SANDIA REPORT
SAND2020-10396
Unclassified Unlimited Release
Printed September 17, 2020

Sandia National Laboratories

Diagnosing and Destroying Non-Markovian Noise

Kevin Young

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185
Livermore, California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

Available to DOE and DOE contractors from

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods

Diagnosing and Destroying Non-Markovian Noise

Kevin Young
Quantum Performance Laboratory
Sandia National Laboratories
Livermore, CA 94550
kyoung@sandia.gov

Stephen Bartlett, Robin Blume-Kohout, John King Gamble,
Daniel Lobser, Peter Maunz, Erik Nielsen,
Timothy Proctor, Melissa Revelle, Kenneth Rudinger

SAND2020-10396

ABSTRACT

Nearly every protocol used to analyze the performance of quantum information processors is based on an assumption that the errors experienced by the device during logical operations are constant in time and are insensitive to external contexts. This assumption is pervasive, rarely stated, and almost always wrong. Quantum devices that do behave this way are termed “Markovian,” but nearly every system we have ever probed has displayed drift or crosstalk or memory effects – they are all non-Markovian. Strong non-Markovianity introduces spurious effects in characterization protocols and violates assumptions of the fault-tolerance threshold theorems.

This SAND report details a three year laboratory-directed research and development (LDRD) project entitled, "Diagnosing and Destroying non-Markovian Noise in Quantum Information Processors." This program was initiated to build tools to study non-Markovian dynamics and quantum systems and develop robust methodologies for eliminating it. The program achieved a number of notable successes, including the first statistically rigorous protocol for identifying and characterizing drift in quantum systems, a formalism for modeling memory effects in quantum devices, and the successful suppression of drift in a Sandia trapped-ion quantum processor.

Acknowledgments

This work could not have been accomplished without the assistance and encouragement of our many collaborators. Much of the theoretical work performed under this project was performed in collaboration with the other members of Sandia's Quantum Performance Laboratory, Robin Blume-Kohout, Erik Nielsen, Timothy Proctor, and Kenneth Rudinger. Theoretical work on memory-augmented tomography was conducted in collaboration with Stephen Bartlett of the University of Sydney. Experimental work on the trapped-ion system was conducted by Peter Maunz, Susan Clark, Melissa Revelle, Dan Lobser, and Craig Hogle. Experimental work on the silicon qubit device was conducted by Dwight Luhman and Chloe Bureau-Oxton.

This work was funded by LDRD Project Number 18-0402, "Diagnosing and Destroying Non-Markovian Noise."

CONTENTS

LIST OF FIGURES

Figure 2-3. **Measuring qubit stability using time-resolved GST**. The results of two time-resolved GST experiments – using the gates G_i, G_x and G_y – with adjustments made after the first experiment aimed at stabilizing the qubit. **A-B.** The evidence for instability in each circuit in the first experiment, quantified by λ_p (see main text). A pixel is colored when λ_p is large enough to be statistically significant, otherwise it is greyscale. Each circuit consists of repeating a *germ* sequence l times in between six initialization and pre-measurement sequences. The data is arranged by germ and approximate circuit length L, and then separated into the 6×6 different preparation and measurement sequence pairs, as shown on the axes of B ("{}" denotes the null sequence). Only long circuits containing repeated applications of G_i exhibit evidence of drift. In the second experiment none of the λ_p are statistically significant (data not shown). **C-D.** The result of time-resolved tomographic reconstructions of the gates in each experiment, summarized by the diamond distance error of each gate, and the decomposition of the coherent errors in the idle gate G_i into rotation angles around $\hat{x}$, $\hat{y}$ and $\hat{z}$, over the duration of each experiment ($t_{\text{max}} \approx 8$ hours and $t_{\text{max}} \approx 2.5$ days for the first and second experiment, respectively). **E.** The power spectrum for each experiment obtained by averaging the individual power spectra for the different circuits, with filled points denoting power above the 5% significance thresholds (not shown). 17

1. INTRODUCTION

The continued rapid progress in qubit technology development depends critically on reliable characterization protocols that quantify as-built qubits' performance, identify error modes, and guide R&D for the next generation. Sandia's gate set tomography (GST) [1, 2] is the most powerful protocol of this sort. All of these protocols, including GST, assume that the qubits' noise is Markovian – it is not changing or correlated in time. But real qubits always suffer from significant "non-Markovian" noise, and in nearly every qubit system we have probed, the noise drifts and/or displays memory effects. When this Markovianity assumption is violated, characterization protocols quickly become unreliable, and in the worst case their results are meaningless and unpredictable. Non-Markovianity therefore poses a massive source of unmitigated risk for future qubit technologies, limiting the performance of all known quantum algorithms and violating core assumptions of the fault-tolerance threshold theorem. It also inhibits qubit characterization – it's impossible to assess quantum hardware in the presence of uncontrolled and unidentified non-Markovianity.

The goal of this program was to solve these problems by creating new, reliable protocols for modeling and characterizing non- Markovian noise. We identified two distinct and experimentally relevant classes of non-Markovianity, low-frequency drift and short-time memory effect, and targeted our our work at identifying and characterizing these particular effects. The result is a suite of non-Markovianity tools to support comprehensive assessment of quantum device performance. These tools provide means to augment standard characterization and performance assessment protocols so that they continue to function in the presence of non-Markovianity. Among the most important of these newly-augmented tools is gate set tomography.

