Note on Error Mitigation

Takahiro Yamamoto

July 9, 2019

1 Type of error

Qubit operations are susceptible to various types of errors due to imperfect control pulses, qubit-qubit couplings (crosstalk), and environmental noise. In order to improve qubit performance, it is necessary to identify the types and magnitudes of these errors and reduce them.

1. State preparation and measurement (SPAM)
   (a) intrinsic
   (b) extrinsic
2. Gate infidelity
   (a) 1 qubit operation
   (b) 2 qubit operation

It will be useful to classify SPAM errors into two different types, which we will call *intrinsic* and *extrinsic*. Intrinsic SPAM errors are those that are inherent in the state preparation and measurement process. One example is an error initializing the $|0\rangle$ state due to thermal populations of excited states. Another is dark counts when attempting to measure, say, the $|1\rangle$ state. Extrinsic SPAM errors are those due to errors in the gates used to transform the initial state to the starting state (or set of states) for the experiment to be performed.

Intrinsic SPAM errors are of particular relevance to fault-tolerant quantum computing, since it turns out that quantum error correction (QEC) requirements are much more stringent on gates than on SPAM.

1. Dephasing
2. Amplitude and phasing damping
3. Homogeneous depolarizing

1. Localized Markovian

2. Unbiased statistical fluctuation

Below are some concrete examples of quantum noise and quantum operations. They are also important in understanding the practical effects of noise on quantum systems, and how noise can be controlled by techniques such as error-correction.

Quantum operations can be represented in the operator-sum representation:

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger, \tag{1}$$

where the operators $\{E_k\}$ are known as operation elements.

1.1 Depolarizing

Imagine we take a single qubit, and with probability p that qubit is depolarized. That is, it is replaced by the completely mixed state, $I/2$. With probability $(1-p)$ the qubit is left untouched. The state of the quantum system after this noise is:

$$\mathcal{E}(\rho) = \frac{pI}{2} + (1-p)\rho \tag{2}$$

Using the relation

$$\frac{I}{2} = \frac{\rho + X\rho X + Y\rho Y + Z\rho Z}{4} \tag{3}$$

we obtain one-qubit depolarizing channel in the operator-sum representation,

$$\mathcal{E}(\rho) = \left(1 - \frac{3p}{4}\right)\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z) \tag{4}$$

In the same token, for two-qubit depolarizing channel with l, m-th qubits

$$\mathcal{E}(\rho) = \frac{pI}{4} + (1-p)\rho \tag{5}$$

and using the relation

$$\frac{I}{4} = \sum_{\{i,j\}=\{0,1,2,3\}} \frac{\sigma_l^j \sigma_m^i \rho \sigma_m^i \sigma_l^j}{16} \tag{6}$$

where $\sigma_l^i = \{I, X, Y, Z\}$. Thus it can be expressed in the operator-sum representation as

$$\mathcal{E}(\rho) = \left(1 - \frac{15p}{16}\right)\rho + \frac{p}{16}\sum_{i,j} \sigma_l^j \sigma_m^i \rho \sigma_m^i \sigma_l^j \tag{7}$$

where $\{i,j\} = \{0,1,2,3\}$ and $\{i,j\} \neq \{0,0\}$.

With $\xi = 3p/4$ for one-qubit and $\xi = 15p/16$ for two-qubit depolarizing channels acting on the system density matrix ρ, we achive

$$\mathcal{E}(\rho) = (1-\xi)\,\rho + \frac{\xi}{3}\sum_{i=1,2,3}\sigma^i\rho\sigma^i \tag{8}$$

$$\mathcal{E}(\rho) = (1-\xi)\,\rho + \frac{\xi}{15}\sum_{i,j}\sigma_l^j\sigma_m^i\rho\sigma_m^i\sigma_l^j \tag{9}$$

where $\{i,j\} = \{0,1,2,3\}$ and $\{i,j\} \neq \{0,0\}$.

1.2 Amplitude damping

the description of energy dissipation-effects due to loss of energy from a quantum system.

$$\mathcal{E}(\rho) = E_0\rho E_0^\dagger + E_1\rho E_1^\dagger \tag{10}$$

where

$$E_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \tag{11}$$

$$E_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}. \tag{12}$$

$\gamma = 1 - e^{-\tau/T_1}$ can be thought of as the probability of losing energy. The typical coherence values for the qubits of $T_1 = 30\mu$s and τ of 450 ns for an entangler are used in the simulation in [arXiv: 1704.05018].

