to the variance with the number of quantum device calls. For example, for the TI model, fewer quantum device calls per evaluation of the Hamiltonian introduce more covariance terms. In turn, this requires more quantum device calls to obtain the necessary statistical power if we aim to extract a histogram mean with a required level of precision. Thus, this balancing act is highly problem-specific. However, considering Method 3 for the TI model, it clearly shows that great care has to be taken when constructing an intermediate method if the aim is to reduce the overall variance on a given budget of quantum device calls.

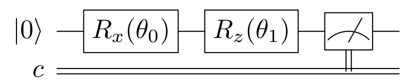
IV. EXPERIMENTAL RESULTS

To demonstrate the experimental applicability of our measurement error-mitigation method, we generate data on IBM quantum hardware using the QISKIT software development kit (SDK) [46]. To assess the performance of our correction procedure, we first simulate the quantum hardware classically using the noise models for the different backends provided by QISKIT before we proceed to the actual hardware.

A. Single-qubit case

To begin with, let us focus on the simplest case of a single qubit. In a first step, we determine the bit-flip probabilities of the qubit. The probability p_0 can be easily obtained by measuring the initial state $|0\rangle$ and recording the number of 1 outcomes, while p_1 requires preparing the state $|1\rangle$ through applying a single X gate to the initial $|0\rangle$ state and recording the number of 0 outcomes. To account for statistical fluctuations, we repeat this procedure several times and average over the bit-flip probabilities obtained for each run (see Appendix C 1 for details).

After obtaining the bit-flip probabilities, we measure $\langle\psi|Z|\psi\rangle$ for a randomly chosen $|\psi\rangle$. Starting from the initial state $|0\rangle$, we can prepare any state on the Bloch sphere by first applying a rotation gate around the x-axis followed by a rotation around the z-axis. Hence, we choose the circuit



in our experiments, where the angles θ_0, θ_1 are both drawn uniformly from the interval $[0, 2\pi]$. Our measurement outcomes allow us to determine the noisy expectation value of Z, $\mathbb{E}(\tilde{Z})$. Subsequently, we can apply our correction procedure using Eq. (8). To acquire statistics for $\mathbb{E}(\tilde{Z})$, we repeat the process for 1050 randomly chosen $|\psi\rangle$ and monitor the mean and the standard deviation of the absolute error,

$$\Delta_{1,\mathrm{abs}} = |\langle\psi|\tilde{Z}|\psi\rangle_{\mathrm{measured}} - \langle\psi|Z|\psi\rangle_{\mathrm{exact}}|, \tag{38}$$

for both the noisy expectation value and the corrected expectation value. Moreover, each individual measurement for fixed values of θ_0 and θ_1 requires running the circuit multiple times to get the probability distribution of basis states in $|\psi\rangle$. Thus we also explore the dependence of our results on the number of shots s.

1. Classical simulation of quantum hardware

To benchmark the performance of our correction procedure, we first simulate ibmq_london [47] and ibmq_burlington [48] classically. The QISKIT SDK provides a noise model for each of the respective chips comprising various sources of error, including readout errors during the measurement process, which can be switched on and off individually. To begin with, we simulate the quantum hardware incorporating the measurement errors only, and subsequently we use the full noise model to see the effect of the various other errors. Our results for the mean and the standard deviation of the absolute error as a function of s are shown in Figs. 5(a), 5(b) and Figs. 5(d), 5(e).

First, let us comment on our results before applying the mitigation procedure. For the case in which the only noise present on the device is the readout error, corresponding to Figs. 5(a) and 5(d), we can derive a relation between the bit-flip probabilities and the saturation values for the mean value of the absolute error in Eq. (38) (open blue dots). As we are interested in the plateau value of the average absolute error in the limit $s \to \infty$, we neglect in the following the statistical uncertainty due to a finite number of shots. Starting from Eq. (8), we can express the absolute error as

$$|\mathbb{E}\langle\tilde{Z}\rangle - \langle Z\rangle| = |-(p_0 + p_1)\langle Z\rangle + p_1 - p_0|. \tag{39}$$

For our choice of ansatz $|\psi(\theta_1, \theta_0)\rangle = R_z(\theta_1)R_x(\theta_0)|0\rangle$, a short analytic calculation yields that $\langle Z\rangle = \cos^2(\theta_0/2) - \sin^2(\theta_0/2)$. Note that this expression is independent of θ_1. Inserting this result into the equation above, we find

$$|\mathbb{E}\langle\tilde{Z}\rangle - \langle Z\rangle| = |-2p_0\cos^2(\theta_0/2) + 2p_1\sin^2(\theta_0/2)|. \tag{40}$$

Averaging this quantity over the angles θ_0 and θ_1, we find the following expression for the average absolute error:

$$\begin{aligned}&\overline{|\mathbb{E}\langle\tilde{Z}\rangle - \langle Z\rangle|}\\&= \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} |-2p_0\cos^2(\theta_0/2)\\&\quad + 2p_1\sin^2(\theta_0/2)|d\theta_0 d\theta_1\\&= \frac{2(p_0+p_1)}{\pi}\left(2\arccos\sqrt{\frac{p_1}{p_0+p_1}} + 2\frac{\sqrt{p_0p_1}}{p_0+p_1} - \frac{\pi}{2}\right)\\&\quad + \frac{4p_1\left(\frac{\pi}{2} - 2\arccos\sqrt{\frac{p_1}{p_0+p_1}}\right)}{\pi}. \qquad (41)\end{aligned}$$

For our numerical simulations using the noise model that keeps the readout error only, the bit-flip probabilities from the QISKIT for the ibmq_london device are $p_0 = 0$ and $p_1 = 0.0317$. Thus, we can insert these values into Eq. (41), which results in a value of $\overline{|\mathbb{E}\langle\tilde{Z}\rangle - \langle Z\rangle|} = p_1$, compatible with the plateau value of the absolute error (open blue dots) in Fig. 5(a). For ibmq_burlington, the noise model has the values $p_0 = 0.015$ and $p_1 = 0.034$, which yields $\overline{|\mathbb{E}\langle\tilde{Z}\rangle - \langle Z\rangle|} \approx 0.0336$, in agreement with our results in Fig. 5(d).

Now, let us comment on our results after applying the mitigation procedure. Focusing on the case with readout error only in Figs. 5(a) and 5(d), we see that correcting our results according to Eq. (8) clearly reduces the mean and the standard deviation of the absolute error in both cases. Without correction, the mean (standard deviation) of the absolute error converges to a value around 0.03 (0.02), and increasing s beyond 1024 does not significantly improve the results. In particular, this stagnation already happens for values of s below the maximum one possible on real hardware, hence showing that the readout error severely limits the accuracy that can be achieved. On the contrary, the corrected results show a significant improvement and a power-law decay of these quantities with s. In particular, in the ideal, completely noise-free case, performing a projective measurement on $|\psi\rangle$ is nothing but sampling from a probability distribution, thus one would expect the mean error to decay as $\propto s^{-1/2}$. To check for that behavior, we can fit the same functional form as in Sec. III B to our data; the resulting exponents are shown in Table I. Indeed, we recover $\alpha = 1/2$, thus demonstrating that our correction procedure essentially allows us to recover the noise-free case.

TABLE I. Exponents α obtained from fitting the power law $Cs^{-\alpha}$ to our simulator data for the mean absolute error in Figs. 5(a), 5(b), and 5(d), 5(e) after applying the correction.

Readout error only	First four points	Full range
ibmq_london	0.519	0.501
ibmq_burlington	0.503	0.499
Full noise model	**First four points**	**Full range**
ibmq_london	0.508	0.500
ibmq_burlington	0.459	0.503

TABLE II. Exponents α obtained from fitting the function $Cs^{-\alpha}$ to our hardware data for the mean absolute error in Figs. 5(c) and 5(f) after applying the correction.

Chip name	First four points	Full range
ibmq_london	0.460	0.298
ibmq_burlington	0.405	0.217

Taking into account the full noise model in our simulations, which contains, for instance, gate errors and decoherence, we obtain the results in Figs. 5(b) and 5(e). Compared to the case with readout errors only, the picture is very similar, which shows that the dominant error contribution for the single-qubit case is coming from the readout procedure. The mean and the standard deviation of the absolute error without any correction only approach marginally higher values than previously. Again, we observe a significant reduction of the mean and the standard deviation of the absolute error after applying the correction procedure, and a power-law decay with s. Fitting a power law to our data yields once more exponents around $1/2$ (see Table I).

2. Quantum hardware

Our experiments can be readily carried out on quantum hardware, and we repeat the same simulations on ibmq_london and ibmq_burlington. The only difference with respect to the classical simulation is that s on those two devices is limited to a maximum number of 8192. Figures 5(c) and 5(d) show our results obtained on real devices.

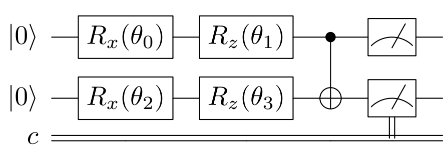
Comparing our data for the chip imbq_london in Fig. 5(c) to the classical simulation of the quantum hardware in Figs. 5(a) and 5(b), we observe qualitative agreement for $s \leqslant 1024$. Compared to the classical simulation of the quantum device, the mean value and the standard deviation of the absolute error are in general larger on the hardware. Correcting for the readout error yields again a significant improvement and reduces the mean and the standard deviation of absolute error considerably. As before, we can fit our data to a power law. While for a small number of shots about $s < 500$ we observe again an exponent of about $1/2$, for a larger number of measurements the curve for the corrected result starts to flatten out, and the exponent obtained for fitting the entire range is considerably smaller than $1/2$ (see Table II for details). Since increasing s should decrease the inherent statistical fluctuations of the projective measurements, and readout errors can be managed with our scheme, this might be an indication that in addition to readout errors, also other sources of noise play a significant role. Their effects cannot be corrected with our procedure and thus dominate from a certain point on.

Looking at the results for imbq_burlington in Fig. 5(f) and comparing them with the classical simulation of the quantum hardware in Figs. 5(d) and 5(e), we see that the discrepancies in this case are more severe and the data are less consistent. Applying our mitigation to the data again yields an improvement, which is less pronounced than in the case of imbq_london. For a small number of shots, the mean of the absolute error after correction shows again roughly a power-law decay. The exponent obtained from a fit to our data in that range is smaller compared to the one from our data from ibmq_london (see Table II for details). From $s = 1024$ on, the uncorrected data are already less consistent. Making use of our mitigation scheme still yields an improvement, however the corrected results scatter similarly to the original ones and do not follow the same power law as for a small number of shots, as a fit to our data reveals. This suggests that noise other than the one resulting from the measurement has a considerable contribution.

B. Two-qubit case

Since our bit-flip correction procedure is not limited to the single-qubit setup, we can straightforwardly apply it to multiple qubits. To assess the performance for that case, we repeat the same procedure that we did previously but now for a circuit encompassing two qubits. Since we assume the bit-flip probabilities $p_{q,b}$ (with $q = 1, 2$, $b = 0, 1$) of the qubits to be independent of each other, we apply the same procedure that we used to obtain the bit-flip probabilities in the single-qubit case, but this time for each qubit individually.

Subsequently, we prepare a two-qubit state using the following circuit:



where the angles $\theta_0, \dots, \theta_3$ are again random numbers drawn uniformly from $[0, 2\pi]$, and the final CNOT gate allows for creating entanglement between the two qubits. Analogous to the single-qubit case, we first simulate the quantum hardware classically before we eventually carry out our experiments on a real quantum device. In both cases, we measure the noisy expectation value of $Z_2 \otimes Z_1$, $\mathbb{E}(\tilde{Z}_2 \otimes \tilde{Z}_1)$, and apply Eq. (17) to correct for noise caused by readout errors. Again, we repeat the procedure for 1050 randomly chosen sets of angles and compute the mean and the standard deviation of the absolute error,

$$\Delta_{2,\mathrm{abs}} = |\langle\psi|\tilde{Z}_2 \otimes \tilde{Z}_1|\psi\rangle_{\mathrm{measured}} - \langle\psi|Z_2 \otimes Z_1|\psi\rangle_{\mathrm{exact}}|, \tag{42}$$

as a function of the number of shots, s, with and without applying the mitigation scheme.

1. *Classical simulation of quantum hardware*

As for the single-qubit case, we use the QISKIT SDK to classically simulate the chips imq_london and ibmq_burlington first with readout error only and subsequently using the full noise model. Figures 6(a), 6(b) and 6(d), 6(e) show our results for both cases.