1.1. Non-Markovian gate set tomography

Quantum process tomography is a technique designed to learn a process-matrix representation of the logic operations of a quantum devices. This protocol is relatively straightforward, but it suffers from self- consistency problems and cannot cope with state preparation and measurement (SPAM) errors. These problems motivated the development of randomized benchmarking (RB), which solves the problems with SPAM error, but is only able to measure a single reduced metric of processor performance. While convenient for reporting and tracking broad performance trends, this particular fidelity is not obviously relevant to the long term goal of quantum error correction. RB also provides no diagnostic information about the gates being characterized, which limits its utility for debugging and improving qubit performance.

Gate set tomography (GST) was designed to provide these diagnostics. GST reconstructs all aspects of the quantum logic gates, and then any metric of interest can be computed from that reconstruction. Standard GST is designed to characterize generic qubit systems, making minimal

assumptions about the underlying physics. This has allowed the use of GST to successfully characterize many physical systems, including Si/SiGe quantum dots, trapped ions, phosphorus impurities in Si and superconducting transmons. These successful experimental demonstrations have shown that GST is a useful tool for assessing and debugging qubits. The Sandia GST team has released a well-tested, open-source python GST package called pyGSTi.

Standard GST does make one strong, critical assumption about the underlying physics of the system: that the gates being characterized are stationary and Markovian. In other words, a gate's effect on the quantum processor does not depend on time, context, or other gates in the past. Unfortunately, there is overwhelming experimental evidence that this assumption is strongly violated in many physical systems. Standard GST can detect this "model violation", but has no ability to characterize the nature of the non-Markovianity or provide debugging hints.

GST is an extremely flexible platform, however, and in this report we detail methods for extending it to encompass two distinct kinds of non-Markovianity: drift and memory. In Sec. 2 we report our techniques and tools for identifying and characterizing drift in quantum information processors. In Sec. 3, we report on a new class of quantum process models that are able to capture a wide range of memory effects, including ion heating, pulse duration spillover, and coupling to classical degrees of freedom.

2. TIME DEPENDENCE

Quantum devices are prone to errors. Further, they are prone to *inconsistent* errors. Quantum information processors are extremely sensitive to their environment, and – almost invariably – this environment is unstable. Laser amplitudes change slowly, laser phases change quickly, and surface electromagnetic fields fluctuate erratically and across vastly disparate timescales. These environmental fluctuations conspire to leave quantum information processors exhibiting an always-changing cacophony of errors.

Though drift in experimental quantum hardware is ubiquitous, theoretical treatments of quantum devices are nearly always predicated on stable, Markovian models. Almost every quantum characterization protocol is similarly built on an unstated assumption that target system is stable, but drift in the target system violates these a assumptions and introduces uncontrolled errors in the output metrics and estimates. Furthermore, near-term quantum algorithms, such as the variational quantum eigensolver [3] or the quantum approximate optimization algorithm [4], are particularly sensitive to drift. These hybrid heuristic algorithms work by pairing a classical optimizer to a quantum accelerator – if the output of this accelerator is drifting, the entire algorithm can fail.

A particularly important casualty of drift is the stability assumption of the various fault-tolerance theorems [5, 6, 7]. These theorems are among the crown jewels of quantum information theory but they generally demand that errors be sufficiently uncorrelated and stable [5, 6]. Beyond the obvious consequences of increased average error rates, the role of drift in this case is somewhat subtle. If an environmental degree of freedom common to many qubits is stable, then the errors on these qubits can be perfectly uncorrelated, even if the environment is not optimally tuned. For instance, if the global, quantizing magnetic field of a set of spin qubits is a little higher than expected, then all of the qubits will experience small, coherent Z rotations. These errors are local and uncorrelated (since they are not random). But if the field is unstable, then all of the qubits will experience the same instability, and the errors will become *correlated*. As an extreme example, if the field suddenly grew sufficiently large to generate a phase flip in a single clock cycle, that phase flip would be experienced by every single qubit. For many quantum codes, this corresponds to a logical phase flip – an uncorrectable error. Less-dramatic fluctuations will still be correlated across all qubits, increasing the probability of high-weight errors against which the code is not fault tolerant.

In order for quantum computers to be trusted, their temporal variability must be controlled and certified. While unstable *classical* systems can be readily measured and monitored for drift, quantum information processors are intrinsically and fundamentally limited by shot noise. Overcoming this shot noise either simultaneous measurement of many qubits, or sequential measurements of one or a few qubits over a long time. The first approach is experimentally demanding [8, 9], while the second raises complex questions about how to extract meaningful time-dependent estimates from shot-noise-corrupted time series data. In this program, we investigated this latter approach and

developed a robust, statistically rigourous framework to assess the stability of quantum systems. We discuss this work in Sec. 2.1.

In Sec. 2.2 we discuss a family of lightweight protocols for *controlling* drift in quantum information processors based on shot-noise-tolerant stochastic optimization. Our work in this area is preliminary and will be developed further under a DOE ECRP award.

2.1. Characterizing drifting quantum information processors

This thrust focused on development of the first statistically rigorous algorithms for identifying and characterizing drift in quantum information processors. These tools have been released publicly as a key component of the pyGSTi software package. Our methods have also been adopted by the IARPA LogiQ program as sufficient for satisfying the program requirements for device stability, and have seen broad adoption by the experimental quantum computing community. This work is captured in the following paper and its supplementary material.