The E_1 operation changes a $|1\rangle$ state into a $|0\rangle$ state, corresponding to the physical process of losing a quantum of energy to the environment. E_0 leaves $|0\rangle$ unchanged, but reduces the amplitude of a $|1\rangle$ state; physically, this happens because a quantum of energy was not lost to the environment, and thus the environment now perceives it to be more likely that the system is in the $|0\rangle$ state, rather than the $|1\rangle$ state. The strength of the channels is set by the experimental coherence times and the length of the gates.

$\mathcal{E}_{\mathrm{GAD}}$, called generalized amplitude damping, is defined for single qubits by the operation elements

$$E_0 = \sqrt{p}\begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \tag{13}$$

$$E_1 = \sqrt{p}\begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}, \tag{14}$$

$$E_2 = \sqrt{1-p}\begin{bmatrix} \sqrt{1-\gamma} & 0 \\ 0 & 1 \end{bmatrix}, \tag{15}$$

$$E_3 = \sqrt{1-p}\begin{bmatrix} 0 & 0 \\ \sqrt{\gamma} & 0 \end{bmatrix}. \tag{16}$$

where the stationary state ρ_∞, which statisfies $\mathcal{E}_{\mathrm{GAD}}(\rho_\infty) = \rho_\infty$ is,

$$\rho_\infty = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix}. \tag{17}$$

When γ is replaced with a time-varying function like $1 - e^{t/T_1}$, you can visualize the effects of amplitude damping as a flow on the Bloch sphere, which moves every point in the unit ball towards a fixed point at $|0\rangle$.

1.3 Phase damping

A noise process that is uniquely quantum mechanical, which describes the loss of quantum information without loss of energy, is phase damping. The energy eigenstates of a quantum system do not change as a function of time, but do accumulate a phase which is proportional to the eigenvalue. When a system evolves for an amount of time which is not precisely known, partial information about this quantum phase – the relative phases between the energy eigenstates – is lost. A phase kick, the angle of rotation θ is random. The randomness could originate, for example, from a deterministic interaction with an environment, which never again interacts with the system and thus is implicitly measured. Let us assume that the phase kick angle θ is well represented as a random variable which has a Gaussian distribution $e^{-\theta^2/4\lambda}$ with mean 0 and variance 2λ.

The random phase kicking causes the expected value of the off-diagonal elements of the density matrix to decay exponentially to zero with time, $e^{-\lambda}$. That is a characteristic result of phase damping.

$$\mathcal{E}(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger \tag{18}$$

where

$$E_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \tag{19}$$

$$E_1 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\gamma} \end{bmatrix}. \tag{20}$$

By applying the unitary freedom of quantum operations, we find that a unitary recombination of E_0 and E_1 gives a new set of operation elements for phase damping;

$$E_0' = \sqrt{\alpha} I \tag{21}$$

$$E_1' = \sqrt{1-\alpha} Z, \tag{22}$$

where $\alpha = (1 + \sqrt{1-\lambda})/2$. Thus the phase damping quantum operation is exactly the same as the phase flip channel.

Phase damping is often referred to as a T_2 relaxation process, for historical reasons, where dephasing time is related to γ as $\sqrt{1-\gamma} = e^{-t/2T_2} = e^{-t/T_\phi}$, where $T_\phi = 2T_2^* T_1/(2T_1 - T_2^*)$. The typical coherence values for the qubits of $T_2^* = 20\mu$s and τ of 450 ns for an entangler are used in the simulation in [arXiv: 1704.05018].

As a function of time, the amount of damping increases, corresponding to an inwards flow towards σ_z-axis.

1.4 Phase flip

$$\mathcal{E}(\rho) = p\rho + (1-p)Z\rho Z \tag{23}$$

1.5 Bit flip

The bit flip channel flips the state of a qubit from $|0\rangle$ to $|1\rangle$ (and vice versa) with probability $1-p$. It has operation elements

$$\mathcal{E}(\rho) = \sqrt{p}I\rho\sqrt{p}I + \sqrt{1-p}X\rho\sqrt{1-p}X \tag{24}$$

1.6 Bit-phase flip

$$\mathcal{E}(\rho) = \sqrt{p}I\rho\sqrt{p}I + \sqrt{1-p}X\rho\sqrt{1-p}Y \tag{25}$$

2 Unitary 2-design and twirling of quantum channel

3 Error models

The most general form of a single-qubit unitary:

$$U(\theta,\phi,\lambda) = \begin{bmatrix} \cos(\theta/2) & -e^{i\lambda}\sin(\theta/2) \\ e^{i\phi}\sin(\theta/2) & e^{i(\lambda+\phi)}\cos(\theta/2) \end{bmatrix}, \tag{26}$$

It is implemented using three frame changes and two $X_{\pi/2}$ pulses.

