Self-Consistent Quantum Process Tomography

Seth T. Merkel,1 Jay M. Gambetta,1 John A. Smolin,1 S. Poletto,1 A. D. Corcoles,1 B. R. Johnson,2 Colm A. Ryan,2 and M. Steffen1
1IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA
2Raytheon BBN Technologies, Cambridge, MA 02138, USA
(Dated: May 22, 2014)

Quantum process tomography is a necessary tool for verifying quantum gates and diagnosing faults in architectures and gate design. We show that the standard approach of process tomography is grossly inaccurate in the case where the states and measurement operators used to interrogate the system are generated by gates that have some systematic error, a situation all but unavoidable in any practical setting. These errors in tomography can not be fully corrected through oversampling or by performing a larger set of experiments. We present an alternative method for tomography to reconstruct an entire library of gates in a self-consistent manner. The essential ingredient is to define a likelihood function that assumes nothing about the gates used for preparation and measurement. In order to make the resulting optimization tractable we linearize about the target, a reasonable approximation when benchmarking a quantum computer as opposed to probing a black-box function.

INTRODUCTION

The design and control of scalable quantum architectures at the level of precision necessary for fault tolerant computation remains a considerable challenge. There has been remarkable progress in numerous physical systems (e.g. superconducting circuits [1–3], atomic systems [4, 5], etc.) to improve metrics such as the gate fidelity [6], and incrementally these values are approaching the fault-tolerant threshold. Diagnosing errors and improving designs of gates and architectures however, often requires much more information than a single scalar value. Instead we need a full characterization of a quantum process, and that necessitates quantum process tomography (QPT) [7].

The essential idea of QPT is as follows: prepare an initial quantum state, apply the operation we would like to reconstruct, measure the expectation value of some observable until it is possible to retrodict the process and then repeat for different initial states and observables. The resulting process may not be physical (i.e. a completely positive trace preserving map). When this non-physicality arises from Gaussian noise on the measurement outcomes it can be corrected using maximum likelihood estimation or through Bayesian inference techniques [9–16].

In addition to stochastic measurement noise, quantum process tomography should be consistent with respect to systematic noise. Most quantum computing architectures allow only for fixed initial states and measurement operators so that the state preparation and measurement (SPAM) described for tomography involves the application of gates that may have the same degree of error as the process being interrogated. These errors in QPT can not be corrected by gathering more measurement statistics or enforcing physicality alone. Here we show that with standard maximum likelihood techniques these errors, and in particular coherent errors, lead to fidelity estimates from QPT that can be very poor. Even more problematic is that the ratio of the QPT to SPAM errors actually can grow as the SPAM error decreases implying that as the gate fidelities approach the threshold it can be even harder to estimate them correctly. SPAM errors can be accounted for in fidelity estimation by using randomized benchmarking [5, 17–20] but this does not provide full tomographic information about the gates in order to diagnose and correct errors. We observe a disagreement between benchmarking [20] and tomography [1] experiments on similar samples which suggests that SPAM errors are a current limitation that needs to be addressed.

In this manuscript we explore techniques to compensate for SPAM errors while still obtaining a full characterization of our quantum gates. One simple technique is to oversample, not by increasing the repetitions of a given experiment, but by increasing the number of gates used for SPAM. We show that this technique shows no improvement for general unitary errors, however it does yield an improvement for a uniform error such as a global frame transformation. The primary subject of this work is an alternative approach to tomography that places the SPAM gates on the same footing as the gate under investigation, and then retrodicts an entire library of unitary gates in a self-consistent manner. The ‘cost’ of this method in terms of the number of independent measurements and post-processing is polynomially equivalent to performing standard QPT on each gate in the
library being reconstructed. Our approach assumes very little about the gate error model, in contrast to the self-correcting state tomography proposed in [21] or the bootstrap method in [22] which otherwise compensate for the same SPAM issues.

The remainder of the manuscript proceeds as follows. In Section 4 we give a detailed description of QPT and simulate the effects of SPAM errors for a range of error models and system parameters. In Section 5 we show the effects of oversampling using unitary 2-designs and in Section 6 we develop self-consistent gate-set tomography. Finally, in Section 7 we apply our self-consistent tomography to experimental data obtained from two independent experiments on a system consisting of a single superconducting qubit.