Detecting, tracking, and eliminating drift in quantum information processors
Timothy Proctor, Melissa Revelle, Erik Nielsen, Kenneth Rudinger, Daniel Lobser, Peter Maunz, Robin Blume-Kohout, and Kevin Young
Sandia National Laboratories
Albuquerque, NM and Livermore, CA
In press at Nature Communications (accepted June 24, 2020)
Available at arXiv:1907.13608 [10]

Abstract If quantum information processors (QIPs) are ever to fulfill their potential, the diverse errors that impact them must be understood and suppressed. But errors fluctuate over time in most processors and the most widely used tools for characterizing them assume static error modes and rates. This mismatch can cause unheralded failures, misidentified error modes, and significant waste of experimental effort. Here, we demonstrate a fast spectral analysis technique for resolving time dependence in QIPs. Our method is simple and statistically sound, and it can be applied to time-series data from any repeated QIP experiment. We use data from both numerical simulations and trapped-ion qubit experiments to show how our method can resolve time dependence when coupled to popular characterization protocols, including randomized benchmarking, gate set tomography, and Ramsey spectroscopy. In the experiments, we detected instability and localized its source, then implemented drift control techniques to mitigate it, and finally used our methods to demonstrate that the instability had been eliminated. See Figs. 2-1, 2-2, and 2-3.

2.2. Controlling drift in quantum information processors

Computationally useful quantum information processors (QIP's) are expected to contain millions of qubits, each of which will undergo precisely tuned preparation, gate, and measurement operations. The experimentally tunable parameters that describe these operations are numerous and may include, for example: the strength of an externally applied magnetic field; the amplitudes, phases,

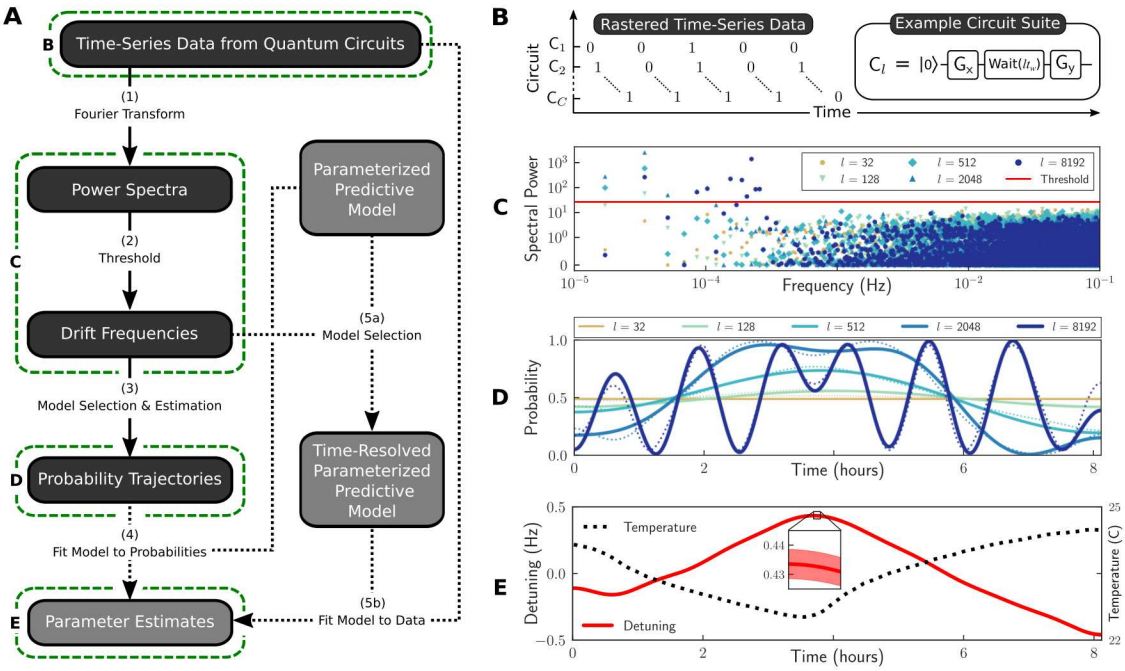

Figure 2-1. Diagnosing time-dependent errors in a quantum information processor. A. A flowchart of our methodology for detecting and quantifying drift in a QIP, using time-series data from quantum circuits. The core steps (1–3) detect instability, identify the dominant frequencies in any drift, and estimate the circuit outcome probabilities over time. Additional steps (4 and/or 5) estimate time-varying parameters (e.g., error rates) whenever a time-independent parameterized model is provided for predicting circuit outcomes. B. An example of circuits on which this technique can be implemented – Ramsey circuits with a variable wait time lt_w – as well as an illustration of data obtained by "rastering" (each circuit is performed once in sequence and this sequence is repeated N times). C-E. Results from performing these Ramsey circuits on a $^{171}Yb^{+}$ ion qubit ($l = 1, 2, 4, \cdots, 8192$, $t_w \simeq 400\mu s$, $N = 6000$), and the corresponding stages of the flowchart. C. The power spectra observed in this experiment for selected values of l. Frequencies with power above the threshold almost certainly appear in the true time-dependent circuit probabilities, $p_l(t)$. D. Estimates of the probability trajectories (unbroken lines are estimates from step 3 of the flowchart; dotted lines are the probabilities implied by the time-resolved detuning estimate shown in E). E. The standard Ramsey model $p_l(t) = A + B\sin(2\pi l t_w \Omega)$, where Ω is the qubit detuning, is promoted to a time-resolved parameterized model (step 5a) and fit to the data (step 5b), resulting in a time-resolved detuning estimate. The shaded area in the inset is a 2σ confidence region. The estimated detuning closely mirrors ambient laboratory temperature.

frequencies, and polarizations of laser or microwave pulses; the electrical potentials applied to surface electrodes; etc. Deviations of these parameters from their nominal values will necessarily result in increased error rates for quantum operations. In order to function properly, a large QIP must then be equipped with an automated suite of tools for tuning these parameters to their nominal values. Furthermore, as these parameters may change over time, these tools must be able to compensate for drift without introducing a significant increase in the error rate.