$$U_1(\lambda) = U(0,0,\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\lambda} \end{bmatrix}, \tag{27}$$

$$U_2(\phi,\lambda) = U(\pi/2,\phi,\lambda) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -e^{i\lambda} \\ e^{i\phi} & e^{i(\lambda+\phi)} \end{bmatrix}, \tag{28}$$

In the IBM Q Experience, this is implemented by a pre- and post-frame change and a $X_{\pi/2}$ pulse

3.1 Coherence

The decay time, (T_1) and dephasing time, (T_2)

$$f(t) = Ae^{-t/T_1} + C \tag{29}$$

for unknown parameters A, C, and T_1. If there are no SPAM errors, $A = 1$ and $C = 0$. Similarly, for T_2 and T_2^*, the ground state population is expected to behave like

$$f(t) = Ae^{-t/T_2^*} \cos(2\pi ft + \phi) + C \tag{30}$$

respectively; both with $A = C = 1/2$ in the lack of SPAM errors.

3.2 Hamiltonian Characterization

Measuring ZZ perform an experiment to measure ZZ between a pair of qubits. ZZ here is defined as the energy shift on the $|11\rangle$ state,

$$H = \frac{\omega_0}{2}(1 - \sigma_{Z,0}) + \frac{\omega_1}{2}(1 - \sigma_{Z,1}) + \xi \, |11\rangle \langle 11| \tag{31}$$

The experiment to measure ξ is to perform a Ramsey experiment on Q0 (H-t-H) and repeat the Ramsey with Q1 in the excited state. The difference in frequency between these experiments is the rate ξ. ZZ rates are typically ≈100kHz so we want Ramsey oscillations around 1MHz. 12 numbers ranging from 10 to 1000, logarithmically spaced

3.3 Amplitude Error Characterization for Single Qubit Gates

Measure the amplitude error in the single qubit gates. Here this measures the error in the $\pi/2$ pulse. Note that we can run multiple amplitude calibrations in parallel. This shows the sequence of the calibration, which is repeated application of $Y_\pi = e^{i(\pi/2)Y} = U_2(0, 0)$. Note that the measurements are mapped to a minimal number of classical registers in order of the qubit list.

Suppose error model where each Y_π gate has error

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}. \tag{32}$$

Excited state population can be fit as:

$$C - \frac{1}{2} \cos\left[\left(\theta + \frac{\pi}{2}\right)(x + 1)\right] \tag{33}$$

where x is the number of gate repetitions and θ is the error for the pulse (amplitude/error). [TODO: fill the gap]

3.4 Angle Error Characterization for Single Qubit Gates

Measure the angle between the X and Y gates

Gate sequence for measuring the angle error $U_2(0,0)U_1(2\theta)U_2(-\pi/2,\pi/2)U_2(-\pi/2,\pi/2)\cdots$ where the U_1 gates are added errors to test the procedure

3.5 Amplitude Error Characterization for CNOT Gates

This looks for a rotation error in the CNOT gate, ie., if the gate is actually $CR_X(\pi/2+\delta)$ measure δ. This is very similar to the single qubit amplitude error calibration except we need to specify a control qubit (which is set to be in state $|1\rangle$) and the rotation is a π.

Suppose error model where each CNOT gate has error

$$\begin{bmatrix} I & O \\ O & e^{i\theta X} \end{bmatrix}. \tag{34}$$

$$C + \frac{1}{2}\sin\left[(\theta+\pi)\,x\right] \tag{35}$$

where x is the number of gate repetitions and θ is the amplitude error for the pulse. See: [https://qiskit.org/documentation/ignis/characterization.html]

3.6 Angle Error Characterization for CX Gates

Measure the angle error θ in the CNOT gate, i.e., $CR_{\cos(\theta)X+\sin(\theta)Y}(\pi/2)$ with respect to the angle of the single qubit gates.

3.7 Measurement Error and Mitigation

The last step of a typical quantum experiment is to perform a measurement on the qubits in the circuit. Although the qubit state $|\psi\rangle$ (or more generally the density matrix ρ) is the general description of the quantum state, in a typical strong projective measurement our measurement projects the general state into a specific computational state $|x\rangle$ (where x is a bitstring, e.g., 1001010) The probability of measuring bitstring x is given by:

$$P_x = \mathrm{Tr}(\langle x|\,\rho\,|x\rangle) \tag{36}$$

Therefore, the measurement process is stochastic. The above distribution of x given a state ρ is true only in the absence of measurement errors. There are multiple sources of possible measurement error, all of which are dependent on the physical mechanism of measurement in the system. For superconducting qubits coupled to readout cavities [1,2,3,4,5] the state of the qubit is determined by measurement the response of a microwave tone incident on the readout cavity. The cavity signal is measured for some time where $V(t)$ is the

complex amplitude of the signal which is converted to a single complex number based on a measurement kernel

$$V = \int_0^T V(t)K(t)dt \tag{37}$$

which is then turned into a bit by a nonlinear discriminator [6]. The simplest example being if $|V| < V_0$ then the qubit was in state 0 and otherwise the qubit was in state 1.