THE TROUBLE WITH TOMOGRAPHY

A general quantum process on a finite $d$-dimensional Hilbert space $\mathcal{H}_d$ ($d = 2^n$ for qubit systems) is a completely positive trace-preserving (CPTP) map. There are many ways to represent such a map, $\Lambda$, but in this manuscript we will primarily use two of them: the Choi matrix [23] which is given by

$$\rho = \frac{1}{d} \sum_{ij} E_{ij} \otimes \Lambda(E_{ij}),$$

where $E_{ij} = |i\rangle \langle j|$, and the Pauli transfer matrix (PTM) [1], which is

$$(R_\Lambda)_{ij} = \frac{1}{d} \operatorname{Tr}(P_i \Lambda(P_j)),
$$

where $P_j$ denotes the set of Pauli matrices, $\{I, X, Y, Z\}^\otimes n$, though one could extend this to qudit systems by considering any Hermitian basis for the $P_j$'s. The transformation between these two representations is the linear mapping,

$$(R_\Lambda)_{ij} = \operatorname{Tr}(\rho A_j P_i^T \otimes P_k),$$

$$\rho = \frac{1}{d^2} \sum_{ij} (R_\Lambda)_{ij} P_j^T \otimes P_k,$$

and each representation have useful properties. The Choi matrix is positive semidefinite if and only if $\Lambda$ is completely positive. If we define $P_0 = I_d$, then the map is trace-preserving if and only if the first row of $R_\Lambda$ is the vector $(1, 0, 0, \ldots)$ and is unital if the first column has the same structure. Furthermore, in the PTM picture composition becomes matrix multiplication, $R_{\Lambda_1 \circ \Lambda_2} = R_{\Lambda_1} R_{\Lambda_2}$, since the PTM is the standard superoperator group representation defined over the Pauli basis.

Let us define an experiment by the triple $\{\rho_j, M_j, m_j\}$ which describes preparing the system in state $\rho_j$ measuring the operator $M_j$ and obtaining an expectation value $m_j$ according to

$$m_j = \operatorname{Tr}(M_j \Lambda(\rho_j)).$$

We can define a vector $|\rho\rangle$ whose elements are $\langle j | \rho \rangle = \operatorname{Tr}(P_j \rho)/d$ as well as $\langle M_j | \rho \rangle = \operatorname{Tr}(M_j P_j)$ (the omission of the dimensional factor in the measurement is intentional and is a consequence of bounding the values of the $R$ matrix between $\pm 1$). In this form,

$$m_j = \langle M_j | R_\Lambda | \rho \rangle = \operatorname{Tr}(|M_j\rangle \langle \rho_j| R_\Lambda).$$

This final expression is simply another inner product of the matrix $R_\Lambda$ with the super-operator $M_j$ and $s_j$ as the vectorized $|M_j\rangle \langle \rho_j|$ so that

$$m_j = s_j^T r_\Lambda.$$  

We can now express the entire set of experiments as

$$m = S^T r_\Lambda,$$

where the columns of the rectangular matrix $S$ are the vectors $s_j$. The first estimate for our quantum process, $r_{\text{bare}}$, arises from linear inversion. While $S$ is typically not square we can invert $SS^T$ or, in the case of an incomplete set of measurements, find its pseudo inverse. This leads to the bare estimate

$$r_{\text{bare}} = (SS^T)^{-1} S m.$$  

The astute reader may have noticed that at no point have we enforced that our estimate $r_{\text{bare}}$ correspond to a physical map, and we will find that for many error models it does not.

The dominant source of error, in the tomography literature at least [30], is statistical error due to finite measurement statistics. Instead of direct access to $m_j$ we measure

$$m_j = s_j^T r_\Lambda + \sqrt{N_j W_j},$$

where $W_j$ is a Gaussian random variable with mean zero and unit variance and $N_j$ is the noise power. Due to the central limit theorem we can approximate the error model as Gaussian if we repeat the measurement sufficiently many times in which case $N_j$ scales as $1/(\# \text{ of repetitions of experiment } j)$. The transformation $m_j' = m_j/\sqrt{N_j}$ and $s_j' = s_j/\sqrt{N_j}$ simplifies the Gaussian likelihood function to

$$\mathcal{L}(m'|r) = \exp \left(-\frac{1}{2} \sum_j |m_j' - s_j^T r|^2 \right).$$

For this error model the most rigorous estimate of the gate would be to perform some sort of Bayesian estimation where $r_{\text{Bayes}} = \int r \mathcal{L}(m'|r) dr / \int \mathcal{L}(m'|r) dr$ but these
integrals are often extremely hard to calculate, especially since there is no uniquely good measure over quantum channels $d\mathbf{r}$. A much simpler technique is to report the maximum of the likelihood function. If the global maximum lies inside the space of physical maps then it is simply the $\mathbf{r}_{\text{bare}}$ from Eq. (8), and if most of the support of $\mathcal{L}(\mathbf{m}'|\mathbf{r})$ is also physical then the Bayesian estimate and the maximum likelihood estimate will be approximately the same. In the case where the bare estimate is unphysical maximizing $\mathcal{L}(\mathbf{m}'|\mathbf{r})$ over the space of physical $\mathbf{r}$ is a semidefinite program \cite{24} and can be solved for small instances easily with optimizers such as SeDuMi \cite{24}. This is because the likelihood function is of the form of a 2-norm distance between vectors $\mathbf{m}'$ and $S^{T} \mathbf{r}$ and the constraints are that the Choi matrix is positive semidefinite. We use the approach outlined in the supplement of \cite{1} which is an extension of the state tomography techniques in \cite{26,27} as well as \cite{12}.