Looking at Figs. 6(a) and 6(d), we see that the two-qubit case with just readout error behaves like the single-qubit case. Without applying any correction, the mean and the standard deviation of the absolute error initially decrease with increasing s, before eventually converging to fixed values that are slightly higher than for the single-qubit case [compare Fig. 5(a) with Fig. 6(a) and Fig. 5(b) with Fig. 6(b)]. Applying the correction procedure, we can significantly decrease the values and observe again a power-law decay with an exponent of $1/2$ over the entire range of s that we study, as a fit to our corrected data reveals (see also Table III).

Repeating the same simulations, but this time with the full noise model, yields the results in Figs. 6(b) and 6(e). Comparing this to the case with readout error only, we see a more pronounced effect than in the single-qubit case. Applying the correction still reduces the mean and the standard deviation of the absolute error considerably, nevertheless one can observe that data after correction converge to a fixed value with increasing s. In particular, the power-law decay with $\alpha = 1/2$ is only present for a small number of shots. Considering the entire range of s that we study, the classical simulation of ibmq_london predicts that the data are not very compatible with a power law. In contrast, our simulation data for ibmq_burlington are still reasonably well described by a power law, however with an exponent of 0.38 and thus considerably smaller than $1/2$ (see Table III for details). Most notably, from a comparison between the results for classically simulating two qubits using the full noise model to the single-qubit case in Figs. 5(b) and 5(e), we see that noise has a substantially larger effect in the two-qubit case. This can be partially explained by the CNOT gate in the circuit, as the error rates for two-qubit gates are in general much larger than for single-qubit rotations.

2. *Quantum hardware*

For the two-qubit case, we can carry out the simulations on real quantum hardware as well. Using again imbq_london and ibmq_burlington, we obtain the data depicted in Figs. 6(c) and 6(f).

Our results for ibmq_london in Fig. 6(c) show qualitative agreement with the classical simulation. Once more, we see that the mean and the standard deviation of the absolute error obtained on the hardware converge to higher values than the ones obtained from the simulation [compare Figs. 6(b) and 6(c)]. Correcting our data according to Eq. (17), the mean of the absolute error and its standard deviation are significantly reduced. Comparing the reduction to the single-qubit case in Fig. 5, we observe that for the two-qubit case, the improvement is even larger. In particular, for our largest number of shots $s = 8192$, the mean and the standard deviation of the absolute error are reduced by approximately one order of magnitude. The corrected data are again well described by a power law. Fitting the first four data points, we obtain an exponent of 0.48. Using the entire range of s for the fit, the exponent only decreases moderately to 0.39 (see also Table IV), thus showing that the readout error still has a significant contribution to the overall error.

Turning to our results for ibmq_burlington in Fig. 6(f), we see that the data for this chip are significantly worse. For one, the mean value (standard deviation) of the absolute error without applying any correction procedure is roughly a factor 3 (2) larger than the one obtained on ibmq_london. Applying the correction procedure still yields an improvement, however this time it is a lot smaller than for ibmq_london, as a comparison between Figs. 6(c) and 6(f) shows. While for a small number of shots the mean value of the absolute error after correction

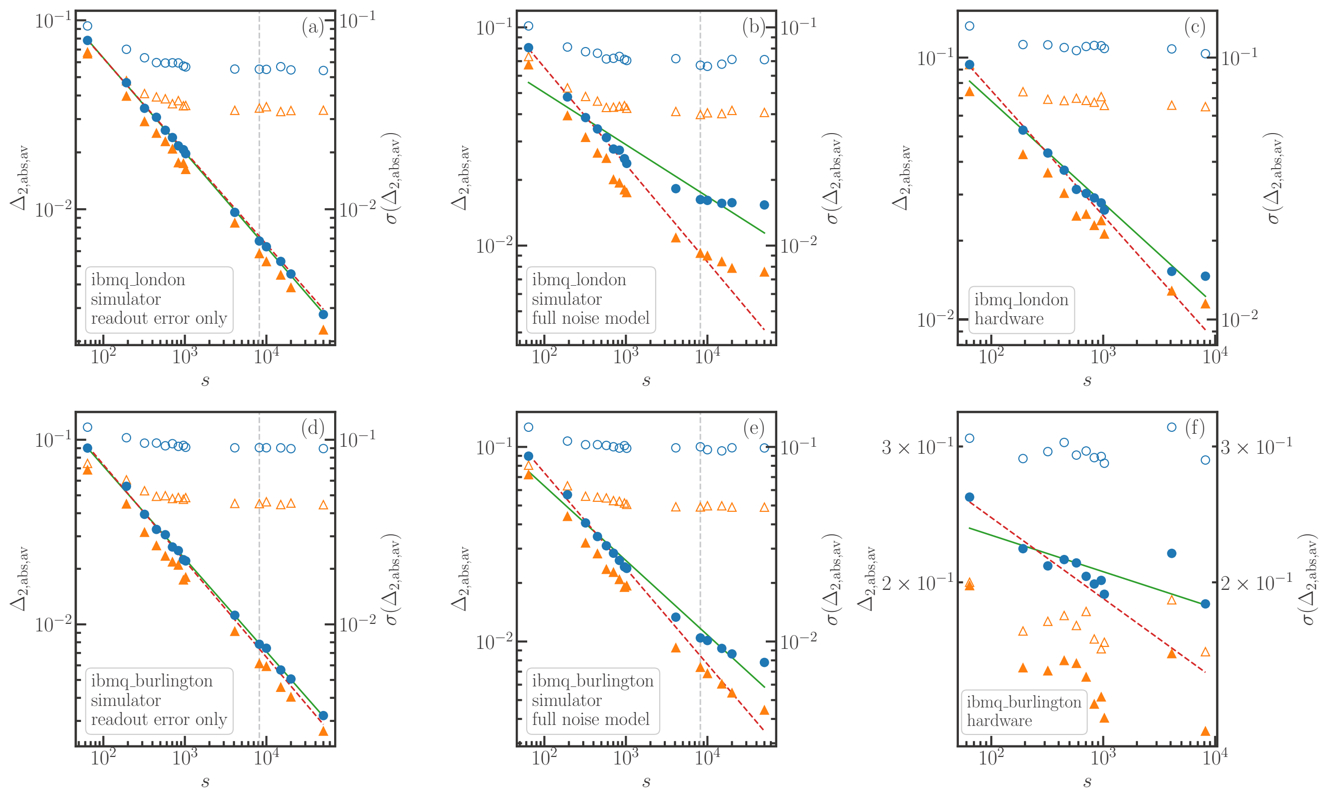


FIG. 6. Mean value $\Delta_{2,\mathrm{abs,av}}$ (blue dots, left y-axes) and standard deviation $\sigma(\Delta_{2,\mathrm{abs,av}})$ (orange triangles, right y-axes) of the absolute error in Eq. (42) after applying the correction procedure (filled symbols) and without it (open symbols) as a function of the number of shots s. The panels in the upper row correspond to data obtained for ibmq_london, the panels in the lower row to data obtained for ibmq_burlington. The different columns correspond to a classical simulation of the quantum device taking into account only readout error (first column), the full hardware noise model (second column), and data obtained on actual quantum hardware (third column). The solid green lines correspond to a power-law fit to all our data points for the mean absolute error, while the red dashed lines correspond to a fit including the lowest four number of shots. The vertical gray dashed lines in panels (a),(b) and (d),(e) indicate the maximum number of shots that can be executed on the actual hardware.

still shows a power-law decay, albeit with an exponent a lot smaller than $1/2$, for a large number of shots this trend stops, as fits to our data reveal (see also Table IV). This behavior indicates that for ibmq_burlington, the readout error is not the dominant one, but rather other errors have a significant contribution that cannot be corrected for using our scheme.

V. DISCUSSION

After demonstrating the applicability of our mitigation method to real quantum hardware, we discuss our results here in greater detail. We comment on the relation to previous works on error mitigation, and we address how our scheme allows for extensions beyond those. In particular, we discuss the inclusion of multiqubit correlations and the generalization to other types of errors such as relaxation. Moreover, we address some questions regarding the practical implementation, such as the overhead costs introduced, preprocessing versus postprocessing, and the possibility of doing probabilistic error mitigation.

TABLE III. Exponents α obtained from fitting the power law $Cs^{-\alpha}$ to our simulator data for the mean absolute error in Figs. 6(a), 6(b) and 6(d), 6(e).

Readout error only	First four points	Full range
ibmq_london	0.492	0.501
ibmq_burlington	0.522	0.503
Full noise model	**First four points**	**Full range**
ibmq_london	0.446	0.238
ibmq_burlington	0.492	0.383

A. Comparison to previous work

One way to mitigate measurement errors that has been put forward in the literature (see, e.g., Ref. [46]) is to construct

TABLE IV. Exponents α obtained from fitting the power law $Cs^{-\alpha}$ to our hardware data for the mean absolute error in Figs. 6(c) and 6(f).

Chip name	First four points	Full range
ibmq_london	0.478	0.390
ibmq_burlington	0.105	0.047

a linear map, which relates the observed measurement outcomes for each computational basis state to the state that was actually prepared. To this end, one prepares all computational basis states $|i\rangle$, $i = 0, \dots, 2^N - 1$, on the quantum device and records the probabilities p_{ji} of obtaining the computational basis state $|j\rangle$ after a projective measurement. The linear map $\omega = (p_{ji})_{i,j=0}^{2^N-1}$ now relates the observed probability distribution of basis states $\tilde{P}$ in a noisy measurement to the ideal distribution P as $\tilde{P} = \omega P$. Thus, one can in principle obtain the exact solution P from the observed results by inverting ω and postprocessing $\tilde{P}$. Obviously, the method scales exponentially with N in terms of the number of measurements and memory requirements. In addition, ω can be singular, and a direct inversion might not be possible. Even if ω^{-1} exists, it is not guaranteed to be stochastic, such that the result obtained might not be a valid probability distribution.

To overcome these shortcomings, it has been proposed to mitigate measurement errors by expressing the error-corrected result in terms of a sum of noisy outcomes and combinations of bit-flip probabilities [12,39,40]. This was first studied by Kandala *et al.* [12] for the case of single-qubit Z-operator measurements. An extension of this method has been provided by Yeter-Aydeniz *et al.* [39,40] for multiqubit $Z_Q \otimes \cdots \otimes Z_1$ operators with expectation values measured from bit-string distributions. Our approach provides an alternative proof for some of the results in Refs. [39,40], which offers an implementation beyond bit-string distributions and allows for several further generalizations, which are discussed in the following subsections. In particular, our results can be extended to multiqubit correlation errors (see Sec. V B), relaxation errors (see Sec. V C), and probabilistic mitigation schemes (see Sec. V D). Moreover, while previous results all rely on postprocessing of the error-mitigation scheme, our method allows for preprocessing as well. Thus, it can be readily integrated into hybrid quantum-classical algorithms, such as the variational quantum eigensolver (VQE) (see Sec. V E).

B. Inclusion of multiqubit correlations

In our paper, we have assumed for simplicity that there are no multiqubit correlations in multiqubit $Z_Q \otimes \cdots \otimes Z_1$-operator measurements. This is because most of the physically relevant Hamiltonians only contain local interaction terms (see the discussion in Sec. V F). As such, the number of qubits for each multiqubit $Z_Q \otimes \cdots \otimes Z_1$-operator measurement is relatively small and the correlations are negligible.

However, as more qubits are measured simultaneously, multiqubit correlations can become significant and have to be taken into account. This has not been incorporated into the above-mentioned mitigation schemes [12,39,40], and, to our knowledge, it has only been addressed with methods that are exponentially costly with the number of qubits [49,50].

Our measurement mitigation scheme can easily take multiqubit correlations into account, because the fundamental step of our approach is the replacement of the operator to be measured with a probability distribution of operators. Adding multiqubit correlations into this probability distribution is straightforward and only requires multiqubit calibration results similar to the single-qubit calibrations discussed in Appendix C. For example, while the single-qubit calibrations required measuring $p(|j\rangle||k\rangle)$ for $j, k \in \{0, 1\}$, the two-qubit calibrations would require measuring $p(|j\rangle||k\rangle)$ for $j, k \in \{00, 01, 10, 11\}$. Since we are interested in n-local Hamiltonians with at most n-qubit interactions (see also the discussion in Sec. V F), the calibration cost scales polynomially in the number of qubits N, as no more than n qubits are measured simultaneously. Indeed, the multiqubit calibration method requires the calibration of $\binom{N}{n}$ n-qubit systems with fixed n, which requires $O(N^n)$ calibrations. Thus, incorporating multiqubit correlations into our mitigation scheme is straightforward and requires relatively small overhead costs compared to previous approaches.