As discussed in [6] there are classical sources of noise on the signal that lead to misidentification of the qubit state, but it can also happen that the qubit decays due to T_1 during the measurement. There are other sources of crosstalk (to numerous to enumerate) such as classical crosstalk on the lines and crosstalk between resonantors on chip. All of these issues lead to a new probability distribution $\tilde{P}_\rho$ for a given state. Given certain assumptions about these errors and appropriate calibration we can attempt to correct the skew in the probability distribution on average.

References [1] Alexandre Blais, Ren-Shou Huang, Andreas Wallraff, S. M. Girvin, and R. J. Schoelkopf, Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation, https://arxiv.org/abs/cond-mat/0402216

[2] Jay Gambetta, Alexandre Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Qubit-photon interactions in a cavity: Measurement induced dephasing and number splitting https://arxiv.org/abs/cond-mat/0602322

[3] Alexandre Blais, Jay Gambetta, A. Wallraff, D. I. Schuster, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf. Quantum information processing with circuit quantum electrodynamics. https://arxiv.org/abs/cond-mat/0612038

[4] Jay Gambetta, W. A. Braff, A. Wallraff, S. M. Girvin, R. J. Schoelkopf. Protocols for optimal readout of qubits using a continuous quantum nondemolition measurement. https://arxiv.org/abs/cond-mat/0701078

[5] Jay Gambetta, Alexandre Blais, M. Boissonneault, A. A. Houck, D. I. Schuster and S. M. Girvin. Quantum trajectory approach to circuit QED: Quantum jumps and the Zeno effect. https://arxiv.org/abs/0709.4264

[6] Colm A. Ryan, Blake R. Johnson, Jay M. Gambetta, Jerry M. Chow, Marcus P. da Silva, Oliver E. Dial and Thomas A. Ohki. Tomography via Correlation of Noisy Measurement Records. https://arxiv.org/abs/1310.6448

3.8 Constructing a Full Calibration Matrix

The assumption of the error mitigation technique is that we can prepare each of the basis states with very low error. Given this assumption, in separate experiments we can prepare one of the 2^n states and then measure the outputs in all 2^n states. Normalizing these outputs and making each set of output probabilities for a given prepared state the columns of a matrix we obtain the matrix $\mathbf{A}$ which translates the ideal probability distribution of

the state ρ (P_ρ) into the experimental probability distribution $\tilde{P}_\rho$

$$\tilde{P}_\rho = \mathbf{A} \cdot P_\rho \tag{38}$$

3.9 Error models

$$\Lambda(\rho) = \sum_k E_k \rho E_k^\dagger \tag{39}$$

$$E_1 = \frac{1+\sqrt{1-\gamma-\lambda}}{2} I + \frac{1-\sqrt{1-\gamma-\lambda}}{2} Z \tag{40}$$

$$E_2 = \frac{\sqrt{\gamma}}{2}(X + iY) \tag{41}$$

$$E_1 = \frac{\sqrt{\lambda}}{2}(I - Z) \tag{42}$$

$$\Lambda(\rho) \to \sum_{A \in \mathcal{A}} A^\dagger \Lambda(A\rho A^\dagger) A = \sum_{A \in \mathcal{A}} p_A A^\dagger \rho A \tag{43}$$

$$\mathcal{A} = \{I, X, Y, Z\} \tag{44}$$

4 Qubit characterization methods

Several methods of qubit characterization are currently available[1]. In chronological order of their development, the main techniques are:

1. quantum state tomography (QST)
2. quantum process tomography (QPT)
3. randomized benchmarking (RB)
4. quantum gate set tomography (GST)

4.1 Quantum state tomography (QST)

4.2 Quantum process tomography (QPT)

A quantum operation on a d-dimensional quantum system can be completely determined by experimentally measuring the output density matrices produced from d^2 pure state inputs.

[1] Introduction to Quantum Gate Set Tomography, D. Greenbaum, arXiv:1509.02921

4.3 Randomized benchmarking (RB)

4.4 Quantum Volume

Quantum Volume (QV) is a single-number metric that can be measured using a concrete protocol on near-term quantum computers of modest size. The QV method quantifies the largest random circuit of equal width and depth that the computer successfully implements. Quantum computing systems with high-fidelity operations, high connectivity, large calibrated gate sets, and circuit rewriting toolchains are expected to have higher quantum volumes.