While there are issues regarding maximum likelihood estimation for QPT, there is a deeper problem that affects any reconstruction method that utilizes the likelihood function. To calculate the likelihood function we require $s_{j}$ which in turn implies we have a complete characterization of $\rho_{j}$ and $M_{j}$. In any experimental implementation of tomography there will be SPAM errors and in many situations the magnitude of the SPAM errors is the same order as the size of the gate errors we are trying to estimate. If the SPAM errors are stochastic, then the situation is reducible to the previous case by effectively treating the SPAM errors as additional sampling noise on the measurement. When the errors are systematic we may write $|\rho_{j}\rangle\rightarrow R_{E_{j}}|\rho_{j}\rangle$, $|M_{j}\rangle\rightarrow R_{F_{j}}|M_{j}\rangle$ and therefore $s_{j}\rightarrow R_{E_{j}}\otimes R_{F_{j}}s_{j}$.

For the remainder of this manuscript we will consider a slightly less general form of tomography that is applicable to many experimental implementations: a fixed initial state $|\rho_{0}\rangle$ and measurement operator $\langle\langle M_{0}|$ and a library of gates $G = \{R_{1}, R_{2}, \ldots R_{N}\}$. In this picture we describe experiments according to the convention

$$m_{ij} = \langle \langle M_{0}|R_{j}R_{\Lambda}R_{i}|\rho_{0}\rangle\rangle = \langle \langle M_{j}|R_{\Lambda}|\rho_{i}\rangle\rangle. \hspace{1cm} (11)$$

Systematic errors are described $G^{(\text{env})} = \{R_{E_{1}}R_{1}, R_{E_{2}}R_{2}, \ldots R_{E_{N}}R_{N}\}$, under the assumption that the initial gate-set was composed of unitary maps.

In Fig. 1 we simulate the effects of both systematic and stochastic measurement noise in the case where the measured gate is a perfect identity gate. This test is a good primitive for both theory and experiment since the identity is the one gate that should be perfectly implementable in any experiment by immediately performing measurement after state preparation (i.e. doing nothing for no time). In this figure the systematic SPAM error comes from a depolarizing channel of varying strengths $\mathcal{E}_{\text{dep}}$. We do not impose the CPTP constraint on the outcome and therefore measure the difference between the reconstructed gate and the identity by a diamond norm distance \cite{28}, which is calculated through the semidefinite programming technique in \cite{24}. In this simulation the gates map the $|0\rangle$ state to the four points on a tetrahedron which is essentially the most symmetric, minimal set of gates. We observe that for large noise powers the error in the estimate decreases exponentially with respect to decreasing noise power (and thus increasing repetitions) until it hits a floor determined by the systematic error in the interrogating gates.

In Fig. 1B we look at a similar plot for different libraries of unitary SPAM gates. The first two gate-sets are defined in terms of the set of states to which they map the qubit ground state $|0\rangle$: a tetrahedron or the six cardinal directions on the Bloch sphere. The second two
Since the two gates under comparison are physical we can
different models of the systematic errors. We find the
tility, in the absence of stochastic measurement errors, for
in the following simulations due to its ease of calculation.

Covariance method for our problem we use a flat metric
proven to be optimal. Since there is no gain in using the
metric over any other metric on operator space (e.g. the
flat metric). Minimizing the distance over the flat metric
then there is no reason to trust the covariance matrix
use to minimize the distance between a physical state and
of the Gaussian likelihood function as a metric which we
the space of physical maps using the covariance matrix
finite programming techniques constrain the estimator to
physical in the absence of any stochastic noise. Semidef-
error and over-rotation) as well three incoherent processes
(amplitude damping, dephasing and depolarizing errors).

groups are unitary 2-designs: a twelve element subgroup
of the Clifford group and the entire 24 element Clifford
group. Both of these groups correspond to the solid rota-
tions of a cube, with the twelve element subgroup consist-
ing of 180° rotations about the faces and 120° rotations
about the corners of the cube. Generically, increasing
the number of gates and maximizing the entropy of
SSP speeds convergence to the estimation floor, but the actual
limit remains unchanged.

By introducing systematic errors into our interrogating
gate-set the maximum of the original estimator Eq. 1
does not yield a faithful reconstruction of the unidentified
gate. In fact, the resulting estimate can be highly non-
physical in the absence of any stochastic noise. Semidefi-
nite programming techniques constrain the estimator to
the space of physical maps using the covariance matrix
of the Gaussian likelihood function as a metric which we
use to minimize the distance between a physical state and
the unphysical estimate. However, if this likelihood esti-

mator is incorrect in the absence of any stochastic noise
then there is no reason to trust the covariance matrix
metric over any other metric on operator space (e.g. the
flat metric). Minimizing the distance over the flat metric
can be preformed using the methods in [15], which was
proven to be optimal. Since there is no gain in using the
covariance method for our problem we use a flat metric
in the following simulations due to its ease of calculation.