In this section, we introduce a set of tools for tuning and drift compensation (TDC) targeted specif-

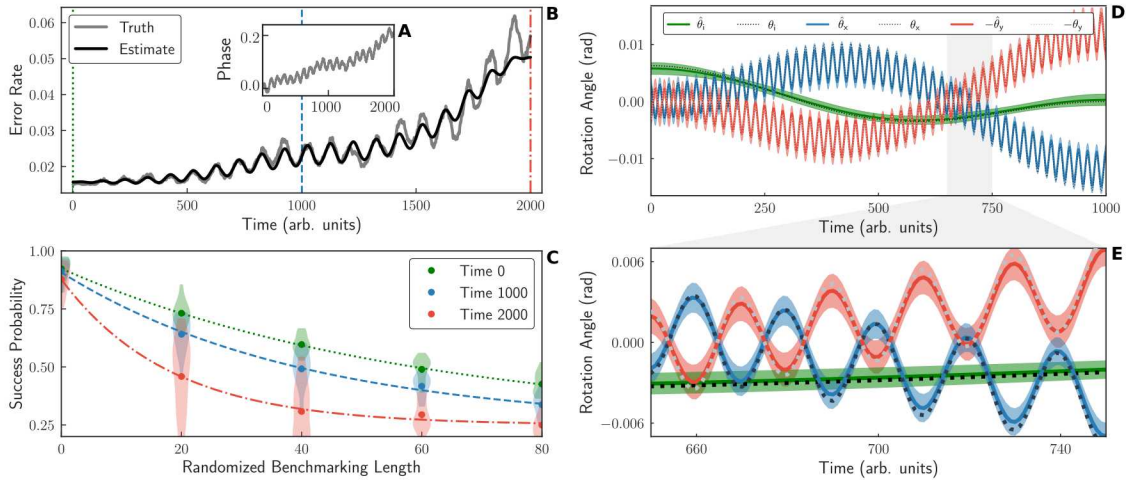

Figure 2-2. Time-resolved randomized benchmarking & tomography on simulated data. A-C. Time-resolved RB on simulated data for gates with time-dependent phase errors. A. The simulated phase error over time. B. The RB error rate (r) versus time and an estimate obtained from simulated data. C. Estimated instantaneous average-over-circuits (points) and per-circuit (distributions) success probabilities at each circuit length, and fits to an exponential (curves), for the three times denoted in B. Each instantaneous estimate of r is a rescaling of the decay rate of the exponential fit at that time. D-E. Time-resolved GST on simulated data, for three gates G_i, G_x and G_y that are subject to time-dependent coherent errors around the $\hat{z}$, $\hat{x}$ and $\hat{y}$ axes, respectively, by angles θ_i, θ_x and θ_y. The estimates of these rotation angles (denoted $\hat{\theta}_i$, $\hat{\theta}_x$ and $\hat{\theta}_y$) closely track the true values. The shaded areas are 2σ confidence regions.

ically to gate-model QIPs. Despite the wide variety of technologies used to implement quantum devices, our tools utilize only experimental resources that all such QIPs must admit: state preparation, application of quantum circuits, and measurement. The state preparation and measurements will be limited to the computational basis, while the quantum circuits, $\mathcal{S}_j$, will be composed of gates, $\mathcal{G}_i$, from some small set, $\{\mathcal{G}\}$. We don't demand that the gates be universal, or even generate a unitary t-design. The protocols we propose here will consist of specific quantum circuits as well as classical post-processing algorithms for updating the control parameters conditional on the circuit measurement results. For ease of presentation, we shall restrict our discussion to single-qubit experiments, but our results are easily generalizable to many qubits.

The quantum circuits we utilize fall into two broad classes. The first class consists of those circuits which, when perfectly implemented, yield a deterministic outcome (either 0 or 1). The second class consists of those for which the ideal outcome is random and unbiased (50% chance of either 0 or 1). For reasons that will be made clear shortly, we shall refer to the class of deterministic circuits as *quadratic*, and the class of unbiased random circuits as *linear*.

We indicate the state of the control parameters with a vector, $\boldsymbol{\eta}(t)$. At any given time, the error rate of the QIP would be minimized if the controls were set to their nominal values, $\boldsymbol{\eta}_0(t)$. It will be convenient to define the difference, $\boldsymbol{\delta}(t) = \boldsymbol{\eta}(t) - \boldsymbol{\eta}_0(t)$, which we call the *deviation*. The deviation characterizes how far the current control parameters of the QIP are from their optimal configuration. Importantly, all state preparations, $\rho(\boldsymbol{\delta}(t))$; gates, $\mathcal{G}(\boldsymbol{\delta}(t))$; and measurements, $\mathcal{M}(\boldsymbol{\delta}(t))$ depend on the deviation.