The Quantum Volume Protocol A QV protocol (see [1]) consists of the following steps:

4.4.1 Step 1: Generate QV sequences

It is well-known that quantum algorithms can be expressed as polynomial-sized quantum circuits built from two-qubit unitary gates. Therefore, a model circuit consists of d layers of random permutations of the qubit labels, followed by random two-qubit gates (from $SU(4)$). When the circuit width m is odd, one of the qubits is idle in each layer.

More precisely, a QV circuit with depth d and width m, is a sequence $U = U^{(d)} \cdots U^{(2)} U^{(1)}$ of d layers:

$$U^{(t)} = U^{(t)}_{\pi_t(m'-1),\pi_t(m)} \otimes \cdots \otimes U^{(t)}_{\pi_t(1),\pi_t(2)} \tag{45}$$

each labeled by times $t = 1...d$ and acting on $m' = 2\lfloor n/2 \rfloor$ qubits. Each layer is specified by choosing a uniformly random permutation $\pi_t \in \mathcal{S}_m$ of the m qubit indices and sampling each $U^{(t)}_{a,b}$, acting on qubits a and b, from the Haar measure on $SU(4)$.

4.4.2 Step 2: Simulate the ideal QV circuits

The quantum volume method requires that we know the ideal output for each circuit, so we use the statevector simulator in Aer to get the ideal result.

4.4.3 Step 3: Calculate the heavy outputs

To define when a model circuit U has been successfully implemented in practice, we use the heavy output generation problem. The ideal output distribution is $p_U(x) = |\langle x| U |0\rangle|^2$, where $x \in \{0,1\}^m$ is an observable bit-string.

Consider the set of output probabilities given by the range of $p_U(x)$ sorted in ascending order $p_0 \leq p_1 \leq \cdots \leq p_{2^m-1}$. The median of the set of probabilities is $p_{\text{med}} = (p_{2^{m-1}} + p_{2^{m-1}-1})/2$, and the heavy outputs are

$$H_U = \{x \in \{0,1\}^m \mid p_U(x) > p_{\text{med}}\}\,. \tag{46}$$

The heavy output generation problem is to produce a set of output strings such that more than two-thirds are heavy.

4.4.4 Step 4: Define the noise model

We define a noise model for the simulator. To simulate decay, we add depolarizing error probabilities to the CNOT and U gates.

4.4.5 Step 5: Calculate the average gate fidelity

The average gate fidelity between the m-qubit ideal unitaries U and the executed U' is:

$$F_{\text{avg}}(U, U') = \frac{|\text{Tr}(U^\dagger U')|^2/2^m + 1}{2^m + 1} \tag{47}$$

The observed distribution for an implementation U' of model circuit U is $q_U(x)$, and the probability of sampling a heavy output is:

$$h_U = \sum_{x \in H_U} q_U(x)$$

4.4.6 Step 6: Calculate the achievable depth

The probability of observing a heavy output by implementing a randomly selected depth d model circuit is:

$$h_d = \int_U h_U dU \tag{48}$$

The achievable depth $d(m)$ is the largest d such that we are confident that $h_d > 2/3$. In other words,

$$h_1, h_2, \cdots, h_{d(m)} > \frac{2}{3} \text{ and } h_{d(m+1)} \leq \frac{2}{3} \tag{49}$$

We now convert the heavy outputs in the different trials and calculate the mean h_d and the error for plotting the graph.

4.4.7 Step 7: Calculate the Quantum Volume

The quantum volume treats the width and depth of a model circuit with equal importance and measures the largest square-shaped (i.e., $m = d$) model circuit a quantum computer can implement successfully on average.

The quantum volume V_Q is defined as

$$\log_2 V_Q = \arg\max_m \min(m, d(m)) \tag{50}$$

We list the statistics for each depth. For each depth we list if the depth was successful or not and with what confidence interval. For a depth to be successful the confidence interval must be $> 97.5\%$.

References [1] Andrew W. Cross, Lev S. Bishop, Sarah Sheldon, Paul D. Nation, and Jay M. Gambetta, Validating quantum computers using randomized model circuits, https://arxiv.org/pdf/1811.12926 [2] S. Aaronson, L. Chen, Complexity-Theoretic Foundations of Quantum Supremacy Experiments, https://arxiv.org/pdf/1612.05903 [3] Quantum Supremacy and the Complexity of Random Circuit Sampling, https://arxiv.org/pdf/1803.04402

4.5 Quantum gate set tomography (GST)

GST arose from the observation that QPT is inaccurate in the presence of SPAM errors. In QPT, the starting states must form an informationally complete basis of the Hilbert-Schmidt space on which the gate being estimated acts. These are typically created by applying gates to a given initial state, usually the $|0\rangle$ state, and these gates themselves may be faulty.