In Fig. 2 we look at the error in reconstructing the iden-
tity, in the absence of stochastic measurement errors, for
different models of the systematic errors. We find the
closest physical state to the estimate under a flat metric.
Since the two gates under comparison are physical we can

QPT WITH OVERSAMPLING

From the previous section one can observe that to elim-
inate fully stochastic errors it is sufficient to oversample,
repeating experiments many times in order to accurately
measure expectation values. Oversampling is also an in-
tuitive approach to dealing with SPAM errors if instead of
increasing repetitions we alternatively increase the num-
ber of independent experiments (e.g. by increasing the
size of the gate library G). In this section we show that
this method does not compensate for general SPAM er-
rors, however when the gate library is a unitary 2-design
some uniform SPAM errors are removed.

Starting from Eq. 7 we can derive the bare estimate
in a slightly different manner that will be more conclusive
to the following discussion. First, note that we can mul-
ty both sides by $|M_j\rangle\langle\rho_i|$ and sum over $i$ and $j$ to obtain

$$
\sum_{ij} m_{ij} |M_j\rangle\langle\rho_i| = \left( \sum_j |M_j\rangle\langle\rho_j| \right) \mathcal{R}_A \\
\times \left( \sum_i |\rho_i\rangle\langle\rho_i| \right).
$$

(12)
In this case the bare estimate is
\[
\mathcal{R}_A^{(\text{est})} = \left( \sum_{j' | M_{j'}} \langle |M_{j'}| \rangle \langle \langle M_{j'} | \rangle \right)^{-1} \left( \sum_{ij} m_{ij} |M_{ij}| \langle \langle \rho_i | \rangle \right) \\
\times \left( \sum_{\rho_i} \langle \langle \rho_i | \rangle \right)^{-1}.
\] (13)

If there are SPAM errors the term \( m_{ij} \) has the form
\[
m_{ij} = \langle |M_0| \mathcal{R}_0 |M_j\rangle | \mathcal{R}_{i} | \mathcal{R}_j | \rho_0 \rangle \langle \langle \rho_i | \rangle \rangle.
\] (14)
which leads to,
\[
\mathcal{R}_A^{(\text{est})} = \left( \sum_{j' | M_{j'}} \langle |M_{j'}| \rangle \langle \langle M_{j'} | \rangle \right)^{-1} \left( \sum_{j} \mathcal{R}_j |M_0\rangle \langle \langle \rho_0 | \rangle \rangle \langle \langle \rho_i | \rangle \rangle \right) \\
\times \mathcal{R}_\Lambda \sum_{i} | \mathcal{R}_\mathcal{E} | \mathcal{R}_i | \rho_0 \rangle \langle \langle \rho_i | \rangle \rangle \mathcal{R}_i \left( \sum_{\rho_i} \langle \langle \rho_i | \rangle \right)^{-1}.
\] (15)

There are some very interesting consequences of the final expression which we shall examine in a few special cases. If the SPAM error is independent of the gate, then due to our error conventions we get a cancellation of the final term in Eq. (15) yielding,
\[
\mathcal{R}_A^{(\text{est})} = \left( \sum_{j | M_{j}} \langle |M_{j}| \rangle \langle \langle M_{j} | \rangle \right)^{-1} \left( \sum_{j' | M_{j'}} \langle |M_{j'}| \rangle \langle \langle M_{j'} | \rangle \rangle \right) \\
\times \mathcal{R}_\Lambda | \mathcal{R}_\mathcal{E} |.
\] (16)

We could derive similar expressions if the constant error occurred before the gate or in some combination of pre and post-gate error. If in addition the error commutes with all of the gates in the library (as is the case with depolarizing errors) the result is, for example
\[
\mathcal{R}_A^{(\text{est})} = \mathcal{R}_\text{dep} \mathcal{R}_\Lambda \mathcal{R}_\text{dep},
\] (17)

independent of the set of gates used for tomography. This form of the reconstructed map lends some intuition to Fig. 2, where the depolarizing error and other incoherent errors lead to a tomography error that had polynomial scaling with respect to the error on the gate-set. Qualitatively, errors of this form 'commute' with the reconstruction procedure.

If the error does not commute with all of the gates but the library forms a unitary 2-design \[30\] we can also simplify equation Eq. (16). We start by independently analyzing the two quantities in parenthesis in Eq. (16). These sums are both twirls over a unitary 2-design, \( \mathcal{W} \), which have the following generic form
\[
\mathcal{W}(A) = |A\rangle \langle A| \mathcal{E}(\langle |A| \rangle + \langle \langle |A| |A\rangle \rangle) \frac{\text{Tr}(A) - \langle |A| \rangle \langle |A| |A\rangle \rangle}{d^2 - 1},
\] (18)