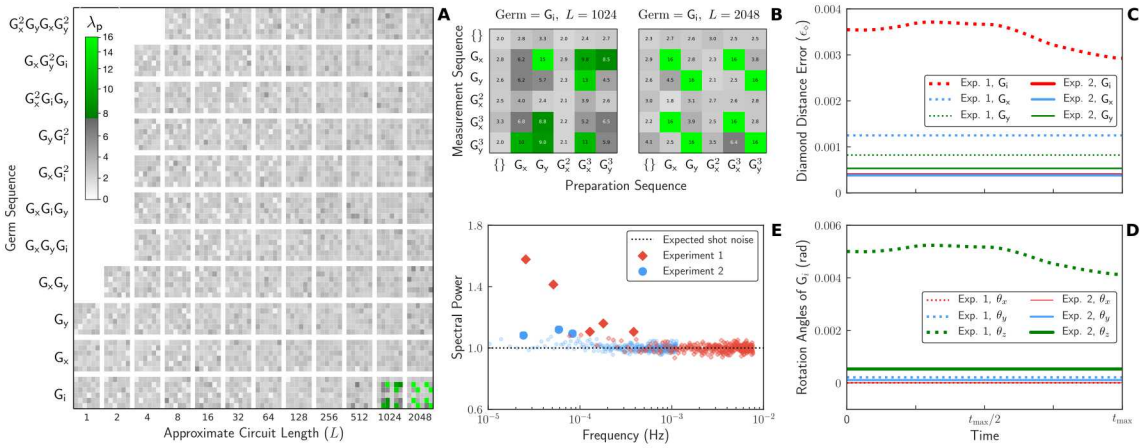

Figure 2-3. Measuring qubit stability using time-resolved GST. The results of two time-resolved GST experiments – using the gates G_i, G_x and G_y – with adjustments made after the first experiment aimed at stabilizing the qubit. A-B. The evidence for instability in each circuit in the first experiment, quantified by λ_p (see main text). A pixel is colored when λ_p is large enough to be statistically significant, otherwise it is greyscale. Each circuit consists of repeating a *germ* sequence l times in between six initialization and pre-measurement sequences. The data is arranged by germ and approximate circuit length L, and then separated into the 6×6 different preparation and measurement sequence pairs, as shown on the axes of B ("$\{\}$" denotes the null sequence). Only long circuits containing repeated applications of G_i exhibit evidence of drift. In the second experiment none of the λ_p are statistically significant (data not shown). C-D. The result of time-resolved tomographic reconstructions of the gates in each experiment, summarized by the diamond distance error of each gate, and the decomposition of the coherent errors in the idle gate G_i into rotation angles around $\hat{x}$, $\hat{y}$ and $\hat{z}$, over the duration of each experiment ($t_{\max} \approx 8$ hours and $t_{\max} \approx 2.5$ days for the first and second experiment, respectively). E. The power spectrum for each experiment obtained by averaging the individual power spectra for the different circuits, with filled points denoting power above the 5% significance thresholds (not shown).

Explicitly, the probability of measuring outcome j of a POVM, $\{\mathcal{M}_i\}$, after performing a sequence of quantum gates, $\mathcal{S}_i = \mathcal{G}_{i_n} \circ \ldots \circ \mathcal{G}_{i_2} \circ \mathcal{G}_{i_1}$, is:

$$p_j^{(i)}(\boldsymbol{\delta}) = \mathrm{Tr}\left(\mathcal{M}_j(\boldsymbol{\delta})\mathcal{S}_i(\rho(\boldsymbol{\delta});\boldsymbol{\delta})\right) \tag{2.1}$$

$$= \left\langle\left\langle \mathcal{M}_j(\boldsymbol{\delta})\right| \mathcal{S}_i(\boldsymbol{\delta}) \left|\rho(\boldsymbol{\delta})\right\rangle\right\rangle. \tag{2.2}$$

The second equality introduces an extension of the bra-ket notation to the Liouville space of superoperators on density matrices[11]. When the deviation is small, the probability of measuring 1 after some quantum circuit is well approximated by the Taylor expansion

$$p^{(i)}(\boldsymbol{\delta}) = p^{(i)}(0) + p^{(i)}_{,j}(0)\delta_j + \frac{1}{2}p^{(i)}_{,jk}(0)\delta_j\delta_k + O\left(\boldsymbol{\delta}^3\right) \tag{2.3}$$

where we have used the Einstein summation convention (implicit summation over repeated indices) and the comma derivative, defined as

$$p_{,j}(0) = \frac{\partial}{\partial\delta_j}p(\boldsymbol{\delta})\bigg|_{\boldsymbol{\delta}=0} \qquad p_{,jk}(0) = \frac{\partial^2}{\partial\delta_j\partial\delta_k}p(\boldsymbol{\delta})\bigg|_{\boldsymbol{\delta}=0}$$

By appropriate choice of the circuit, $\mathcal{S}_i$, one can make either the linear term or the quadratic term dominate the other in the Taylor series. For example, consider two very simple circuits composed only of $\pi/2$-pulses along σ_x:

$$\mathcal{S}_1(\delta) = \mathcal{G}_x(\delta) \tag{2.4}$$
$$\mathcal{S}_2(\delta) = \mathcal{G}_x(\delta) \circ \mathcal{G}_x(\delta) \tag{2.5}$$

In this example, the unknown parameter is the pulse amplitude, so $\mathcal{G}_x(\delta) = \exp(-i(\pi/4+\delta)\sigma_x)$ is a unitary operator. Assuming the qubit is prepared in the $|0\rangle$ state, then the probability of finding it in $|0\rangle$ state after the circuit is

$$p_1 = 1/2 - 1/2\sin(2\delta) \simeq 1/2 - \delta \tag{2.6}$$
$$p_2 = \sin(2\delta)^2 \simeq 4\delta^2 \tag{2.7}$$

It is here that we make connection to the previously introduced linear and quadratic circuit classes. Restricting to linear circuits, those for which the zero-deviation outcomes are uniformly random, small unitary errors can admit linear changes to circuit outcome probability. For quadratic circuits, those for which the zero-deviation outcomes are deterministic, small unitary errors can result in quadratic changes to the circuit outcome probability.