According to recent results from IBM, a 50-fold increase in intrinsic SPAM error reduces the surface code threshold by only a factor of 3-4. Therefore QPT – the accuracy of which degrades with increasing SPAM – would not be able to determine if a qubit meets threshold requirements when the ratio of intrinsic SPAM to gate error is large.

This is not an issue for extrinsic SPAM errors, which go to zero as the errors on the gates go to zero. Nevertheless, extrinisic SPAM error interferes with diagnostics: as an example, QPT cannot distinguish an over-rotation error on a single gate from the same error on all gates. In addition, Merkel, et al. have found that, for a broad range of gate error – including the thresholds of leading QEC code candidates – the ratio of QPT estimation error to gate error increases as the gate error itself decreases. This makes QPT less reliable as gate quality improves.

Extrinsic SPAM error is also unsatisfactory from a theoretical point of view: QPT assumes the ability to perfectly prepare a complete set of states and measurements. In reality, these states and measurements are prepared using the same faulty gates that QPT attempts to characterize. One would like to have a characterization technique that takes account of SPAM gates self-consistently. We shall see that GST is able to resolve all of these issues.

Another approach to dealing with SPAM errors is provided by randomized benchmarking. RB is based on the idea of twirling – the gate being characterized is averaged in a such a way that the resulting process is depolarizing with the same average fidelity as the original gate. The depolarizing parameter of the averaged process is measured experimentally, and the result is related back to the average fidelity of the original gate. This technique is independent of the particular starting state of the experiment, and therefore is not affected by SPAM errors. However, RB has several shortcomings which make it unsatisfactory as a sole characterization technique for fault-tolerant QIP. For one thing, it is limited to Clifford gates, and so cannot be used to characterize a universal gate set for quantum computing. For another, RB provides only a single metric of gate quality, the average fidelity. This can be insufficient for determining the correct qubit error model to

use for evaluating compatibility with QEC. Several groups have shown that qualitatively different errors can produce the same average gate fidelity, and in the case of coherent errors the depolarizing channel inferred from the RB gate fidelity underestimates the effect of the error. Finally, RB assumes the errors on subsequent gates are independent. This assumption fails in the presence of non-Markovian, or time-dependent noise. GST suffers from this assumption as well, but the long sequences used in RB make this a more pressing issue.

Despite these apparent shortcomings, RB has been used with great success by several groups to measure gate fidelities and to diagnose and correct errors. RB also has the advantage of scalability – the resources required to implement RB (number of experiments, processing time) scale polynomially with the number of qubits being characterized. QPT and GST, on the other hand, scale exponentially with the number of qubits. As a result, these techniques will foreseeably be limited to addressing no more than 2-3 qubits at a time.

GST and RB may end up complementing each other as elements of a larger characterization protocol for any future multi-qubit quantum computer.

TODO: summarize them in table.

method	assumption	advantage	disadvantage	scalability
quantum state tomography				
quantum process tomography				
randomized benchmarking				
quantum gate set tomography				

5 Type of error mitigation technics

There are four types of errors that can result in infidelity in the variational quantum simulation:

1. errors due to limited generality of the trial wave function, which may only be able to describe the simulated system approximately;

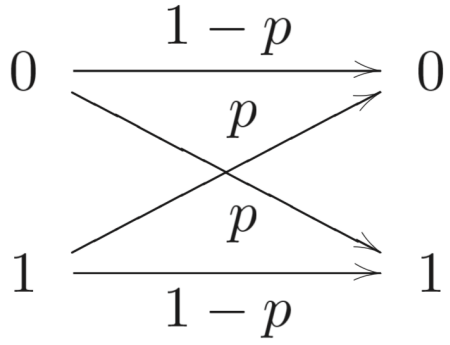
2. shot noise (TODO: definition) in measurments

3. errors due to noise in the quantum machine, e.g., decoherence and quantum gate infidelity.

Type of error mitigation technics are listed as:

1. Error extrapolation

2. Quasiprobability decomposition



3. Quantum subspace expansion. quantum channels
4. Process tomography protocols
   (a) Gate set tomography

5.1 Homogeneous and inhomogeneous scaling

5.1.1 Theory

Consider the case in which the effect of machine noise is to depolarize the ancillary qubit at a fixed level; *i.e.*, the output of the quantum computer becomes $\langle X\rangle = \eta \langle X\rangle^{(0)}$, where η is a constant independent of the quantum circuit.