[\text{S}]. Therefore, the first factor is given by
\[
\mathcal{W}(|M_0\rangle \langle \langle M_0 | \rangle \rangle) = |M_0\rangle \langle \langle |M_0| \rangle \rangle^2 \\
+ (\langle |M_0| \rangle) \left( \langle |M_0| \rangle - \langle |M_0| \rangle^2 \right),
\] (19)
and the second by
\[ W(|M_0\rangle\langle M_0|\mathcal{R}_\mathcal{E}) = |I\rangle\langle I|(|I\rangle\langle M_0|\langle M_0|R_{\mathcal{E}}|I\rangle) + (\mathbb{I} - |I\rangle\langle I|) \frac{\langle\langle M_0|R_{\mathcal{E}}|M_0\rangle - \langle\langle I|M_0\rangle\langle M_0|R_{\mathcal{E}}|I\rangle}{d^2 - 1}. \] (20)

Both of these are proportional to depolarizing channels which generically have the form \(|I\rangle\langle I| + \epsilon(1 - |I\rangle\langle I|). The product \(W(|M_0\rangle\langle M_0|^{-1}W(|M_0\rangle\langle M_0|R_{\mathcal{E}}|I\rangle)\) is therefore proportional to a depolarizing channel with a strength
\[ \epsilon = \frac{\langle\langle M_0|R_{\mathcal{E}}|M_0\rangle\langle\langle I|M_0\rangle\rangle - \langle\langle I|M_0\rangle\langle M_0|R_{\mathcal{E}}|I\rangle\rangle}{\langle\langle I|M_0\rangle\langle\langle I|M_0\rangle\rangle - \langle\langle I|M_0\rangle\langle M_0|R_{\mathcal{E}}|I\rangle\rangle}, \] (21)

and proportionality constant
\[ \alpha = \frac{\langle\langle M_0|R_{\mathcal{E}}|I\rangle\rangle}{\langle\langle I|M_0\rangle\rangle}. \] (22)

The bare estimate for a constant error on a gate-set composed of a unitary 2-design is thus given by
\[ R_{\lambda}^{(\text{est})} = \alpha R_{\mathcal{E}}^{(\text{dep})} R_{\lambda} R_{\mathcal{E}}. \] (23)

The previous theoretical analysis suggests that there may be a fundamental difference when the gate library is a unitary 2-design, and this is confirmed in Fig. 3. When we measure the errors in terms of the gate fidelity we see a clear difference with respect to the unitary 2-designs. In that case a global unitary error has the same effect on the fidelity as a depolarizing error, as expected from the previous argument. QPT errors resulting from independent unitaries on each of the gates remains qualitatively the same for all four sets. What we describe as a global unitary error is a strange object and should not be confused for control errors such as over rotation or detuning errors, but is instead more akin to a global frame misalignment.

Interestingly, the errors look very different when measured by the diamond norm (the right and left plots of Fig. 3). The diamond norm distance of the reconstruction scales polynomially with the gate-set diamond norm error for all forms of the SPAM error and gate-sets. The fact that the rate of the depolarizing error spans twice as large a domain as the coherent errors in the diamond norm is also curious. These two metrics clearly yield very different results and it is an important question as to which is appropriate for an experiment to report. For this reason, we consider both for the remainder of this manuscript.

**Self-consistent tomography**

We have thus far failed at tomography in this manuscript for the simple reason that our methodology is based on the likelihood function, and we have used the wrong one. A proper likelihood function should incorporate the SPAM errors, but characterizing the SPAM errors themselves requires QPT and therefore a likelihood function. This is an impasse for standard QPT methods. The alternative we present in this section is to assume nothing about the SPAM errors (though later we will bound their magnitude) and instead estimate an entire library of gates in a self-consistent manner.

Let us write the set of experimental gates as \(G^{(\text{exp})} = \{\tilde{R}_{\mathcal{E}} \cup \mathcal{R}_1, \tilde{R}_{\mathcal{E}} \cup \mathcal{R}_2, \ldots \tilde{R}_{\mathcal{E}} \cup \mathcal{R}_N\}\) which is a faulty version of the library \(G^{(\text{ideal})} = \{\mathcal{R}_1, \mathcal{R}_2, \ldots \mathcal{R}_N\}\). The free parameters of the this model are the error terms \(\tilde{R}_{\mathcal{E}}\) while the ideal gates are fixed. We begin by performing a series of \(N^3\) experiments of the form
\[ m_{ijk} = \langle\langle M_0|\tilde{R}_{\mathcal{E}} R_i \tilde{R}_{\mathcal{E}} R_j \tilde{R}_{\mathcal{E}} R_k|\rho_0\rangle\rangle. \] (24)

which in essence consists of the tomography data necessary to reconstruct each element of \(G^{(\text{exp})}\).