C. Extension to relaxation errors

Our method of replacing noisy operators with random operators that model the noise behavior, as presented in Sec. III, can in principle be generalized to other types of errors on noisy quantum computers. For example, if we wish to measure the operator Z and consider the relaxation error T_1 (decay of $|1\rangle$ to $|0\rangle$) [46], then we have a probability distribution of measuring Z (not yet decayed) and $\mathbb{1}$ (decayed). Hence, the measurement outcome subject to the T_1 error is described by $\tilde{Z} = p(t)Z + [1 - p(t)]\mathbb{1}$, where $p(t) = \exp(-\frac{t}{T_1})$ is the probability that $|1\rangle$ has not yet decayed. With our scheme of replacing noisy with random operators, this T_1 error can be corrected as $Z = \frac{1}{p(t)}\tilde{Z} - \frac{1-p(t)}{p(t)}\mathbb{1}$. As such, our approach is generalizable to other types of errors beyond measurement errors, and it needs to be adapted accordingly in order to correctly incorporate the parameters underlying the specific type of error.

D. Probabilistic implementation of the measurement error-mitigation scheme

The probabilistic description of the noisy operator naturally lends itself to a probabilistic implementation of the mitigation scheme. While deterministic mitigation schemes require the measurement of all mitigation terms, a probabilistic protocol allows for partial error mitigation if the full mitigation is too costly. For example, if the corrected operator $Z \otimes Z$ is given by $\alpha_1\tilde{Z}\otimes\tilde{Z} + \alpha_2\tilde{Z}\otimes\mathbb{1} + \alpha_3\mathbb{1}\otimes\tilde{Z} + \alpha_4\mathbb{1}\otimes\mathbb{1} = A(p_1s_1\tilde{Z}\otimes\tilde{Z} + p_2s_2\tilde{Z}\otimes\mathbb{1} + p_3s_3\mathbb{1}\otimes\tilde{Z} + p_4s_4\mathbb{1}\otimes\mathbb{1})$ with $A := \sum_j |\alpha_j|$, $s_j := \mathrm{sgn}(\alpha_j)$, $p_j := \frac{|\alpha_j|}{A}$, and $\sum_j p_j = 1$, then we may randomly draw the operator to measure; namely, $As_1\tilde{Z}\otimes\tilde{Z}$ with probability p_1, $As_2\tilde{Z}\otimes\mathbb{1}$ with probability p_2, etc. This can be used on a single-operator measurement, commuting sets of operators, as well as on the level of drawing random Hamiltonians. It can also be generalized to other types of errors, such as the inclusion of multiqubit correlations.

E. Preprocessing the mitigation scheme

All of the previously known measurement mitigation schemes rely on postprocessing, that is, first measuring without taking the error into account and afterwards manipulating the obtained data (see, e.g., Refs. [12,39,40,46,49,50]). However, this is not always possible using "black box subroutines," such as VQE routines provided by SDKs. Such routines typically ask for the Hamiltonian to be passed as an argument, and they will return the optimized parameter set. They do not allow for user-supplied error-mitigation methods to be incorporated. In contrast, the approach presented in this work allows for preprocessing. Hence, rather than passing on the Hamiltonian we are interested in solving, we can pass on the bit-flip corrected Hamiltonian instead. The user is therefore able to manually insert an error-mitigation scheme into the "black-box subroutine."

F. Moderate overhead costs

For local Hamiltonians, the computational cost of our mitigation routine scales polynomially with the number of qubits. For nonlocal Hamiltonians, our mitigation routine does not add any computational cost with respect to the measurement itself, as the measurement of the expectation value already exhibits exponential complexity. We will explain and exemplify both of these cases in the following.

For nonlocal Hamiltonians, let us consider the example of a generic operator acting on N qubits. This operators is a linear combination of all 4^N N-qubit Pauli matrices, i.e., a tensor product of N 2×2 Pauli matrices, which include the 2×2 identity. Our mitigation method now replaces each tensor product by a sum of up to 2^N operators, which already need to be measured for the full Hamiltonian measurement. Thus, the replacement only changes the coefficients of these operators but does not incur any overhead on the quantum device. Moreover, Hamiltonians with nonlocal interactions are likely to incur exponential complexity already in the evaluation of the expectation value, thus making them unfeasible to measure, let alone error-correct.

For n-local Hamiltonians, the individual Pauli terms do not act on all N qubits but on a given number $Q \leqslant n \leqslant N$ of qubits, which is independent of the total number N of qubits. For example, the Ising model, the Heisenberg model, and the Schwinger model (after integrating out the gauge field; see, e.g., Refs. [51–54]) exhibit at most two-qubit interaction terms and thus have $Q \leqslant n = 2$. For our mitigation method, we now need to replace each tensor product of the Q nonidentity Pauli matrices by up to 2^Q operators. For each $Q \leqslant n$, there are polynomially many of these terms, $\binom{N}{Q}$, and we can estimate the total number using the upper bound $\sum_{Q \leqslant n} \binom{N}{Q} \leqslant (n+1)N^n$. For each of these terms, the error-correction matrix ω is of magnitude 2^Q, i.e., bounded by a constant of magnitude 2^n. Each of the matrices ω are triangular and thus can be inverted with a computational cost of $O(4^n)$, which is still constant. The entire operation requires fewer than $(n+1)N^n$ times $O(1)$ operations, i.e., the computational complexity is $O(N^n)$ and thus scales polynomially in the number of qubits N. Note that for n-local interactions between *adjacent* qubits, the computational complexity gets reduced even further, because we then only have nN instead of $\sum_{Q=1}^{n} \binom{N}{Q}$ different terms.

For the Ising, Heisenberg, and Schwinger models mentioned above, we can explicitly estimate the number of terms required for our error-mitigation method. For each qubit q, we have at most three single-qubit Paulis (X_q, Y_q, Z_q), as well as three two-qubit Paulis (X_qX_{q+1}, Y_qY_{q+1}, Z_qZ_{q+1}). Hence, for N qubits, our mitigation method requires the inversion of $3N$ matrices ω (size 2×2) for the single-qubit Paulis and $3N$ matrices ω (size 4×4) for the two-qubit Paulis. In other words, the overall complexity of the error mitigation for N qubits is bounded by $3N$ triangular matrix inversions of size 2×2 plus $3N$ triangular matrix inversions of size 4×4.

In some cases, the number of additional terms that have to be measured on the quantum device can be reduced even further. For example, the LI model discussed in Sec. II A only contains the Paulis Z_qZ_{q+1} and Z_q for each qubit q. While error-correcting the Z_qZ_{q+1} terms, we automatically error-correct the Z_q terms as well. Hence, the overall complexity for the error mitigation of the LI model for N qubits is equivalent to N triangular matrix inversions of size 4×4. Thus, the LI model incurs no overhead cost on the quantum device.

VI. CONCLUSIONS

In this paper, we proposed a classical bit-flip correction method to mitigate measurement errors on noisy quantum computers. This method relies on cancellations of different erroneous measurement outcomes and requires knowledge of the different bit-flip probabilities during readout for each qubit. We tested the performance of this method by correcting the noisy energy histograms of the longitudinal and transversal Ising models. Moreover, we demonstrated that the method can be applied to any operator, any number of qubits, and any realistic bit-flip probability. For the single-qubit case, we also provided a density matrix description and a visualization scheme of the measurement noise. Finally, we tested our method both numerically and experimentally for the IBM quantum devices ibmq_london and ibmq_burlington for both a single qubit and two qubits. We observe that our method is able to improve the data significantly for both cases and to reduce the error by up to one order of magnitude.

Our method of replacing noisy operators with random operators that model the noise behavior, as we presented in Sec. III, is generally applicable to arbitrary observables and could also be applied to other error sources, such as relaxation errors. As stated in Sec. V, the computational cost of our mitigation routine is moderate (i.e., polynomial) for local Hamiltonians, even if multiqubit correlations are included. For nonlocal Hamiltonians, our mitigation routine does not add any computational cost with respect to the measurement itself, as the measurement of the expectation value already exhibits exponential complexity.

In addition to the moderate overhead cost, another advantage of our mitigation scheme is that it can be readily integrated into hybrid quantum-classical algorithms, as, for example, the quantum approximate optimization algorithm [55] and VQS. After initially measuring the bit-flip probabilities, one can simply correct the values obtained for the cost function from the quantum device, before passing them on to a classical algorithm for optimizing the variational parameters. Moreover, in contrast to previous mitigation schemes, our method also allows for preprocessing. Thus, the user can manually insert the bit-flip corrected Hamiltonian into "black-box subroutines" such as VQS routines provided by SDKs, which allows for on-the-fly error mitigation.

Finally, our method is completely platform-independent and lends itself not only to superconducting qubits, but also to other architectures such as trapped ions. As long as the measurement errors are constant to a certain degree and not excessively large, they can be reliably corrected for with our procedure. These advantages make our mitigation method promising for various applications on NISQ devices but also beyond.

Note added. Recently, we became aware of a related mitigation method [56], which also assumes uncorrelated measurement errors and thus replaces the exponential scaling of common methods by a polynomial one. The implementation in the current paper goes further by constructing operators that correct for the measurement error on the operator level. While the current paper demonstrates the performance of the method with two-qubit experiments, Ref. [56] provides a demonstration on up to 42 qubits.

ACKNOWLEDGMENTS

The authors thank Xu Feng, Georgios Polykratis, Philipp Stratmann, and Tom Weber for valuable discussions and comments. We also thank Tom Weber for pointing out the probabilistic implementation of the mitigation scheme. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Industry Canada and by the Province of Ontario through the Ministry of Colleges and Universities. S.K. acknowledges financial support from the Cyprus Research and Innovation Foundation under project "Future-proofing Scientific Applications for the Supercomputers of Tomorrow (FAST)," Contract No. COMPLEMENTARY/0916/0048. P.S. thanks the Helmholtz Einstein International Berlin Research School in Data Science (HEIBRiDS) for funding. We acknowledge the use of IBM Quantum services for this work. The views expressed are those of the authors, and do not reflect the official policy or position of IBM or the IBM Quantum team.

APPENDIX A: ILLUSTRATION OF THE MITIGATION METHOD FOR SIMPLIFIED ENERGY HISTOGRAMS

In this Appendix, we provide a pedagogic explanation of our mitigation method. We describe in Appendix A 1 how the measured energy histograms can be described by a binomial distribution in certain cases. We then discover in Appendix A 2 that the mean energy of the distribution vanishes if all qubits have equal bit-flip probabilities, $p = 0.5$. We finally discuss the deviation $\Delta\tilde{E}$ between the measured mean energy and the noiseless true energy for the specific cases of noninteracting Hamiltonians (Appendix A 3), interacting Hamiltonians (Appendix A 4), and the example of the TI model (Appendix A 5).

1. Binomial distribution of measurements

To start with, let us focus on a diagonal Hamiltonian $\mathcal{H}$ with eigenstate $|\psi\rangle$. To evaluate the corresponding energy, $E = \langle\psi|\mathcal{H}|\psi\rangle$, on a quantum device, we have to (i) run the quantum circuit preparing the state $|\psi\rangle$, (ii) projectively measure the energy in the (computational) basis, and (iii) record the outcomes. Performing s shots, we record k correct results with $\tilde{E} = E$ but $s - k$ incorrect results with $\tilde{E} \neq E$, where the tilde denotes a noisy outcome. For simplicity, we assume that (i) a wrong measurement originates from a *single* bit flip with probability p, and (ii) each bit flip yields the *same* deviation from E. We will see later that these assumptions will need to be modified in the presence of multiqubit interactions, for example for the LI and TI models.

The probability of getting k correct measurement results is given by the probability mass function

$$f(k, s, 1-p) = \binom{s}{k} (1-p)^k \, p^{s-k}, \tag{A1}$$

where p is the probability of *incorrectly* measuring the energy. The resulting noisy energy histograms can be described in terms of the number k of correct measurements,

$$\begin{aligned} \tilde{E}(k) &= E + (s-k)\Delta\tilde{E} \\ &= \begin{cases} E & \text{for } k = s, \\ E + s\Delta\tilde{E} & \text{for } k = 0, \end{cases} \end{aligned} \tag{A2}$$

where $\Delta\tilde{E}$ is the deviation from E per bit flip. In terms of the bit-flip probability p, the resulting noisy expectation $\mathbb{E}$ of the measured energy $\tilde{E}$ reads

$$\begin{aligned} \mathbb{E}\tilde{E} &= E + sp\Delta\tilde{E} \\ &= \begin{cases} E & \text{for } p = 0, \\ E + s\Delta\tilde{E} & \text{for } p = 1. \end{cases} \end{aligned} \tag{A3}$$

We note that "expectation" here means the expectation with respect to the probability p, which should not be confused with the quantum-mechanical expectation value of the Hamiltonian, $\langle\psi|\mathcal{H}|\psi\rangle = E$. Thus, the expectation $\mathbb{E}\tilde{\mathcal{H}}$ is the expected value (as an operator to be measured subject to bit flips; see also Sec. III) for the noisy Hamiltonian $\tilde{\mathcal{H}}$, while $\mathbb{E}\langle\psi|\tilde{\mathcal{H}}|\psi\rangle = \mathbb{E}\tilde{E}$ is the expected value for the noisy (quantum-mechanical) expectation value $\langle\psi|\tilde{\mathcal{H}}|\psi\rangle = \tilde{E}$.