An example of the homogeneous scaling is the case of balanced measurement errors. Errors in the measurement on the ancillary qubit (see Fig.?) can be modeled as follows:

1. If the state of the qubit is $|0\rangle$ the measurement outcome is correct with the probability $(1 - p_0)$ and the outcome is incorrect with the probability p_0
2. If the state of the qubit is $|1\rangle$ the measurement outcome is correct with the probability $(1 - p_1)$ and the outcome is incorrect with the probability p_1

Then the expectation value is $\langle X\rangle = (p_1 - p_0) + (1 - p_0 - p_1)\langle X\rangle^{(0)}$,

If measurement errors are balanced, *i.e.*, $p_0 = p_1$, the effect of measurement errors is a fixed scaling factor $\eta = 1 - p_0 - p_1$, which does not result in computing errors. Therefore, the hybrid algorithm is inherently insensitive to measurement errors on the ancillary qubit if these errors are balanced. Note also that if single-qubit gates are reliable, one can flip the qubit before the measurement so that measurement errors are effectively balanced. Measurement errors can be corrected even if they are not balanced. If p_0 and p_1 can be evaluated by benchmarking measurement operations, one can easily work out the true value $\langle X\rangle^{(0)}$ using the value obtained from the real machine: $\langle X\rangle^{(0)} = [\langle X\rangle - (p_1 - p_0)]/(1 - p_0 - p_1)$ We note that when error probabilities are higher, the denominator is smaller, which means

that we need to evaluate [$\langle X\rangle$ with a higher accuracy in order to achieve the same accuracy of $\langle X\rangle^{(0)}$[2]

5.1.2 Experiment

5.1.3 Summary

5.2 Error extrapolation

5.2.1 Motivation

5.2.2 Theory

Errors in an operation are stochastic if the operation is described by a superoperator $\mathcal{NU}$ and $\mathcal{N}$ has the form $\mathcal{N} = (1-\epsilon)\mathcal{I} + \mathcal{E}$. Here, $\mathcal{U}$ is the ideal operation without errors, $\mathcal{N}$ is the superoperator describing the effect of the noise, $\mathcal{I}$ is an identity operation, and errors $\mathcal{E}$ occur with the probability ϵ. Here, $\mathcal{E}$ is a valid quantum operation, *i.e.*, trace-preserving completely positive map.

Given an initial state ρ, after a sequence of operations, the final state of the quantum computer is $\rho = \mathcal{N}_L\mathcal{U}_L \cdots \mathcal{N}_\ell\mathcal{U}_\ell \cdots \mathcal{N}_1\mathcal{U}_1\rho$

where $\mathcal{N}_\ell\mathcal{U}_\ell$ denotes the ℓ-th operation. Taking into account the fact that errors are stochastic, the outcome can be rewritten in the form

$$\langle X\rangle = \left(1 - r\sum \epsilon_\ell\right)\langle X\rangle^{(0)} + r\langle X\rangle^{(1)} + \mathcal{O}(r^2), \tag{51}$$

where r is a convenient scale factor. For more detailed discussion see [PhysRevX.7.021050]. If probabilities of errors are tunable, we can infer the value of $\langle X\rangle^{(0)}$ by measuring values of $\langle X\rangle$ of a set of different factors r [see Fig.?].

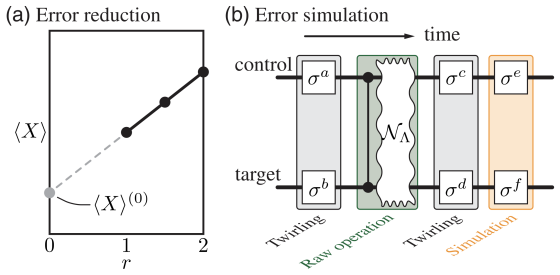
To infer the value of $\langle X\rangle^{(0)}$, first, we take N_X values of r and measure $\langle X\rangle$ with $r_1, \cdots, r_X$, where we can take $r_1 = 1$. Then, we fit the output as $\langle X\rangle(r) = \langle X\rangle^{(0)} + \chi r$ and obtain the value of $\langle X\rangle^{(0)}$. In this way, the first-order contribution of machine noise can be corrected. Similarly, by considering second-order terms in the expansion Eq.?, we can fit data using a function with second-order terms.

Using the extrapolation, we can reduce the effect of the machine noise. However, the final estimation of $\langle X\rangle^{(0)}$ may still be different from its actual value, and the error in the extrapolation depends on the shot noise in estimating each $\langle X\rangle(r)$.

The error-reduction protocol only works for small-size circuits, which are used in the hybrid algorithm, while the Trotterization algorithm usually needs large-size circuits. The true value $\langle X\rangle^{(0)}$ can be inferred when the contribution of high-order terms is much smaller than the contribution of lower-order terms.