The likelihood of a trial set of gates \(\tilde{G} = \{\tilde{R}_{\mathcal{E}}, \mathcal{R}_1, \tilde{R}_{\mathcal{E}} \cup \mathcal{R}_2, \ldots \tilde{R}_{\mathcal{E}} \cup \mathcal{R}_N\}\) is now given by
\[ \mathcal{L}(\tilde{G}) = \exp \left( - \sum_{i,j,k} |m_{ijk}| - \langle\langle M_0|\tilde{R}_{\mathcal{E}} R_i \tilde{R}_{\mathcal{E}} R_j \tilde{R}_{\mathcal{E}} R_k|\rho_0\rangle\rangle \right)^2. \] (25)

The constraints such that the gate library is physical is that each of the maps \(\tilde{R}_{\mathcal{E}}\) are completely positive and trace preserving. Finding a maximum likelihood estimate is equivalent to minimizing the least square estimator,
\[ \text{LSQ}(G) = \sum_{i,j,k} \left| m_{ijk} - \langle\langle M_0|\tilde{R}_{\mathcal{E}} R_i \tilde{R}_{\mathcal{E}} R_j \tilde{R}_{\mathcal{E}} R_k|\rho_0\rangle\rangle \right|^2, \] (26)
but is highly non-trivial due to the fact that this function is 6th order in the gate-set.

We can reduce the order of the estimator by linearizing the evolution about the ideal set of gates. This is a reasonable assumption if we are testing components of a faulty quantum computer as opposed to probing some unknown physics or interaction. In this situation we allow \(\tilde{R}_{\mathcal{E}} = R_i + \tilde{E}_i\) where \(\tilde{E}_i\) is not necessarily physical but is small, \(||\tilde{E}_i|| \ll 1\). Expanding the evolution to first order in \(\tilde{E}_i\) we have
\[ \tilde{m}_{ijk} = \langle\langle M_0|\tilde{R}_{\mathcal{E}} R_i \tilde{R}_{\mathcal{E}} R_j \tilde{R}_{\mathcal{E}} R_k \rangle\rangle \] (27)

or by transforming back to the physical \(\tilde{R}_{\mathcal{E}}\)
\[ \tilde{m}_{ijk} = \langle\langle M_0| - 2R_i \tilde{R}_{\mathcal{E}} R_j \tilde{R}_{\mathcal{E}} R_k \rangle\rangle \] (28)
which implies that the matrix corresponding to the interference which is that the optimal fit is no longer unique, and the finite program. In fact, it is isomorphic to the original QPT problem on a larger space where the physicality constraint is another example of a semidefinite program. Therefore self-consistent reconstruction of a set of gates can be described as the following least square estimation

\[
\min_{\tilde{R}_{E_1}, \tilde{R}_{E_2}, \ldots} \sum_{ijk} |m_{ijk} + \langle \langle M_0 \rangle \rangle R_k R_j R_i - \tilde{R}_{E_k} R_k R_j R_i - R_k \tilde{R}_{E_j} R_j R_i - R_k R_j \tilde{R}_{E_i} R_i| |\rho_0| \rangle \rangle^2.
\]

subject to the constraints that the error maps \( \tilde{R}_{E_i} \) are all physical.

Minimizing this linearized least-square fit under the physicality constraint is another example of a semidefinite program. In fact, it is isomorphic to the original QPT problem on a larger space where \( \mathcal{R}_A \rightarrow \mathcal{R}_{E_1} \oplus \mathcal{R}_{E_2} \oplus \ldots \oplus \mathcal{R}_{E_N} \). There is however one major difference which is that the optimal fit is no longer unique, which implies that the matrix corresponding to \( SS^T \) in this space is not full-rank. As an example take any unitary transformation \( G^{(\text{ideal})} = \{ R_1, R_2, \ldots, R_N \} \rightarrow G^{(\text{err})} = \{ R_1^T R_1 R_U, R_2^T R_2 R_U, \ldots, R_N^T R_N R_U \} \) where \( R_U | \rho_0 \rangle \rangle \langle \langle M_0 \rangle \rangle R_U^T = | \rho_0 \rangle \rangle \langle \langle M_0 \rangle \rangle \). The measurement outcomes are invariant to such a transformation since

\[
m_{ijk} = \text{Tr} \left( | \rho_0 \rangle \rangle \langle \langle M_0 \rangle \rangle R_i R_j R_k \right) \n- \text{Tr} \left( | \rho_0 \rangle \rangle \langle \langle M_0 \rangle \rangle R_U^T R_i R_j R_k R_U \right) \n= \text{Tr} \left( R_U | \rho_0 \rangle \rangle \langle \langle M_0 \rangle \rangle R_U^T R_i R_j R_k R_U \right) = m_{ijk} \quad (30)
\]

When we prepare and measure in the computational basis this is equivalent to a frame invariance with respect to a diagonal unitary transformation. This is not analogous to moving to the interaction picture since the resulting frame is stationary, and does not generate an inertial correction to the Hamiltonian. It is more akin to an energy rescaling. Such a frame is physically irrelevant and in the following simulations we calculate distance between gate-sets by optimizing to find the most generous frame.

In Fig. 4 we compare the results of standard QPT and our self-consistent approach. The error model is that of individual random unitary errors, the uncorrectable error from the previous section, and the gate library is the six element gate-set that maps \( |0 \rangle \rangle \) to the cardinal directions on the Bloch sphere. We plot the errors from both the fidelity and the diamond norm of the estimates versus the error of the physical gate-set with respect to the ideal target. For each we show the average result for the six gates. Additionally, we also plot the ratio of the error in the estimate to the ideal for both measures.