Because the response to deviation is different for these two classes of circuits, the TDC procedures which rely on such circuits will be different as well. Ultimately we will require methods to handle four broad classes of tuning and drift situations: single and multiple parameter TDC, each with either linear or quadratic circuits. As we mentioned above, this work remains under active development, and the remainder of this section will discuss:

1. Single-parameter, quadratic TDC
2. Single-parameter, linear TDC
3. Multi-parameter, linear TDC

We emphasize that we are *not* attempting optimal parameter estimation. The techniques we present are almost certainly not the most efficient probes for the deviation parameters, and they are not intended to be.

2.2.1. Single-parameter, quadratic tuning and drift control

In this section we consider the tuning of a single parameter using quadratic circuits. Recall that quadratic circuits are those for which the probability of an unexpected result should be quadratic for small deviation:

$$p_u(\delta) \simeq s\,\delta^2 \tag{2.8}$$

We call s the *sensitivity* of the circuit and will assume that it is known. Such circuits abound in quantum information – in the no noise/zero-deviation limit, π-gates, XY-Ramsey sequences, and syndrome extraction circuits all have have deterministic outcomes and are thus quadratically sensitive to parametric errors. The goal of a TDC protocol is to find a value of the parameter, η, for which the probability of an unexpected outcome is less than some specified error, ε.

Our algorithm for minimizing the error proceeds as follows. Take data until n_1 1's are obtained (with n_0 being the number of 0's recorded). The distribution of n_0 should be given by the negative binomial distribution. An unbiased estimator for the probability is $\hat{p} = (n_1 - 1)/(n_0 + n_1 - 1)$. From this we can estimate the magnitude of the deviation: $|\hat{\delta}| = \sqrt{\hat{p}/s}$. Since we don't know the sign of the deviation, we guess, assigning a value of plus- or minus-one to a coin, c, and adjust the parameter $\eta \leftarrow \eta + cg|\hat{\delta}|$, where g is a gain parameter. This procedure is repeated, alternating the sign of the coin each time, until the estimate of the probability is consistently below the target ε.

This procedure is very accurate and surprisingly robust. Alternating the sign of the coin as described above imposes a degree of self correction. If one choses the wrong direction at first, then the deviation is increased by the correction step, increasing the probability of error. In the next round, the probability estimate will reflect this increased error rate and the change in the coin's sign will ensure that this subsequent correction is in the right direction.

2.2.1.1. Example

In this section we present the simple case where only a single parameter is being adjusted and we restrict our quantum circuits to those for which $p(\mathbf{0}) = 0$ or 1. To illustrate our approach, we consider tuning the action of a π pulse:

$$G_x(\eta) = \exp(-i\eta\,\sigma_x). \tag{2.9}$$

In this case, the ideal value for the parameter is $\eta_0 = \pi/2$. Rewriting the gate in terms of $\delta = \eta - \eta_0$,

$$G_x(\delta) = \sigma_x \exp(-i\delta\sigma_x) \tag{2.10}$$

After repeated applications of the gate, the probability of an error is given simply by $p_{\text{err}} = \sin(\delta n)^2$. For small values of the deviation, this is simply $p_{\text{err}} = \delta^2 n^2$. So the error is quadratic in the deviation, as expected. For small deviation, the probability of measuring a 1 is

$$p_1(\delta) = \frac{1}{2}\delta^2 \frac{d^2}{d\delta^2} p(\mathbf{0}) \equiv s\delta^2 \tag{2.11}$$

where the sensitivity is equal to $s = \partial_\delta^2 p(\delta)|_{\delta=0}$ and will depend on the particular sequence being implemented. Our approach can be summarized with the following pseudocode:

2.2.2. Single-parameter, linear tuning and drift control

Consider a qubit prepared in the state $|+\rangle$ and allowed to freely evolve. If the frequency of the AC control field reference η_0 is not exactly equal to the Larmor frequency η, the qubit's Bloch vector will precess in the x/y-plane at an angular frequency, $\omega = (\eta - \eta_0)$. By applying a $\pi/2$-pulse about the x-axis after some time, t, the qubit is left in the state

$$|\psi\rangle = \cos(\pi/2 + (\eta - \eta_0)t)\,|1\rangle + \cos(\pi/2 - (\eta - \eta_0)t)\,|-1\rangle\,, \tag{2.12}$$

ALGORITHM 2.1. SINGLEPARAMETERQUADRATICTDC

1.1

Input: $\eta, \varepsilon \in (0,1), k_0 \in \mathbb{Z}_{>1}$

```
c ← 1, k ← 1, n ← 1
while k/n > ε do
    n ← 0, k ← 0
    while k < k_0 do
        PREPARE |0⟩
        APPLY G_x(η)
        if MEASURE |0⟩⟨0| then
            k ← k + 1
        end if
        n ← n + 1
    end while
    η ← η + c√((k − 1)/(s(n − 1)))
    c ← −c
end while
```

where we have departed from the usual $|0/1\rangle$ qubit state notation in favor of $|\pm 1\rangle$. The probability of measuring $r \in \{-1, 1\}$ is then

$$\Pr(r|\eta) = \frac{1}{2}\left(1 + r\sin((\eta - \eta_0)t)\right) \simeq \frac{1}{2}(1 + t(\eta - \eta_0)) \tag{2.13}$$

For small, θ, the probability is approximately linear in $\eta - \eta_0$. Note that if we observe $r = 1$, it is more likely that $\eta > \eta_0$ that the reverse (and visa-versa). This suggests a stochastic optimization approch to stabilizing the η parameter.