The total rate of errors in the circuit with N_g gates is approximately $1-(1-\epsilon)^{N_g} = N_g\epsilon + (N_g\epsilon)^2/2 + \cdots$.

[2] PhysRevX.7.021050



When there are too many gates in the circuit or the error rate is too high, $N_g\epsilon \gtrsim 1$, the quantum state will be populated with errors, and one cannot retrieve the true value $\langle X \rangle^{(0)}$ even if we consider high-order terms in the interpolation.

5.2.3 Experiment

5.2.4 Summary

5.3 Error twirling

TODO: fill in

5.4 Quasiprobability decomposition

5.4.1 Motivation

5.4.2 Theory

Utility of "twirling" operations in minimizing the cost[3]. For the extrapolation method, their optimisation is to observe that typically for the classes of noise most common in experiments it is appropriate to assume that the expected values of the observation will decay exponentially with he severity of the circuit noise, rather than polynomial.

5.4.3 Experiment

5.4.4 Summary

5.5 Process tomography protocols

Localized Markovian errors

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

5.5.1 Motivation

5.5.2 Theory

Single-qubit Clifford gates and measurements are universal in computing expectation values. Any quantum operation is a linear map. Single qubit Clifford gates and measurements yield a complete set of linear independent maps. Any error can be simulated or subtracted by decomposition of the error using complete operation set.

By combining GST and the complete set decomposition, any localized Markovian errors in the QC can be systematically mitigated, so that the error in the final computational output is due to unbiased statistical fluctuation.

5.5.3 Experiment

5.5.4 Summary

6 Summary

1. Error mitigation method
2. Applicable error type
3. Efficiency
4. Cost (per qubit)

A Topological QC

Fault-tolerant quantum computing can be achieved by encoding qubits in non-Abelian anyons in topological materials[4]

B Quantum error correction

However, quantum error correction involves a substantial multiplication of resources; the number of physical qubits required may be orders of magnitude greater than the number of error-free logical qubits seen by the algorithm. A recent study audited the cost of implementing Shor's algorithm to solve a classically infeasible task and found that, even with state-of-the-art techniques for magic-state distillation, the machine would need over six million of today's highest- quality qubits[5].

[4] A. Kitaev, Fault-Tolerant Quantum Computation by Anyons, Ann. Phys. (Amsterdam) 303, 2 (2003).

[5] J. O'Gorman and E. T. Campbell, Quantum Computation with Realistic Magic State Factories, Phys. Rev. A 95, 032338 (2017)

C Dynamics simulation of quantum system

Dynamics must be studied when properties cannot be determined from static features. This has motivated dynamical versions of many well-known techniques, e.g., nonequilibrium dynamical mean-field theory[6], the time-dependent variational quantum Monte Carlo method[7], time-dependent tensor network methods[8], and, of course, time-dependent density functional theory[9].

A new approach[10] is based on the variational method, and our hope is that it could be implemented using small- size quantum circuits, i.e., quantum circuits with a small number of quantum operations that suffer significant noise compared with fault-tolerant quantum computers.

Variational methods have numerous applications in the numerical study of many-body quantum systems: for example, density functional theory[11], the matrix product state method[12], and simulating molecular dynamics using the variational principle[13]

[6]H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner, Nonequilibrium Dynamical Mean-Field Theory and Its Applications, Rev. Mod. Phys. 86, 779 (2014).

[7]G. Carleo, F. Becca, M. Schiro, and M. Fabrizio, Localization and Glassy Dynamics of Many-Body Quantum Systems, Sci. Rep. 2, 243 (2012).

[8]A. J. Daley, C. Kollath, U. Schollwock, and G. Vidal, Time-Dependent Density-Matrix Renormalization-Group Using Adaptive Effective Hilbert Spaces, J. Stat. Mech. (2004) P04005., M. C. Banuls, M. B. Hastings, F. Verstraete, and J. I. Cirac, Matrix Product States for Dynamical Simulation of Infinite Chains, Phys. Rev. Lett. 102, 240603 (2009).

[9]E. Runge and E. K. U. Gross, Density-Functional Theory for Time-Dependent Systems, Phys. Rev. Lett. 52, 997 (1984).

[10]Ying Li and Simon C. Benjamin, Efficient Variational Quantum Simulator Incorporating Active Error Minimization, Physical Review X 7, 021050 (2017)

[11]R. O. Jones, Density Functional Theory: Its Origins, Rise to Prominence, and Future, Rev. Mod. Phys. 87, 897 (2015).

[12]D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Matrix Product State Representations, Quantum Inf. Comput. 7, 401 (2007).

[13]H. Feldmeier and J. Schnack, Molecular Dynamics for Fermions, Rev. Mod. Phys. 72, 655 (2000).