In both cases there appears to be an exponential improvement in accuracy over standard QPT with respect to decreasing gate error. This is fundamentally different from the results of the previous section where the fidelity error changed substantially but the diamond norm remained qualitatively the same, and so this method is not an effective depolarization of the error.

To clarify the differences between the two approaches
we plot the least square estimator Eq. (25) for the two estimates in Fig. 5. Again we look at the ratio of the value of the least square fit for the estimated value divided by the fit of the ideal gates to the data. The estimate from QPT is flat with respect to lowering gate error which corresponds to the fact that QPT is minimizing the wrong quantity. Our estimator does not properly minimize this quantity either, due to the linear approximation, but the scaling is dramatically more favourable.

There are some obvious shortcomings of this method. We have increased the number of experiments and the complexity of the reconstruction algorithm. These increases are both polynomial in the dimension of the system, but since the dimension is already exponential in the number of qubits the resulting overhead is non-trivial. In principle, we can apply these techniques to two or three qubit systems. However, the reconstructions in Fig. 4 are already taxing the default numerical precision of our semidefinite optimizer. If we extended the plot in Fig. 5 to the left we would very quickly see the self-consistent method flatten from the numerical tolerance. Also, the maximization of diagonal frames is computationally expensive, generally more so than the SDP solver itself.

**CONCLUSION**

The methods of tomography are very powerful when using a well-characterized system to probe a black-box operation, but this is rarely the case when verifying quantum computing hardware. Instead, the situation is one in which all components are faulty, and in that scenario we have shown that process tomography can fail dramatically. Instead of separating the preparation and measurement phases of tomography we instead must view the experiment as a sequence of faulty gates, and while we have not discussed errors on the fiducial state and fixed measurement operator, those too should probably be brought into question in the future. Without incorporating these errors into our model it is impossible to obtain a more accurate reconstruction through statistical analysis alone.

In this manuscript we have developed a new protocol that self-consistently reconstructs a library of quantum gates by modifying the likelihood function of quantum process tomography to incorporate our uncertainty of the state preparation and measurement gates. In simulation, this self-consistent method outperforms standard QPT in terms of the accuracy of estimation since QPT consistently underestimates gate fidelities in the presence of SPAM errors. This method requires the same number of experiments as QPT and only adds polynomial overhead to the amount of classical post-processing. In fact, in any case where tomography has been performed on the entire set of gates used for state preparation and measurement no further experiments are needed to apply our protocol. Colloquially, we have taken to calling this the ‘overkill’ method of tomography, since the number of the experiments borders on the ridiculous for all but one or two qubit systems, but this will be the case for any protocol that provides full information about a library of
TABLE I. The gate error from QPT and self-consistent tomography for A) a SJT capacitively coupled to a coplanar waveguide and B) a SJT suspended in a copper cavity. We report the fidelity error and diamond norm for both approaches (QPT and self-consistent tomography), and for each gate in the set. The last row in each sample, labeled $\langle U \rangle$, is the average of the previous four as well as the average error obtained from randomized benchmarking.

| Sample | Gate | $1 - F_g(\Lambda_{QPT}, \Lambda_{ideal})$ | $1 - F_g(\Lambda_{SC}, \Lambda_{ideal})$ | $1 - F_g$ (RB) | $\|\Lambda_{QPT} - \Lambda_{ideal}\|_F$ | $\|\Lambda_{SC} - \Lambda_{ideal}\|_F$ |
|--------|------|------------------------------------------|------------------------------------------|----------------|---------------------------------|---------------------------------|
| A      | $I$  | 0.0058                                   | 0.0051                                   | -              | 0.037                           | 0.029                           |
| A      | $X_x$| 0.013                                    | 0.0098                                   | -              | 0.075                           | 0.042                           |
| A      | $X_{\pi/2}$ | 0.0077                   | 0.0047                                   | -              | 0.049                           | 0.018                           |
| A      | $Y_{\pi/2}$ | 0.0096                   | 0.0053                                   | -              | 0.069                           | 0.068                           |
| A      | $\langle U \rangle$ | 0.0090                   | 0.0062                                   | 0.0029         | 0.057                           | 0.039                           |
| B      | $I$  | 0.11                                     | 0.037                                    | -              | 0.43                            | 0.13                            |
| B      | $X_x$| 0.039                                    | 0.0057                                   | -              | 0.31                            | 0.014                           |
| B      | $X_{\pi/2}$ | 0.040                   | 0.014                                    | -              | 0.37                            | 0.090                           |
| B      | $Y_{\pi/2}$ | 0.032                   | 0.021                                    | -              | 0.32                            | 0.090                           |
| B      | $\langle U \rangle$ | 0.057                   | 0.020                                    | 0.0016         | 0.36                            | 0.087                           |

maps. Self-consistently reconstructing a library of gates, as opposed to performing tomography on the individual members, yields more trustworthy and generally higher fidelity estimates in both experiment and simulation and we expect the difference between standard QPT and the self-consistent method presented here to grow more pronounced as we reach lower and lower gate errors.