For a large number of shots s, the noisy energy histograms can be described by a normal distribution with mean energy $\mathbb{E}\tilde{E}$ given by Eq. (A3). The only free parameter of this measurement noise model is $\Delta\tilde{E}$, since s and p are known input parameters.

2. Mean energy vanishes for $p = 0.5$

The first step towards eliminating the free parameter $\Delta\tilde{E}$ is to study the dependence of this parameter on the bit-flip probability p, for example for $p = 0.5$. Let us consider the noise-free Hamiltonian $\mathcal{H}$ acting on the state $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$ and yielding the energy

$$\langle\psi|\mathcal{H}|\psi\rangle = (c_0^*\langle 0| + c_1^*\langle 1|)\, \mathcal{H}\, (c_0|0\rangle + c_1|1\rangle) = E. \tag{A4}$$

The noisy measurement of this energy on the quantum hardware is performed along the basis $Z = \mathrm{diag}(1, -1)$. We note that this noisy measurement yields $\mathbb{E}\tilde{E} = 0$ for $p = 0.5$, due to the opposite signs of the terms resulting from bit-flipping the terms in Eq. (A4).

Let us demonstrate this for a simple example, $\mathcal{H}_X = X = HZH$, where H is the Hadamard gate, and study the possible outcomes of the energy measurements in the single-qubit case:

(i) The absence of any bit flip gives the true energy of the noise-free Hamiltonian:

$$\begin{aligned}\langle\psi|\mathcal{H}_X|\psi\rangle &= \langle\psi|HZH|\psi\rangle \\ &= [c_0^*(\langle 0|+\langle 1|)+c_1^*(\langle 0|-\langle 1|)] \\ &\quad\times Z\,[c_0(|0\rangle+|1\rangle)+c_1(|0\rangle-|1\rangle)] \\ &= |c_0+c_1|^2-|c_0-c_1|^2 = E. \end{aligned} \tag{A5}$$

(ii) The bit flip $|0\rangle \to |1\rangle$, $|1\rangle \to |1\rangle$ changes one sign: $\langle\psi|\tilde{\mathcal{H}}_X|\psi\rangle = -|c_0+c_1|^2-|c_0-c_1|^2$.

(iii) The bit flip $|0\rangle \to |0\rangle$, $|1\rangle \to |0\rangle$ changes the other sign: $\langle\psi|\tilde{\mathcal{H}}_X|\psi\rangle = |c_0+c_1|^2+|c_0-c_1|^2$.

(iv) The bit flip $|0\rangle \to |1\rangle$, $|1\rangle \to |0\rangle$ changes both signs: $\langle\psi|\tilde{\mathcal{H}}_X|\psi\rangle = -|c_0+c_1|^2+|c_0-c_1|^2$ and thus yields the outcome $-E$.

For $p=0.5$, each of these four possible outcomes has the same probability $p^2=0.25$, and thus cancellation yields $\mathbb{E}\tilde{E}=0$. This result holds true for any operator and any number of qubits, as we showed in Eq. (29).

3. $\Delta\tilde{E}$ for noninteracting Hamiltonians

The next step towards eliminating the free parameter $\Delta\tilde{E}$ is to examine the four possible measurement outcomes from the previous section for *any* bit-flip probability p. We observe that the second and third outcomes have *opposite* signs and *equal* probability and thus cancel, given the above assumption that $p(|0\rangle \to |1\rangle) = p(|1\rangle \to |0\rangle)$. For N qubits, one can similarly show that among the 4^N possible measurement outcomes, *all* outcomes cancel apart from the ones corresponding to no bit flip and all bit flips. This justifies our previous assumption that we either measure a correct energy with probability $1-p$ or an incorrect energy with probability p. Crucially, the latter probability is *not* given by p^{2N} as one might expect at first glance. Thus, each incorrect measurement yields the same deviation from the correct energy of $-2E$ with the same probability p. This can be seen by evaluating the probabilities of the four different outcomes above:

(i) The absence of any bit flip, $|0\rangle \xrightarrow{1-p} |0\rangle$, $|1\rangle \xrightarrow{1-p} |1\rangle$, gives $\langle\psi|\mathcal{H}_X|\psi\rangle = E$ with probability $(1-p)^2$.

(ii) The "mixed" bit flips $|0\rangle \xrightarrow{p} |1\rangle$, $|1\rangle \xrightarrow{1-p} |1\rangle$ and $|0\rangle \xrightarrow{1-p} |0\rangle$, $|1\rangle \xrightarrow{p} |0\rangle$ give $\langle\psi|\tilde{\mathcal{H}}_X|\psi\rangle = 0$ with a combined probability of $2p(1-p)$.

(iii) The "total" bit flip $|0\rangle \xrightarrow{p} |1\rangle$, $|1\rangle \xrightarrow{p} |0\rangle$ gives $\langle\psi|\tilde{\mathcal{H}}_X|\psi\rangle = -E$ with probability p^2.

Note that the cancellation $\langle\psi|\tilde{\mathcal{H}}|\psi\rangle = 0$ for the "mixed" bit flips seems to require $|\mathcal{H}|0\rangle| = |\mathcal{H}|1\rangle|$ at first glance. While this is not true in general, the measurement of $\langle\psi|\tilde{\mathcal{H}}|\psi\rangle$ in the Z basis reduces to measuring Pauli strings composed of $\mathbb{1}$ and Z matrices, which are unitary. Thus, after the appropriate postrotation, the condition $|\mathcal{H}|0\rangle| = |\mathcal{H}|1\rangle|$ changes to $|\mathcal{Z}|0\rangle| = |\mathcal{Z}|1\rangle|$ and $|\mathbb{1}|0\rangle| = |\mathbb{1}|1\rangle|$, which is trivially fulfilled.

This yields the simple relation for the mean energy,

$$\begin{aligned}\mathbb{E}\tilde{E} &= (1-p)^2E + p^2(-E) \\ &= (1-2p)E. \end{aligned} \tag{A6}$$

Combining Eqs. (A3) and (A6), we find for the parameter $\Delta\tilde{E}$

$$\mathbb{E}\tilde{E} = E + sp\Delta\tilde{E} \leftrightarrow \Delta\tilde{E} = -\frac{2E}{s}, \tag{A7}$$

where $\Delta\tilde{E}$ is normalized by the number of shots s, i.e., the number of evaluations of the energy (A4) required to produce the energy histogram. For $p=1$, the first three possible measurement outcomes have zero probability, independently of any cancellations, and only the last outcome with $\langle\psi|\tilde{\mathcal{H}}_X|\psi\rangle = -E$ contributes.

As we will discuss in the next subsection, Eq. (A7) *only* applies to noninteracting Hamiltonians, i.e., without any multiqubit interaction terms. For example, for the Hamiltonians $\mathcal{H}_X = h\sum_{i=1}^N X_i$ or $\mathcal{H}_Z = h\sum_{i=1}^N Z_i$ with the ground-state energy $E_0 = -Nh$, we would get $\Delta\tilde{E}_0 = 2Nh/s$ when measuring the ground-state energy. Thus, after measuring the noisy expectation value of any (trivial) noninteracting Hamiltonian on a quantum computer, Eq. (A7) allows us to predict the corresponding true energy.

4. $\Delta\tilde{E}$ for interacting Hamiltonians

For two-qubit interaction terms in the Hamiltonian, e.g., for $\mathcal{H}_{ZZ} = J\sum_{i=1}^N Z_iZ_{i+1}$, our previous considerations need to be modified in two ways: first, we observe that the *one-qubit* bit flips from the previous subsection give the same contribution to the mean energy as before, but now with a probability of $2p(1-p)$ instead of p^2. This is because the one-qubit "total" bit flips yield $\langle\psi|\tilde{\mathcal{H}}_{ZZ}|\psi\rangle = -E$. Here, "one-qubit 'total' bit flip" means that one of the two qubits experiences a bit flip during readout ($|0\rangle \to |1\rangle$, $|1\rangle \to |0\rangle$), while the other qubit has no bit flip ($|0\rangle \to |0\rangle$, $|1\rangle \to |1\rangle$). Second, the mean energy receives small $O(p^2)$ corrections since the parameter $\Delta\tilde{E}$ becomes p-dependent for the interacting case. These $O(p^2)$ corrections come from the *two-qubit* bit flips and have the *opposite* sign of the $O(p)$ terms, because the two minus signs from the measurement bases Z_1 and Z_2 cancel. Indeed, the two-qubit "total" bit flips, i.e., $|0\rangle \to |1\rangle$ and $|1\rangle \to |0\rangle$ for both qubits, yield $\langle\psi|\tilde{\mathcal{H}}_{ZZ}|\psi\rangle = E$ with probability p^2.

Let us demonstrate the latter for the simple two-qubit Hamiltonian $\mathcal{H}_{ZZ} = Z_1Z_2$, which gives

$$\begin{aligned}\langle\psi|\mathcal{H}_{11}|\psi\rangle &= \langle\psi|Z_1Z_2|\psi\rangle \\ &= [c_0^*\langle 00|+c_1^*\langle 01|+c_2^*\langle 10|+c_3^*\langle 11|]Z_1Z_2 \\ &\quad\times[c_0|00\rangle+c_1|01\rangle+c_2|10\rangle+c_3|11\rangle] \\ &= |c_0|^2-|c_1|^2-|c_2|^2+|c_3|^2 = E \end{aligned} \tag{A8}$$

without any bit flip. For two-qubit bit flips with $p=1$, we obtain the same result and thus recover the true energy E,

$$\begin{aligned}\langle\psi|\tilde{\mathcal{H}}_{ZZ}|\psi\rangle &= \langle\psi|\tilde{Z}_1\tilde{Z}_2|\psi\rangle \\ &= [c_0^*\langle 11|+c_1^*\langle 10|+c_2^*\langle 01|+c_3^*\langle 00|]Z_1Z_2 \\ &\quad\times[c_1|11\rangle+c_2|10\rangle+c_3|01\rangle+c_4|00\rangle] \\ &= |c_0|^2-|c_1|^2-|c_2|^2+|c_3|^2 = E, \end{aligned} \tag{A9}$$

since the two minus signs from the Z-matrices cancel, i.e., $\langle 00|Z_1Z_2|00\rangle = \langle 11|Z_1Z_2|11\rangle$.

The contributions from "mixed" bit flips, such as all basis states $|b_1b_0\rangle$ flipping to $|11\rangle$, cancel for any p due to opposite signs and equal probabilities, just as in the noninteracting case. Therefore, the "total" two-qubit bit flips as discussed in Eq. (A9) have a probability of p^2 instead of p^{4N}. This yields for the total mean energy

$$\mathbb{E}\tilde{E} = (1-p)^2 E + 2p(1-p)(-E) + p^2 E$$
$$= E - 4pE + 4p^2 E = (1-2p)^2 E. \quad \text{(A10)}$$

Thus, the parameter $\Delta\tilde{E}$ now has two contributions,

$$\mathbb{E}\tilde{E} = E + sp\Delta\tilde{E} \;\Leftrightarrow\; \Delta\tilde{E} = -\frac{4E}{s}(1-p). \quad \text{(A11)}$$

Equations (A10) and (A11) imply that the two-qubit interacting Hamiltonian yields the correct energy $\mathbb{E}\tilde{E} = E$ for *both* $p = 0$ and 1, in contrast to the noninteracting case where $p = 1$ gave $\mathbb{E}\tilde{E} = -E$ [see Eq. (A6)]. Moreover, $\mathbb{E}\tilde{E} = 0$ is still given for $p = 0.5$.