Explicitly, consider a one-parameter model with circuits whose outcome distributions are linearly sensitive to the parameter η. As above, we can write the probability of measuring r as given by

$$\Pr(r|\eta) = \frac{1}{2}\left(1 + r\sin\left(2s(\eta - \eta_0)\right)\right) \tag{2.14}$$

where η_0 is the ideal operating point and s is a sensitivity parameter characterizing the amplification power of the sequence (equal to $t/2$ in the earlier example). With each experiment, we get a result, $r_n \in \{-1, 1\}$, and adjust the parameter through

$$\eta_{n+1} = \eta_n + g r_n / s \tag{2.15}$$

for g, a gain parameter. The gain parameter should be chosen small, $g \ll 1$, so that the system is never kicked outside the linear response region of of the probability function.

Because r_n is a random variable, these relations define a stochastic difference equation describing the dynamics of the parameter, η, and imply a Fokker-Plank difference equation for its probability

distribution $P(\eta)$.

$$\Pr_{n+1}(\eta) = \int_{-\infty}^{\infty} K(\eta,\eta') \Pr_{n}(\eta') d\eta' \tag{2.16}$$

Where $K(\eta,\eta')$ is the integral kernel defined through Eq. (2.15) as,

$$K(\eta,\eta') = \sum_{r=-1,1} \delta(\eta - (\eta' - gr/s)) \Pr(r|\eta') \tag{2.17}$$

$$= \frac{1}{2}\left[\delta(\eta - (\eta' - g/s))\left(1 + \sin\left(2s(\eta' - \eta_0)\right)\right) + \right. \tag{2.18}$$

$$\left. \delta(\eta - (\eta' + g/s))\left(1 - \sin\left(2s(\eta' - \eta_0)\right)\right)\right] \tag{2.19}$$

To solve for the stationary distribution, we appeal to the central limit theorem and use the ansatz:

$$\Pr_{s}(\eta) = \frac{e^{-(\eta-\eta_0)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}. \tag{2.20}$$

We then solve for the variance, σ^2, through

$$\Pr_{s}(\eta) = \int_{-\infty}^{\infty} K(\eta,\eta') \Pr_{s}(\eta') d\eta' \tag{2.21}$$

$$= \frac{1}{2}\left[(1 + \sin(2s(\eta - \eta_0)))P_s(\eta + g/s) + (1 - \sin(2s(\eta - \eta_0)))P_s(\eta - g/s)\right] \tag{2.22}$$

After some algebra, assuming that the gain parameter, $g < 1$, we find that

$$\langle \eta \rangle = \eta_0, \tag{2.23}$$

$$\left\langle (\eta - \eta_0)^2 \right\rangle = \frac{g}{4s^2}. \tag{2.24}$$

That is, the mean of this process is a well-calibrated device, and the variance can be reduced arbitrarily by increasing the sequence sensitivity and decreasing the gain. Because these results are asymptotic, future work should focus on optimizing the sensitivity and gain schedules to minimize the time required to achieve a given target variance.

Drift in the one-parameter model We now consider a situation where the underlying parameter is no longer static, but drifts under a random walk process with step size, l. The kernel for this process is then:

$$K(\eta,\eta') = \sum_{q,r=-1,1} \delta(\eta - (\eta' - gr/s - lq)) \Pr(q) \Pr(r|\eta') \tag{2.25}$$

Here q is a random variable indicating whether the drift step is to the left or right. We assume an unbiased random walk, for which $\Pr(q) = 1/2$.

The central limit theorem again implies that the stationary distribution will be Gaussian, but with a larger variance due to the underlying parameter drift. Some more algebra gives the stationary variance as:

$$\sigma^2 = \frac{g}{4s^2} + \frac{l^2}{4g} - \frac{2gl}{s} + l^2. \tag{2.26}$$

2.2.3. Multi-parameter, linear tuning and drift control

Now consider a more general set of gate sequences,

$$G_i = \prod_j g_{i,j} \tag{2.27}$$

selected so that $\Pr(r = \pm 1) = 1/2$ if all gates are perfect. For any given gate sequence, we can compute the Jacobian of the sequence with respect to the experimentally adjustable parameters, $\mathbf{c}$:

$$\mathbf{S}_i = \nabla_c G_i \tag{2.28}$$

We call this Jacobian the *sensitivity vector*. Computing the sensitivity vector requires a model for the gate dynamics as functions of the control parameters. We can now use the following proceedure to minimize the error in these parameters.

1. Select and run a gate sequence, G_i
2. Measure the qubit state, r_n
3. Conditional on the state of the qubit, adjust the parameters using the update rule:

$$\mathbf{c}_{n+1} = \mathbf{c}_n + g r_n \mathbf{s}_i / |\mathbf{s}_i|^2 \tag{2.29}$$

Now, assuming that the Jacobians span the space of adjustable parameters, this update rule should converge on the target. The Fokkor-Plank equation describing the probability distribution over the parameters is given by the kernel

$$K(\mathbf{c}, \mathbf{c}') = \sum_{r=\pm 1} \delta(\mathbf{c} - (\mathbf{c}' - g r \mathbf{s}_i / |\mathbf{s}_i|^2)) \Pr(r|\mathbf{c}', \mathbf{s}_i) \tag{2.30}$$