We acknowledge contributions from J. Chow and Thomas Ohki and discussions with R. Blume-Kohout and Graeme Smith. We acknowledge support from IARPA under contract W911NF-10-1-0324. All statements of fact, opinion or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of the U.S. Government.
[1] J. M. Chow, J. M. Gambetta, A. D. Corcoles, S. T. Merkel, J. A. Smolin, C. Rigetti, S. Poletto, G. A. Keefe, M. B. Rothwell, J. R. Rozen, M. B. Ketchen, and M. Steffen, Phys. Rev. Lett. 109, 060501 (2012).

[2] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Nature 460, 240 (2009).

[3] R. C. Bialczak, M. Ansmann, M. Hofheinz, E. Lucero, M. Neeley, A. D. O’Connell, D. Sank, H. Wang, J. Wenner, M. Steffen, A. N. Cleland, and J. M. Martinis, Nat Phys 6, 409 (2010).

[4] H. Häffner, W. Hansel, C. F. Roos, J. Benhelm, D. Chek-al kar, M. Chwalla, T. Korber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Guhne, W. Dur, and R. Blatt, Nature 438, 643 (2005).

[5] J. P. Gaebler, A. M. Meier, T. R. Tan, R. Bowler, Y. Lin, D. Hanneke, J. D. Jost, J. P. Home, E. Knill, D. Leibfried, and D. J. Wineland, Phys. Rev. Lett. 108, 260503 (2012).

[6] M. A. Nielsen, Phys. Lett. A 303, 249 (2002).

[7] I. L. Chuang and M. Nielsen, Journal of Modern Optics 44, 2455 (1997).

[8] J. M. Gambetta, A. D. Corcoles, S. T. Merkel, B. R. Johnson, J. A. Smolin, J. M. Chow, C. A. Ryan, C. Rigetti, S. Poletto, T. A. Ohki, M. B. Ketchen, and M. Steffen, arXiv:1204.6308 (2012).

[9] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A 64, 052312 (2001).

[10] Z. Hadži, J. Rehacek, J. Fiurášek, and M. Jezek, Lect. Notes Phys. 649, 59 (2004).

[11] A. I. Lvovsky, Journal of Optics B: Quantum and Semiclassical Optics 6, S556 (2004).

[12] M. D. de Burgh, N. K. Langford, A. C. Doherty, and A. Gilchrist, Phys. Rev. A 78, 052122 (2008).

[13] R. Blume-Kohout, Phys. Rev. Lett. 105, 200504 (2010).

[14] R. Blume-Kohout, New Journal of Physics 12, 043034 (2010).

[15] J. A. Smolin, J. M. Gambetta, and G. Smith, Phys. Rev. Lett. 108, 070502 (2012).

[16] M. Christandl and R. Renner, Phys. Rev. Lett. 109, 120403 (2012).

[17] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Phys. Rev. A 77, 012307 (2008).

[18] E. Magesan, J. M. Gambetta, and J. Emerson, Phys. Rev. Lett. 106, 180504 (2011).

[19] E. Magesan, J. M. Gambetta, and J. Emerson, Phys. Rev. A 85, 042311 (2012).

[20] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, M. B. Ketchen, and M. Steffen, Phys. Rev. Lett. 109, 080505 (2012).

[21] A. M. Braczyk, D. H. Mahler, A. I. Lvovsky, and D. F. V. James, New Journal of Physics 14, 085003 (2012).

[22] V. V. Dolrovitski, G. de Lange, D. Rist, and R. Hanson, Phys. Rev. Lett. 105, 077601 (2010).

[23] M.-D. Choi, Linear Algebra and its Applications 10, 285 (1975).

[24] S. Boyd and L. Vandenberghe, (Cambridge University Press).

[25] J. F. Sturm, Optimization methods and Software 11, 625 (1999).

[26] A. Silberfarb, P. S. Jessen, and I. H. Deutsch, Phys. Rev. Lett. 95, 030402 (2005).

[27] C. A. Riofro, P. S. Jessen, and I. H. Deutsch, Journal of Physics B: Atomic, Molecular and Optical Physics 44, 154007 (2011).

[28] A. Y. Kitaev, A. Shen, and M. N. Vyalyi, Classical and Quantum Computation (American Mathematical Society, 2002).

[29] J. Watrous, Theory of Computing 5, 11.

[30] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Phys. Rev. A 80 (2009), 10.1103/PhysRevA.80.012304.

[31] C. Rigetti, J. M. Gambetta, S. Poletto, B. L. T. Plourde, J. M. Chow, A. D. Corcoles, J. A. Smolin, S. T. Merkel, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, Phys. Rev. B 86, 100506 (2012).