5. Prediction for the transversal Ising model

Next, we apply our results to the ground-state energy of the TI model with the Hamiltonian

$$\mathcal{H}_{\mathrm{TI}} = J\sum_{i=1}^{N} Z_i Z_{i+1} + h\sum_{i=1}^{N} X_i, \quad \text{(A12)}$$

where we again assume $J < 0$ and $h > 0$ and periodic boundary conditions. The true ground-state energy can be derived as [57–61]

$$E_0 = -\frac{1}{2}\sum_k \gamma\,(\alpha^2 + 4\beta^2)$$
$$= -\frac{1}{2}\sum_k \gamma\left[4h^2 + 4J^2 - 8Jh\cos\left(\frac{2\pi k}{N}\right)\right], \quad \text{(A13)}$$

where the sum runs from $k = -(\frac{N-1}{2})$ to $(\frac{N-1}{2})$, and the constants α, β, and γ are defined as

$$\alpha = 2h - 2J\cos\left(\frac{2\pi k}{N}\right),$$
$$\beta = J\sin\left(\frac{2\pi k}{N}\right),$$
$$\gamma = \frac{\mathrm{sgn}(\alpha)}{\alpha}\sqrt{\frac{\alpha^2}{\alpha^2 + 4\beta^2}}. \quad \text{(A14)}$$

Just as for the LI model (7), the mean energy of the noisy ground-state energy histograms receives three different contributions,

$$\mathbb{E}\tilde{E}_0 = (1-p)^2 E_1 + 2p(1-p)E_2 + p^2 E_3. \quad \text{(A15)}$$

The probabilities of the three different terms in Eqs. (7) and (A15) are the same because they are determined by the number of interacting qubits in the different terms of the respective Hamiltonian. However, the measurement outcomes E_i in Eq. (A15) deviate from the ones in Eq. (7) because E_0 in Eq. (A13) is not simply the sum of the J- and h-dependent parts of the ground-state energy as in Eq. (4).

The different measurement outcomes E_i in Eq. (A15) can be derived in the following way. First, we know that $E_1 = E_0$. Second, we know that $\mathbb{E}\tilde{E}$ vanishes for $|J| = |h|$ and $p = 1$ because the two terms in the Hamiltonian (A12) contribute equally to E_0 and thus cancel for $p = 1$. This cancellation happens due to opposite signs of the noninteracting and interacting energy contributions in the case of a total bit flip, as discussed above. In particular, any mixed terms, such as the mixed Jh-term in Eq. (A13), vanish for $p = 1$, as also discussed above. This fixes E_3. Third, we know that $\mathbb{E}\tilde{E}(p = 0.5) = 0$, so we can find E_2 by solving Eq. (A15) for $p = 0.5$ and inserting the known expressions for E_1 and E_3. In total, we obtain

$$E_1 = E_{ZZ} + E_X,$$
$$E_2 = -E_{ZZ}, \quad \text{(A16)}$$
$$E_3 = E_{ZZ} - E_X,$$

which is similar to Eq. (7), but with E_{ZZ} and E_X given by

$$E_{ZZ} = -\frac{1}{2}\sum_k \gamma\left[4J^2 - 4Jh\cos\left(\frac{2\pi k}{N}\right)\right],$$
$$E_X = -\frac{1}{2}\sum_k \gamma\left[4h^2 - 4Jh\cos\left(\frac{2\pi k}{N}\right)\right]. \quad \text{(A17)}$$

Thus, the mean energy in Eq. (A15) can be brought into a similar form as the true ground-state energy in Eq. (A13),

$$\mathbb{E}\tilde{E}_0 = (1-2p)E_X + (1-2p)^2 E_{ZZ}$$
$$= -\frac{1}{2}\sum_k \gamma\left[(1-2p)4h^2 + (1-2p)^2 4J^2\right.$$
$$\left. - (1 - 3p + 2p^2)8Jh\cos\left(\frac{2\pi k}{N}\right)\right]. \quad \text{(A18)}$$

The resulting parameter $\Delta\tilde{E}_0$ now has three contributions,

$$\Delta\tilde{E}_0 = -\frac{1}{s}(2E_X + 4E_{ZZ} - 4pE_{ZZ}). \quad \text{(A19)}$$

We note that this expression is identical to the one for the LI model but with different E_{ZZ} and $E_{Z/X}$. For the LI model, $\Delta\tilde{E}_0$ rises strictly linearly with N. For the TI model, the sum over k yields N contributions to each E_i in Eq. (A16), which are equal for E_3 but differ for E_1 and E_2 due to the N-dependence of the cosine in Eq. (A18). Thus, $\Delta\tilde{E}_0(N)$ only becomes approximately linear for large N, where these small differences average out.

In Fig. 7, we plot the energy histograms for the ground state of $\mathcal{H}_{\mathrm{TI}}$ with different N, J, h, s, and p, where we again measure the ground state 2048 times for each parameter combination. As before, the noise model with the mean energy from Eq. (A18) and the variance from Eq. (B17) agrees with the data for any choice of parameters we study. Note that the variance is larger compared to the longitudinal case in Fig. 1, because the measurement Z-basis is not an eigenbasis of the X_i operator. Thus, the histograms are wider for the transversal case.

APPENDIX B: PREDICTION FOR THE VARIANCES OF NOISY EXPECTATION VALUES

In this Appendix, we derive the variances of the different noisy expectation values presented in Secs. II and III. To this

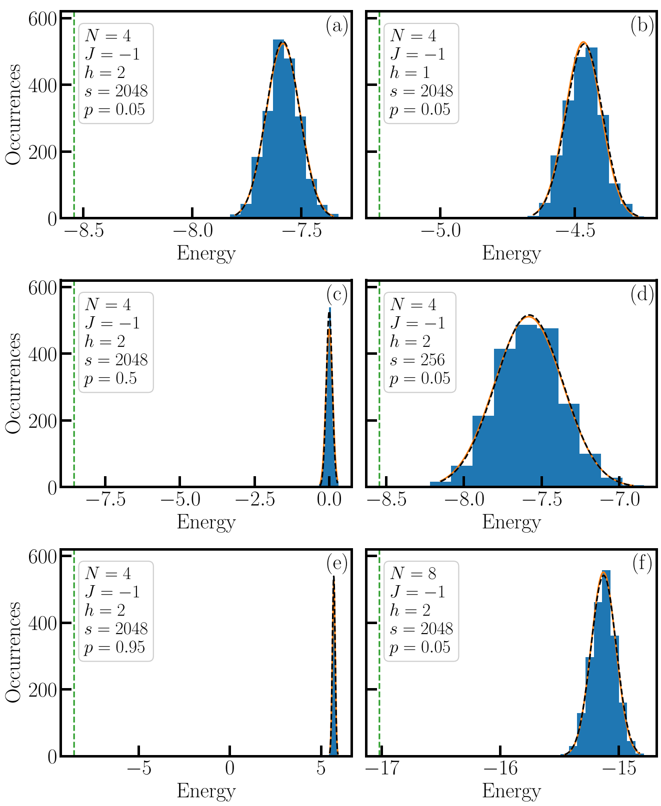


FIG. 7. Energy histograms for the TI model. The vertical dashed green line indicates the true ground-state energy, the solid orange line indicates the prediction, and the dashed black line indicates a fit to the data. The left column corresponds to $N = 4$, $J = -1$, $h = 2$, $s = 2048$ with (a) $p = 0.05$, (c) $p = 0.50$, and (e) $p = 0.95$. The right column shows varied N, h, and s: (b) $h = 1$, (d) $s = 256$, and (f) $N = 8$.

end, we construct again random operators whose expectation value yields the variance with respect to the bit-flip probability. We follow the structure of Sec. III: we first discuss a single Z operator in Appendix B 1, followed by the general case of $Z_Q \otimes \cdots \otimes Z_1$ operators in Appendix B 2. We then simplify our results to the case of equal bit-flip probabilities in Appendix B 3. Finally, we discuss the case of measuring general operators from bit-string distributions of $|\psi\rangle$ in Appendix B 4. In this Appendix, we will also discuss different measurement paradigms and their impact on the variance of means extracted from histogram data. We will eventually return to the TI model for an explicit illustration.

1. Measurement of a single Z operator

For computing the variance $\mathbb{V}\tilde{Z}_q$ of the noisy expectation in Eq. (8),

$$\begin{aligned}\mathbb{V}\tilde{Z}_q &= \mathbb{E}(\tilde{Z}_q \otimes \tilde{Z}_q) - (\mathbb{E}\tilde{Z}_q)^2 \\ &= \Phi'_{\tilde{Z}_q}(0) \otimes \Phi'_{\tilde{Z}_q}(0) - \Phi''_{\tilde{Z}_q}(0),\end{aligned} \tag{B1}$$

we need to evaluate the derivatives $\Phi'_q(0) = i\mathbb{E}\tilde{Z}_q$ and $\Phi''_{\tilde{Z}_q}(0) = -\mathbb{E}(\tilde{Z}_q)^2$ of the characteristic function

$$\Phi_{\tilde{Z}_q}(t) := \mathbb{E}\exp[i\,\mathrm{Tr}(t^*\tilde{Z}_q)]. \tag{B2}$$

This yields

$$\begin{aligned}\Phi'_{\tilde{Z}_q}(0) &= i(1 - p_{q,0} - p_{q,1})Z_q + i(p_{q,1} - p_{q,0})\mathbb{1}_q, \\ \Phi''_{\tilde{Z}_q}(0) &= -(1 - p_{q,0} - p_{q,1} + 2p_{q,0}p_{q,1})Z_q \otimes Z_q \\ &\quad - (p_{q,0} + p_{q,1} - 2p_{q,0}p_{q,1})\mathbb{1}_q \otimes \mathbb{1}_q.\end{aligned} \tag{B3}$$

Thus, the variance operator in Eq. (B1) reads

$$\begin{aligned}\mathbb{V}\tilde{Z}_q = &[(p_{q,0} + p_{q,1})(1 - p_{q,0} - p_{q,1}) + 2p_{q,0}p_{q,1}] \\ &\times Z_q \otimes Z_q \\ &- (1 - p_{q,0} - p_{q,1})(p_{q,1} - p_{q,0})Z_q \otimes \mathbb{1}_q \\ &- (1 - p_{q,0} - p_{q,1})(p_{q,1} - p_{q,0})\mathbb{1}_q \otimes Z_q \\ &+ (p_{q,0} + p_{q,1} - p_{q,0}^2 - p_{q,1}^2)\mathbb{1}_q \otimes \mathbb{1}_q.\end{aligned} \tag{B4}$$

2. Measurement of $Z_Q \otimes \cdots \otimes Z_1$ operators

We now generalize the variance for $Q = 1$ in Eq. (B4) to operators acting on multiple qubits, i.e., $Q > 1$. According to Eq. (16), operators $\tilde{O}_1$ and $\tilde{O}_2$ acting on different qubits are uncorrelated, i.e., the covariance vanishes,

$$\begin{aligned}\mathrm{Cov}_\otimes(\tilde{O}_1, \tilde{O}_2) &:= \mathbb{E}(\tilde{O}_1 \otimes \tilde{O}_2) - \mathbb{E}(\tilde{O}_1) \otimes \mathbb{E}(\tilde{O}_2) \\ &= 0.\end{aligned} \tag{B5}$$

Hence, we obtain the variance operator

$$\begin{aligned}&\mathbb{V}(\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1) \\ &\quad= \mathbb{E}(\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1 \otimes \tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1) \\ &\qquad - \mathbb{E}(\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1) \otimes \mathbb{E}(\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1) \\ &\quad= U^*\Bigg(\mathbb{E}(\tilde{Z}_Q \otimes \tilde{Z}_Q) \otimes \cdots \otimes \mathbb{E}(\tilde{Z}_1 \otimes \tilde{Z}_1) \\ &\qquad - \bigotimes_{q=1}^{Q} (\mathbb{E}\tilde{Z}_q \otimes \mathbb{E}\tilde{Z}_q)\Bigg)U \\ &\quad= U^*\Bigg((\mathbb{V}\tilde{Z}_Q + \mathbb{E}\tilde{Z}_Q \otimes \mathbb{E}\tilde{Z}_Q) \otimes \cdots \\ &\qquad \cdots \otimes (\mathbb{V}\tilde{Z}_1 + \mathbb{E}\tilde{Z}_1 \otimes \mathbb{E}\tilde{Z}_1) \\ &\qquad - \bigotimes_{q=1}^{Q} (\mathbb{E}\tilde{Z}_q \otimes \mathbb{E}\tilde{Z}_q)\Bigg)U,\end{aligned} \tag{B6}$$

where the unitary operation U reorders the tensor products from $|\psi_Q\rangle \otimes \cdots \otimes |\psi_1\rangle \otimes |\psi_Q\rangle \otimes \cdots \otimes |\psi_1\rangle$ to $(|\psi_Q\rangle \otimes |\psi_Q\rangle) \otimes \cdots \otimes (|\psi_1\rangle \otimes |\psi_1\rangle)$. That is, for two qubits the reordering maps the basis state $|b_3b_2b_1b_0\rangle$ to $|b_3b_1b_2b_0\rangle$, and for three qubits the reordering maps $|b_5b_4b_3b_2b_1b_0\rangle$ to $|b_5b_2b_4b_1b_3b_0\rangle$, etc.

3. Measurement of $Z_Q \otimes \cdots \otimes Z_1$ operators assuming equal bit-flip probabilities

For $Q = 1$, the variance in Eq. (B6) reduces to

$$\mathbb{V}\tilde{Z}_q = 2p(1 - p)(Z_q \otimes Z_q + \mathbb{1}_q \otimes \mathbb{1}_q). \tag{B7}$$

For $Q = 2$, the reordering of the tensor product $|\psi\rangle \otimes |\psi\rangle$ in Eq. (B6) becomes important, which yields

$$\begin{aligned}\mathbb{V}(\langle\psi|\tilde{Z}_2 \otimes \tilde{Z}_1|\psi\rangle)\\ &= ((\langle\psi| \otimes \langle\psi|)U^*(\mathbb{V}\tilde{Z}_2 \otimes \mathbb{V}\tilde{Z}_1)U(|\psi\rangle \otimes |\psi\rangle))\\ &+ ((\langle\psi| \otimes \langle\psi|)U^*(\mathbb{V}\tilde{Z}_2 \otimes \mathbb{E}\tilde{Z}_1 \otimes \mathbb{E}\tilde{Z}_1)U(|\psi\rangle \otimes |\psi\rangle))\\ &+ ((\langle\psi| \otimes \langle\psi|)U^*(\mathbb{E}\tilde{Z}_2 \otimes \mathbb{E}\tilde{Z}_2 \otimes \mathbb{V}\tilde{Z}_1)U(|\psi\rangle \otimes |\psi\rangle)).\end{aligned} \tag{B8}$$

For arbitrary Q, we can evaluate the variance operator in Eq. (B6) for the ground state, $|\psi\rangle = |0, \dots, 0\rangle$, and we obtain the expression

$$\begin{aligned}\mathbb{V}(\langle\psi|\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1|\psi\rangle)\\ &= \prod_{q=1}^{Q}\langle\psi|\mathbb{V}\tilde{Z}_q + \mathbb{E}\tilde{Z}_q \otimes \mathbb{E}\tilde{Z}_q|\psi\rangle\\ &\quad - \prod_{q=1}^{Q}\langle\psi|\mathbb{E}\tilde{Z}_q \otimes \mathbb{E}\tilde{Z}_q|\psi\rangle\\ &= \prod_{q=1}^{Q}\left(4p(1-p) + (1-2p)^2\right) - (1-2p)^{2Q}\\ &= 1 - (1-2p)^{2Q}.\end{aligned} \tag{B9}$$

This surprisingly simple result can be verified directly by noting that the measurement of $\langle 0, \dots, 0|\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1|0, \dots, 0\rangle$ yields the values $+1$ with probability p_1 and -1 with probability p_{-1}. Thus, we conclude

$$\begin{aligned}\mathbb{V}(\langle 0, \dots, 0|\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1|0, \dots, 0\rangle)\\ &= \mathbb{E}\left(\langle 0, \dots, 0|\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1|0, \dots, 0\rangle^2\right)\\ &\quad - \mathbb{E}(\langle 0, \dots, 0|\tilde{Z}_Q \otimes \cdots \otimes \tilde{Z}_1|0, \dots, 0\rangle)^2\\ &= p_1 + p_{-1}(-1)^2\\ &\quad - (1-2p)^{2Q}\langle 0, \dots, 0|Z_Q \otimes \cdots \otimes Z_1|0, \dots, 0\rangle^2\\ &= 1 - (1-2p)^{2Q}.\end{aligned} \tag{B10}$$

4. Measurement of general operators $\mathcal{H}$ from bit-string distributions of $|\psi\rangle$

a. Prediction for the variance of operators

While measuring the entire Hamiltonian simultaneously makes no difference for the measured mean value, the variance, on the other hand, is affected by this change in measurement paradigm. If we consider $\mathcal{H}_{ZZ}$ with $N = 2$ and $J = 1$, i.e., $\mathcal{H}_{ZZ} = Z_2Z_1 + Z_1Z_2$, then we would formally compute $\langle\psi|Z_2 \otimes Z_1|\psi\rangle$ twice independently using the approach considered so far, whereas the expectation from the bit-string distribution of $|\psi\rangle$ directly extracts $2\langle\psi|Z_2 \otimes Z_1|\psi\rangle$. Thus the variance using independent histograms for each summand is given by

$$\begin{aligned}\mathbb{V}_{\text{ind}}\langle\psi|\tilde{\mathcal{H}}_{ZZ}|\psi\rangle &= \mathbb{V}\langle\psi|\tilde{Z}_2 \otimes \tilde{Z}_1|\psi\rangle + \mathbb{V}\langle\psi|\tilde{Z}_2 \otimes \tilde{Z}_1|\psi\rangle\\ &= 2\mathbb{V}\langle\psi|\tilde{Z}_2 \otimes \tilde{Z}_1|\psi\rangle,\end{aligned} \tag{B11}$$

whereas the variance using the bit-string distribution of $|\psi\rangle$ is

$$\begin{aligned}\mathbb{V}_{\text{bsd}}\langle\psi|\tilde{\mathcal{H}}_{ZZ}|\psi\rangle &= \mathbb{V}(2\langle\psi|\tilde{Z}_2 \otimes \tilde{Z}_1|\psi\rangle)\\ &= 4\mathbb{V}\langle\psi|\tilde{Z}_2 \otimes \tilde{Z}_1|\psi\rangle\\ &= 2\mathbb{V}_{\text{ind}}\langle\psi|\tilde{\mathcal{H}}_{ZZ}|\psi\rangle.\end{aligned} \tag{B12}$$

In general, if $\tilde{\mathcal{H}} = \sum_\alpha \lambda_\alpha U_\alpha^* \tilde{O}_\alpha U_\alpha$, we are still able to predict the variance $\mathbb{V}_{\text{bsd}}\langle\psi|\tilde{\mathcal{H}}|\psi\rangle$ using the same method as above, albeit the covariance terms no longer vanish (each $\tilde{O}_\alpha$ is a tensor product $\tilde{O}_{\alpha,Q} \otimes \cdots \otimes \tilde{O}_{\alpha,1}$). For $\tilde{O}_{\alpha,q} = \tilde{Z}_q$, $\tilde{O}_{\alpha,q}$ takes one of the possible values $\{Z_q, -\mathbb{1}_q, \mathbb{1}_q, -Z_q\}$, as in Sec. III A. For $\tilde{O}_{\alpha,q} = \tilde{\mathbb{1}}_q$, $\tilde{O}_{\alpha,q}$ always takes the value $\mathbb{1}_q$. Using these replacements for all summands in $\tilde{\mathcal{H}}$, we obtain that $\tilde{\mathcal{H}}$ takes finitely many (up to 2^N) values $\mathcal{H}_\alpha$ with probability p_α. Hence, the characteristic function $\Phi_{\tilde{\mathcal{H}}}$ is given by

$$\begin{aligned}\Phi_{\tilde{\mathcal{H}}}(t) &:= \mathbb{E}\exp[i\text{Tr}(t^*\tilde{\mathcal{H}})]\\ &= \sum_\alpha p_\alpha \exp[i\text{Tr}(t^*\mathcal{H}_\alpha)].\end{aligned} \tag{B13}$$

As such, we can directly conclude

$$\Phi'_{\tilde{\mathcal{H}}}(0) = \sum_\alpha p_\alpha i\mathcal{H}_\alpha = i\mathbb{E}\tilde{\mathcal{H}}, \tag{B14}$$

$$\Phi''_{\tilde{\mathcal{H}}}(0) = -\sum_\alpha p_\alpha \mathcal{H}_\alpha \otimes \mathcal{H}_\alpha, \tag{B15}$$

and we find the variance operator

$$\begin{aligned}\mathbb{V}_{\text{bsd}}\tilde{\mathcal{H}} &= \Phi'_{\tilde{\mathcal{H}}}(0) \otimes \Phi'_{\tilde{\mathcal{H}}}(0) - \Phi''_{\tilde{\mathcal{H}}}(0)\\ &= \left(\sum_\alpha p_\alpha \mathcal{H}_\alpha \otimes \mathcal{H}_\alpha\right) - (\mathbb{E}\tilde{\mathcal{H}}) \otimes (\mathbb{E}\tilde{\mathcal{H}})\\ &= \left(\sum_\alpha p_\alpha \mathcal{H}_\alpha \otimes \mathcal{H}_\alpha\right) - \left(\sum_{\alpha,\beta} p_\alpha p_\beta \mathcal{H}_\alpha \otimes \mathcal{H}_\beta\right).\end{aligned} \tag{B16}$$

Similarly, we can measure the operator $\tilde{\mathcal{H}}$ on the state $|\psi\rangle$ and obtain the variance

$$\begin{aligned}\mathbb{V}_{\text{bsd}}\langle\psi|\tilde{\mathcal{H}}|\psi\rangle &= \left(\sum_\alpha p_\alpha \langle\psi|\mathcal{H}_\alpha|\psi\rangle^2\right)\\ &\quad - \left(\sum_{\alpha,\beta} p_\alpha p_\beta \langle\psi|\mathcal{H}_\alpha|\psi\rangle\langle\psi|\mathcal{H}_\beta|\psi\rangle\right).\end{aligned} \tag{B17}$$

b. Prediction for the variance of histogram means

Lastly, we can combine the bit-flip induced variances with quantum mechanically induced variances to obtain the full variances observed in measuring histogram means. In particular, we will construct the variances for the three methods discussed in Sec. III E. There we measured the bit-flip corrected TI Hamiltonian $\tilde{\mathcal{H}}_{\text{TI,bfc}} = J_p \sum_{j=1}^N \tilde{Z}_j\tilde{Z}_{j+1} + h_p \sum_{j=1}^N \tilde{X}_j$ in Eq. (37) subject to bit flips on the ground state of the "true" TI Hamiltonian $\mathcal{H}_{\text{TI}} = J\sum_{j=1}^N Z_jZ_{j+1} + h\sum_{j=1}^N X_j$. For simplicity, we assumed that all bit-flip probabilities $p_{q,b}$ equal p. The three methods are as follows:

(i) *Method 1.* measure each $\tilde{Z}_j\tilde{Z}_{j+1}$ and $\tilde{X}_j$ in Eq. (37) independently.

(ii) *Method 2.* measure the entire Hamiltonian $\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}$ in Eq. (37) from distributions of $|\psi\rangle$ measurements.

(iii) *Method 3.* measure $\tilde{\mathcal{H}}_{ZZ} := J_p \sum_{j=1}^{N} \tilde{Z}_j\tilde{Z}_{j+1}$ and $\tilde{\mathcal{H}}_X := h_p \sum_{j=1}^{N} \tilde{X}_j$ independently from distributions of $|\psi\rangle$ measurements.

Method 1. Since each $\tilde{Z}_j\tilde{Z}_{j+1}$ and $\tilde{X}_j$ is measured independently, the bit-flip contributions $\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{Z}_j\tilde{Z}_{j+1}|\psi\rangle$ and $\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{X}_j|\psi\rangle$ to the variance can be directly obtained from Eq. (B6), keeping in mind that $\tilde{X}_j = H_j\tilde{Z}_jH_j$, where H_j is the Hadamard gate on qubit j. But since $|\psi\rangle$, in general, will not be an eigenstate of all Z_jZ_{j+1} and X_j simultaneously, we also have a contribution from the quantum-mechanical variances $\mathbb{V}_{\mathrm{QM}}\langle\psi|Z_jZ_{j+1}|\psi\rangle = 1 - \langle\psi|Z_jZ_{j+1}|\psi\rangle^2$ and $\mathbb{V}_{\mathrm{QM}}\langle\psi|X_j|\psi\rangle = 1 - \langle\psi|X_j|\psi\rangle^2$. We therefore obtain the variance of histogram means

$$\begin{aligned}\mathbb{V}_{\mathrm{M1}}\langle\psi|\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}|\psi\rangle &= \frac{J_p^2}{s}\sum_{j=1}^{N}\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{Z}_j\tilde{Z}_{j+1}|\psi\rangle + \frac{J_p^2}{s}\sum_{j=1}^{N}\mathbb{V}_{\mathrm{QM}}\langle\psi|Z_jZ_{j+1}|\psi\rangle \\ &\quad + \frac{h_p^2}{s}\sum_{j=1}^{N}\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{X}_j|\psi\rangle + \frac{h_p^2}{s}\sum_{j=1}^{N}\mathbb{V}_{\mathrm{QM}}\langle\psi|X_j|\psi\rangle. \end{aligned} \tag{B18}$$

In particular, if the state $|\psi\rangle$ is translationally invariant, such as the ground state of $\mathcal{H}_{\mathrm{TI}}$, then this further simplifies to

$$\begin{aligned}\mathbb{V}_{\mathrm{M1}}\langle\psi|\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}|\psi\rangle &= \frac{J_p^2N}{s}\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{Z}_j\tilde{Z}_{j+1}|\psi\rangle + \frac{J_p^2N}{s}\mathbb{V}_{\mathrm{QM}}\langle\psi|Z_jZ_{j+1}|\psi\rangle \\ &\quad + \frac{h_p^2N}{s}\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{X}_j|\psi\rangle + \frac{h_p^2N}{s}\mathbb{V}_{\mathrm{QM}}\langle\psi|X_j|\psi\rangle \end{aligned} \tag{B19}$$

for any choice of j.

Method 2. In this case, the bit-flip contribution $\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}|\psi\rangle$ is given by Eq. (B17) and the quantum-mechanical variance is given by

$$\mathbb{V}_{\mathrm{QM}}\langle\psi|\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}|\psi\rangle = \langle\psi|(\tilde{\mathcal{H}}_{\mathrm{TI,bfc}})^2|\psi\rangle - \langle\psi|\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}|\psi\rangle^2. \tag{B20}$$

Hence, the variance of histogram means is

$$\mathbb{V}_{\mathrm{M2}}\langle\psi|\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}|\psi\rangle = \frac{1}{s}\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}|\psi\rangle + \frac{1}{s}\mathbb{V}_{\mathrm{QM}}\langle\psi|\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}|\psi\rangle. \tag{B21}$$

While this expression appears simpler than its counterpart for Method 1, it is also important to note that $O(4^N)$ terms are required to compute $\mathbb{V}_{\mathrm{M2}}\langle\psi|\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}|\psi\rangle$, whereas the number of terms required to compute $\mathbb{V}_{\mathrm{M1}}\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}$ is only $O(N)$ and can even be reduced to $O(1)$ for translationally invariant states $|\psi\rangle$.

Method 3. Being a combination of Method 1 and Method 2, the variance can be constructed combining the results from Methods 1 and 2. The bit-flip contributions $\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{\mathcal{H}}_{ZZ}|\psi\rangle$ and $\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{\mathcal{H}}_X|\psi\rangle$ follow from Eq. (B17) again. Furthermore, the quantum-mechanical variances contribute as

$$\mathbb{V}_{\mathrm{QM}}\langle\psi|\tilde{\mathcal{H}}_{ZZ}|\psi\rangle = \langle\psi|(\tilde{\mathcal{H}}_{ZZ})^2|\psi\rangle - \langle\psi|\tilde{\mathcal{H}}_{ZZ}|\psi\rangle^2 \tag{B22}$$

and

$$\mathbb{V}_{\mathrm{QM}}\langle\psi|\tilde{\mathcal{H}}_X|\psi\rangle = \langle\psi|(\tilde{\mathcal{H}}_X)^2|\psi\rangle - \langle\psi|\tilde{\mathcal{H}}_X|\psi\rangle^2. \tag{B23}$$

The variance of histogram means is thus

$$\begin{aligned}\mathbb{V}_{\mathrm{M3}}\langle\psi|\tilde{\mathcal{H}}_{\mathrm{TI,bfc}}|\psi\rangle &= \frac{1}{s}\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{\mathcal{H}}_{ZZ}|\psi\rangle + \frac{1}{s}\mathbb{V}_{\mathrm{bf}}\langle\psi|\tilde{\mathcal{H}}_X|\psi\rangle \\ &\quad + \frac{1}{s}\mathbb{V}_{\mathrm{QM}}\langle\psi|\tilde{\mathcal{H}}_{ZZ}|\psi\rangle + \frac{1}{s}\mathbb{V}_{\mathrm{QM}}\langle\psi|\tilde{\mathcal{H}}_X|\psi\rangle. \end{aligned} \tag{B24}$$

Methods 1 and 2 are the two extreme cases, which we discussed in Secs. III A–III C and Sec. III D, respectively. Method 3 is a reasonable compromise, which is closely related to implementations of quantum algorithms that are optimized for the number of calls to the quantum device. In such implementations, only parts of an operator can be measured simultaneously, such that both Methods 1 and 2 are impractical to various degrees.

APPENDIX C: TECHNICAL DETAILS OF THE SIMULATIONS

Here we briefly summarize the details on how to determine the bit-flip probabilities, the simulations, and the data evaluation procedure for the results shown in Sec. IV.

1. Calibration of the bit-flip probabilities

Although the QISKIT SDK [46] provides values for the bit-flip probabilities for the different qubits on the different chips, we choose to calibrate $p_{q,0}$ and $p_{q,1}$ ourselves. To obtain $p_{q,0}$, we simply measure the initial state using $s_{\mathrm{calibration}}$ shots and record the number of 1 outcomes. Similarly, we determine $p_{q,1}$ by first applying an X gate to the qubit q, thus preparing the state $|1\rangle$, and we measure the resulting state again $s_{\mathrm{calibration}}$ times and record the number of 0 outcomes. For all data shown in the main text, we use $s_{\mathrm{calibration}} = 8192$, which is the maximum number of repetitions possible on the real quantum hardware. Moreover, to acquire some statistics on how the obtained values for the bit-flip probabilities fluctuate, we repeat this procedure multiple times. Subsequently, we average all the data obtained for $p_{q,b}$. The resulting bit-flip probabilities are the ones used for correcting the data in Sec. IV.

a. Single-qubit case

In Figs. 8 and 9, we show the bit-flip probabilities we obtained for ibmq_london and ibmq_burlington. Looking at the data resulting from simulating ibmq_london classically with readout noise only in Fig. 8(a), we observe that the bit-flip probabilities our calibration procedure yields scatter around the value provided by the noise model. Using the full noise model does not change the picture a lot; only the values for $p_{0,1}$ scatter slightly more around the value of the noise model, as Fig. 8(b) reveals. The data generated on the actual ibmq_london quantum hardware in Fig. 8(c) do not agree very well with the values of the noise model. Even the values for

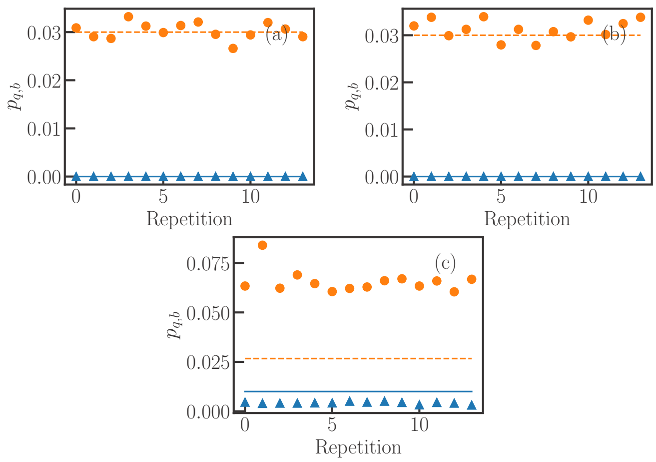


FIG. 8. Bit-flip probabilities $p_{0,0}$ (blue triangles) and $p_{0,1}$ (orange dots) for the single-qubit case measured with the calibration procedure as a function of the repetition for (a) classically simulating ibmq_london with readout errors only, (b) the full noise model, and (c) data obtained on the quantum hardware. The blue solid and the orange dashed line represent the corresponding data provided by the noise model.

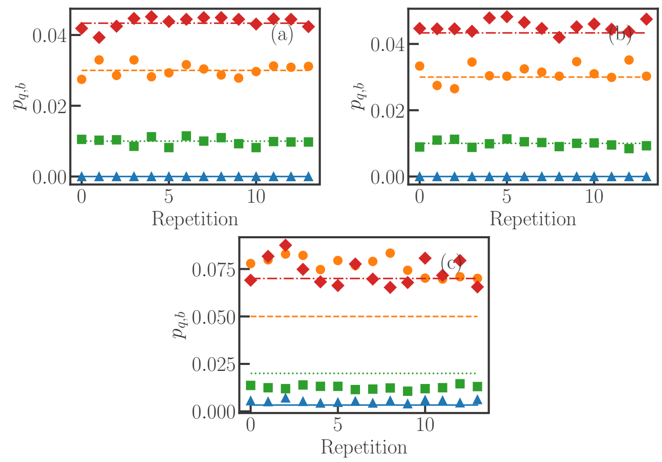


FIG. 10. Bit-flip probabilities $p_{0,0}$ (blue triangles), $p_{0,1}$ (orange dots), $p_{1,0}$ (green squares), and $p_{1,1}$ (red diamonds) for the two-qubit case measured with the calibration procedure as a function of the repetition for (a) classically simulating ibmq_london with readout error only, (b) the full noise model, and (c) data obtained on the hardware. The blue solid, orange dashed, green dotted, and red dash-dotted lines represent the corresponding data provided by the noise model.

$p_{0,0}$, which do not involve a single gate, are in general lower than the value provided by the noise model. In contrast, $p_{0,1}$ exceeds the value of the noise model. Despite the fact that the values for the experimentally obtained bit-flip probabilities deviate from the noise model, they only fluctuate moderately and we can extract a reasonable bit-flip probability by averaging over all repetitions. Comparing the different panels of Fig. 8 closely, one can also observe that the values for the bit-flip probabilities provided by the noise model in panel (c) differ slightly from those in panels (a) and (b). The reason for that is that the data in the noise model are updated every day, and our classical simulations as well as our simulations on real quantum hardware were not carried out the same day.

The corresponding results for imbq_burlington are shown in Fig. 9. Again, the classical simulation of the chip using the noise model produces as expected bit-flip probabilities in agreement with the values provided. Looking at the data from the real chip in Fig. 9(c), we see that these fluctuate over a wide range between different repetitions. Thus, in this case the bit-flip probabilities cannot be extracted as reliably as for imbq_london. Since our correction procedure relies on being able to estimate the bit-flip probabilities precisely, this partially explains why the improvement in Sec. IV A after applying the correction to our data for ibmq_burlington is smaller.

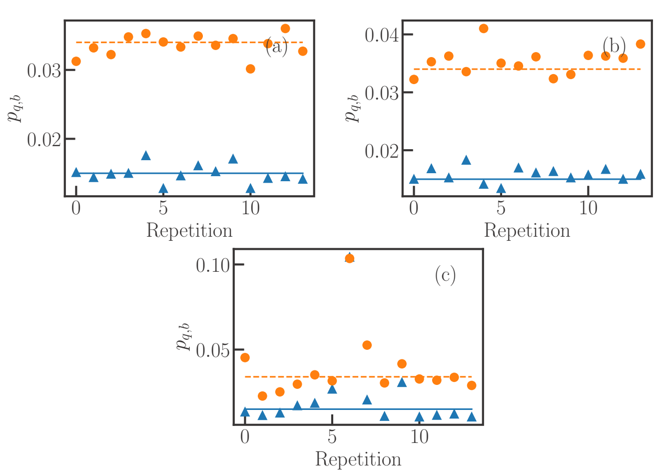


FIG. 9. Bit-flip probabilities $p_{0,0}$ (blue triangles) and $p_{0,1}$ (orange dots) for the single-qubit case measured with the calibration procedure as a function of the repetition for (a) classically simulating ibmq_burlington with readout error only, (b) the full noise model, and (c) data obtained on the hardware. The blue solid and the orange dashed line represent the corresponding data provided by the noise model.

b. Two-qubit case

Analogously to the single-qubit case, Figs. 10 and 11 show the data for extracting the bit-flip probabilities for ibmq_london and ibmq_burlington obtained in our two-qubit simulations. The results for the two-qubit case on ibmq_london in Fig. 10 show a fairly similar behavior to the single-qubit case. The classical simulations in panels (a) and (b) yield as expected good agreement with the values provided in the noise model. In contrast, the data obtained on the real quantum device [Fig. 10(c)] do not agree with the data in the noise model, in particular for $p_{0,1}$ and $p_{1,0}$. Nevertheless, the experimental data are fairly consistent and allow us to reliably determine the bit-flip probabilities for ibmq_london. Again, we see that the theoretical values differ noticeably between the panels in the upper row and the lower row. This is once more due to the fact that the hardware data were taken on a different day from the simulator data, and the noise model was updated in between.

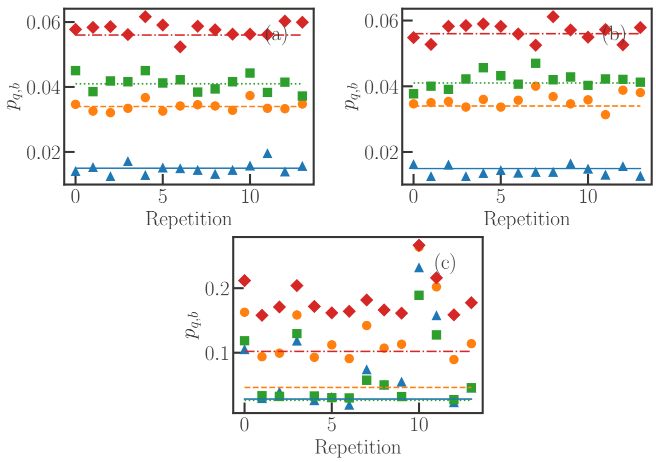


FIG. 11. Bit-flip probabilities $p_{0,0}$ (blue triangles), $p_{0,1}$ (orange dots), $p_{1,0}$ (green squares), and $p_{1,1}$ (red diamonds) for the two-qubit case measured with the calibration procedure as a function of the repetition for (a) classically simulating ibmq_burlington with readout error only, (b) the full noise model, and (c) data obtained on the hardware. The blue solid, orange dashed, green dotted, and red dash-dotted lines represent the corresponding data provided by the noise model.

The bit-flip probabilities obtained from classically simulating ibmq_burlington in Figs. 11(a) and 11(b) show a similar picture to the previous cases, and they agree well with the values provided in the noise model. On the contrary, the data from the real quantum device again do not agree very well with the values provided in the noise model. Moreover, the values for $p_{0,0}$ and $p_{0,1}$ show large fluctuations. In this case as well, the theoretical values for the bit-flip probabilities differ between the simulator data and the hardware data. Most noticeably, the theoretical value for $p_{1,1}$ almost doubled during the time span between carrying out the classical simulations and the experiments on quantum hardware.

2. Technical details for generating the experimental data on quantum hardware

Each of the data points in Figs. 5(c), 5(f) and Figs. 6(c), 6(f) is obtained by preparing 1050 random wave functions $|\psi\rangle$ using the circuits shown in Sec. IV. While running the 1050 circuits is unproblematic for classical simulations, as of completion of this paper one can only submit 75 circuits per job to real quantum hardware. Thus, we divide them into 14 chunks of 75 circuits. This procedure is repeated for every value of s. Since we have to run a considerable number of jobs, which might take some time depending on how busy the queue of the device is, we insert a job running the circuits for determining the bit-flip probabilities before every chunk. This way we can monitor the bit-flip probabilities over the duration of the run and detect potential outliers.

Moreover, before running our circuits, we use the transpiler to optimize them for the hardware we intend to use. To ensure that we have the same mapping between logical and physical qubits in every case, we inspect the transpiler results obtained for the circuits used to extract the bit-flip probabilities and to prepare the random wave function $|\psi\rangle$. For all the data reported in the main text, we checked that the mapping between logical and physical qubits is indeed the same, and the bit-flip probabilities we extract correspond to the qubits we use for generating our random wave functions.

[1] A. Montanaro, Quantum algorithms: An overview, npj Quantum Inf. **2**, 15023 (2016).

[2] F. G. S. L. Brandao and K. M. Svore, Quantum speed-ups for solving semidefinite programs, in *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Piscataway, NJ, 2017), pp. 415–426.

[3] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Quantum cryptography, Rev. Mod. Phys. **74**, 145 (2002).

[4] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. S. Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden, Advances in quantum cryptography, Adv. Opt. Photon. **12**, 1012 (2020).

[5] M. Schuld, I. Sinayskiy, and F. Petruccione, An introduction to quantum machine learning, Contemp. Phys. **56**, 172 (2015).

[6] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, Nature (London) **549**, 195 (2017).

[7] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[8] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[9] A. Peruzzo, J. McClean, P. Shadbolt, M. Yung, X. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[10] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, New J. Phys. **18**, 023023 (2016).

[11] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding *et al.*, Scalable Quantum Simulation of Molecular Energies, Phys. Rev. X **6**, 031007 (2016).

[12] A. Kandala, A. Mezzacapo, K. Temme *et al.*, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, Nature (London) **549**, 242 (2017).

[13] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure, Phys. Rev. A **95**, 020501(R) (2017).

[14] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi, Computation of Molecular Spectra on a Quantum Processor with an Error-Resilient Algorithm, Phys. Rev. X **8**, 011021 (2018).

[15] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, Cloud Quantum Computing of an Atomic Nucleus, Phys. Rev. Lett. **120**, 210501 (2018).

[16] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator, Phys. Rev. X **8**, 031022 (2018).

[17] M. Ganzhorn, D. J. Egger, P. Barkoutsos, P. Ollitrault, G. Salis, N. Moll, M. Roth, A. Fuhrer, P. Mueller, S. Woerner, I. Tavernelli, and S. Filipp, Gate-Efficient Simulation of Molecular Eigenstates on a Quantum Computer, Phys. Rev. Applied **11**, 044092 (2019).

[18] C. Kokail, C. Maier, R. van Bijnen *et al.*, Self-verifying variational quantum simulation of lattice models, Nature (London) **569**, 355 (2019).

[19] T. Hartung and K. Jansen, Zeta-regularized vacuum expectation values, J. Math. Phys. **60**, 093504 (2019).

[20] K. Jansen and T. Hartung, Zeta-regularized vacuum expectation values from quantum computing simulations, Proc. Sci. LATTICE2019 **363**, 153 (2020).

[21] Y. Li and S. C. Benjamin, Efficient Variational Quantum Simulator Incorporating Active Error Minimization, Phys. Rev. X **7**, 021050 (2017).

[22] K. Temme, S. Bravyi, and J. M. Gambetta, Error Mitigation for Short-Depth Quantum Circuits, Phys. Rev. Lett. **119**, 180509 (2017).

[23] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states, Phys. Rev. A **95**, 042308 (2017).

[24] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien, Low-cost error mitigation by symmetry verification, Phys. Rev. A **98**, 062339 (2018).

[25] S. Endo, S. C. Benjamin, and Y. Li, Practical Quantum Error Mitigation for Near-Future Applications, Phys. Rev. X **8**, 03102 (2018).

[26] S. McArdle, X. Yuan, and S. Benjamin, Error-Mitigated Digital Quantum Simulation, Phys. Rev. Lett. **122**, 180501 (2019).

[27] S. Endo, Q. Zhao, Y. Li, S. Benjamin, and X. Yuan, Mitigating algorithmic errors in a Hamiltonian simulation, Phys. Rev. A **99**, 012334 (2019).

[28] A. Kandala, K. Temme, A. D. Córcoles *et al.*, Error mitigation extends the computational reach of a noisy quantum processor, Nature (London) **567**, 491 (2019).

[29] J. R. McClean, Z. Jiang, N. C. Rubin, R. Babbush, and H. Neven, Decoding quantum errors with subspace expansions, Nat. Commun. **11**, 636 (2020).

[30] M. Otten and S. K. Gray, Recovering noise-free quantum observables, Phys. Rev. A **99**, 012338 (2019).

[31] M. Otten and S. K. Gray, Accounting for errors in quantum algorithms via individual error reduction, npj Quantum Inf. **5**, 11 (2019).

[32] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. C. Bultink, X. Fu, C. H. Price, V. P. Ostroukh, N. Muthusubramanian, A. Bruno, M. Beekman, N. Haider, T. E. O'Brien, and L. DiCarlo, Experimental error mitigation via symmetry verification in a variational quantum eigensolver, Phys. Rev. A **100**, 010302(R) (2019).

[33] M. Urbanek, B. Nachman, and W. A. de Jong, Quantum error detection improves accuracy of chemical calculations on a quantum computer, Phys. Rev. A **102**, 022427 (2020).

[34] O. Crawford, B. van Straaten, D. Wang, T. Parks, E. Campbell, and S. Brierley, Efficient quantum measurement of pauli operators in the presence of finite sampling error, Quantum **5**, 385 (2021).

[35] B. Chungheon, O. Tomohiro, T. Seigo, and C. Byung-Soo, Density matrix simulation of quantum error correction codes for near-term quantum devices, Quantum Sci. Technol. **5**, 015002 (2019).

[36] A. D. Córcoles, E. Magesan, S. J. Srinivasan, A. W. Cross, M. Steffen, J. M. Gambetta, and J. M. Chow, Demonstration of a quantum error detection code using a square lattice of four superconducting qubits, Nat. Commun. **6**, 6979 (2015).

[37] S. Sheldon, E. Magesan, J. M. Chow, and J. M. Gambetta, Procedure for systematically tuning up cross-talk in the cross-resonance gate, Phys. Rev. A **93**, 060302(R) (2016).

[38] S. S. Tannu and M. K. Qureshi, Mitigating measurement errors in quantum computers by exploiting state-dependent bias, in *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*, MICRO '52 (Association for Computing Machinery, New York, 2019), p. 279.

[39] K. Yeter-Aydeniz, E. F. Dumitrescu, A. J. McCaskey, R. S. Bennink, R. C. Pooser, and G. Siopsis, Scalar quantum field theories as a benchmark for near-term quantum computers, Phys. Rev. A **99**, 032306 (2019).

[40] K. Yeter-Aydeniz, R. C. Pooser, and G. Siopsis, Practical quantum computation of chemical and nuclear energy levels using quantum imaginary time evolution and lanczos algorithms, npj Quantum Inf. **6**, 63 (2020).

[41] E. van den Berg, Z. K. Minev, and K. Temme, Model-free readout-error mitigation for quantum expectation values, Phys. Rev. A **105**, 032620 (2022).

[42] H. Kaufmann, T. Ruster, C. T. Schmiegelow, M. A. Luda, V. Kaushal, J. Schulz, D. von Lindenfels, F. Schmidt-Kaler, and U. G. Poschinger, Fast ion swapping for quantum-information processing, Phys. Rev. A **95**, 052319 (2017).

[43] Qiskit Aer API documentation and source code, accessed on 2020/07/03.

[44] G. J. Mooney, G. A. L. White, C. D. Hill, and L. C. L. Hollenberg, Generation and verification of 27-qubit greenberger-horne-zeilinger states in a superconducting quantum computer, J. Phys. Commun. **5**, 095004 (2021).

[45] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000), p. 375.

[46] H. Abraham *et al.*, Qiskit: An open-source framework for quantum computing, Zenodo, (2019)10.5281/zenodo.2562110.

[47] *ibmq_london* v1.1.0, IBM Quantum team, Retrieved from https://quantum-computing.ibm.com (2020).

[48] *ibmq_burlington* v1.1.4, IBM Quantum team, Retrieved from https://quantum-computing.ibm.com (2020).

[49] M. R. Geller and M. Sun, Efficient correction of multiqubit measurement errors, Quantum Sci. Technol. **6**, 025009 (2021).

[50] M. R. Geller, Rigorous measurement error correction, Quantum Sci. Technol. **5**, 03LT01 (2020).

[51] M. C. Bañuls, K. Cichy, K. Jansen, and J. I. Cirac, The mass spectrum of the schwinger model with matrix product states, J. High Energy Phys. 11 (2013) 158.

[52] M. C. Bañuls, K. Cichy, J. I. Cirac, K. Jansen, and S. Kühn, Density Induced Phase Transitions in the Schwinger Model: A Study with Matrix Product States, Phys. Rev. Lett. **118**, 071601 (2017).

[53] M. C. Bañuls, K. Cichy, J. I. Cirac, K. Jansen, S. Kühn, and H. Saito, The multi-flavor schwinger model with chemical potential - overcoming the sign problem with matrix product states, PoS (LATTICE 2016) **256**, 316 (2016).

[54] L. Funcke, K. Jansen, and S. Kühn, Topological vacuum structure of the Schwinger model with matrix product states, Phys. Rev. D **101**, 054507 (2020).

[55] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, arXiv:1411.4028.

[56] P. D. Nation, H. Kang, N. Sundaresan, and J. M. Gambetta, Scalable mitigation of measurement errors on quantum computers, PRX Quantum **2**, 040326 (2021).

[57] E. Lieb, T. Schultz, and D. Mattis, Two soluble models of an antiferromagnetic chain, Ann. Phys. **16**, 407 (1961).

[58] P. Pfeuty, The one-dimensional Ising model with a transverse field, Ann. Phys. **57**, 79 (1970).

[59] A. Kitaev, Unpaired majorana fermions in quantum wires, Phys. Usp. **44**, 131 (2001).

[60] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Entanglement in Quantum Critical Phenomena, Phys. Rev. Lett. **90**, 227902 (2003).

[61] J. I. Latorre, E. Rico, and G. Vidal, Ground state entanglement in quantum spin chains, Quantum Inf. Comput. **4**, 48 (2004).