Where $\Pr(r|\mathbf{c}, \mathbf{s}_i) = 1/2(1 + \sin(2\mathbf{s}_i \cdot \mathbf{c}))$, or in the linear response region, $\Pr(r|\mathbf{c}) = \frac{1}{2}(1 + 2\mathbf{s}_i \cdot \mathbf{c}_i)$. We can now average over the sensitivity vectors to get

$$K(\mathbf{c}, \mathbf{c}') = \frac{1}{M} \sum_{r=\pm 1} \sum_{i=1}^{M} \delta(\mathbf{c} - (\mathbf{c}' - g r \mathbf{s}_i / |\mathbf{s}_i|^2)) \Pr(r|\mathbf{c}', \mathbf{s}_i) \tag{2.31}$$

Again, taking the ansatz that the stationary distribution is

$$\Pr(\mathbf{c}) = \mathcal{N}(\mathbf{c}_0, \mathbf{\Sigma}) \tag{2.32}$$

$$= \frac{1}{\sqrt{(2\pi)^N \det(\mathbf{\Sigma})}} \exp\left(-\frac{1}{2}(\mathbf{c} - \mathbf{c}_0)^T \mathbf{\Sigma}^{-1} (\mathbf{c} - \mathbf{c}_0)\right) \tag{2.33}$$

Solving for the stationary variance leads to the following relation between it and and the sensitiviy vectors,

$$\sum_i \mathbf{S}_i \mathrm{Tr}\,(\mathbf{S}_i \cdot \mathbf{\Sigma}) = \sum_j \frac{g}{4 |\mathbf{s}_j|^2} \mathbf{S}_j \tag{2.34}$$

where we have defined $\mathbf{S}_i = \mathbf{s}_i \mathbf{s}_i^T / |\mathbf{s}_i|^2$, the projector onto the sensitivity vector, $\mathbf{s}_i$. This set of projectors spans a subspace, $\mathcal{A} = \text{span}(\{\mathbf{S}_i\})$, of real, symmetric, n-by-n matrices. These projectors further define a frame for $\mathcal{A}$, and we see that the left-hand side of (2.34) is equal to the frame operator's action on $\mathbf{\Sigma}$. Recall the definition of the frame operator:

$$F(Q) = \sum_j \mathbf{S}_i \langle \mathbf{S}_i, Q \rangle \tag{2.35}$$

$$= \sum_j \mathbf{S}_i \text{Tr}\left(\mathbf{S}_i \cdot Q\right) \tag{2.36}$$

The operator F is invertible over $\mathcal{A}$, giving the stationary covariance matrix as:

$$\mathbf{\Sigma} = F^{-1}\left(\sum_j \frac{g}{4\left|\mathbf{s}_j\right|^2} \mathbf{S}_j\right). \tag{2.37}$$

This relation has been demonstrated in simulation. The extension to drifting situations is straightforward, but also somewhat cumbersome.

3. MEMORY DEPENDENCE

3.1. Non-Markovian noise in quantum information processors

The time-dependent characterization and tuning protocols discussed in the previous section account for drift on timescales of entire quantum circuits. In this section, we address non-Markovian behavior with correlation times that are on the timescales of one or more quantum *gates*. In this regime, a quantum gate $\mathcal{G}$ that is applied several times within a single circuit may not act the same way each time. Current process-matrix-based approaches are incompatible with this kind of noise, despite the ubiquity with which this phenomenon occurs in physical systems – this particular non-Markovian behavior can arise from high-frequency electromagnetic noise, coupling to neighboring qubits, serial context dependence, heating, leakage, or any of a number of other physical sources. In this section we detail our approach to modeling, and eventual characterization of, fast non-Markovianity.

3.2. Augmented completely-positive maps

Our approach to constructing simple models for describing non-Markovian noise in QIPs is to begin with a description for Markovian noise, and then to generalize as needed. The most general model for time-stationary Markovian noise in a QIP is one wherein each quantum logic gate (including state preparation and measurement) is described by a completely positive, trace preserving (CPTP) map Gi, where i denotes the specific logic gate. For a quantum system with Hilbert space H_S and dimension d_S , CPTP maps can be represented as $d_S^2 \times d_S^2$ matrices, or equivalently as superoperators on the state space $\mathcal{B}(H_S)$.

A key advantage of using the CPTP map model for describing noisy quantum logic gates, as opposed to using an ensemble of ideal gates distributed over the values of some noise parameter and then simulating trajectories of noise processes using Monte Carlo, is that the CPTP map model for Markovian noise permits probabilities to be calculated exactly without Monte Carlo. It is a time/space tradeoff; more memory is required to store the density matrix of the QIP instead of a pure state, but the simulation time scales only logarithmically with the desired precision (instead of linearly or quadratically as for Monte Carlo).

We construct a simple model for non-Markovian noise, that similarly avoids the trajectories approach, by extending the state space on which the gate matrices act. Our model is the simplest possible: the CPTP map applied to the system is conditioned on a finite-state classical environment, which evolves stochastically following each gate in a Markovian way. That is, although the system dynamics is conditioned on a hidden environment and can therefore be non-Markovian, the environment dynamics itself is Markovian and independent of the system.

Specifically, our model is defined as follows:

1. Each logic gate labelled by i is described by a parameterized CPTP map $G_i(k)$, where $k \in \mathbb{Z}_K$ for some integer K represents the classical state of the K-state environment.

2. All logic gates $\{G_i(k), i = 1,2,\ldots\}$ are conditioned on the same K-state environment.

3. Subsequent to each gate, the environment is updated via a stochastic map J.

(In this representation, at each clock cycle the conditional CP-map is applied before the environment is updated; this could be done the other way around, but it allows for the possibility of starting with the environment in a deterministic state, from which it equilibrates toward its stationary distribution.)

We define such a model to be a discrete-environment augmented CP map model, and gates are now described by augmented completely-positive (ACP) maps. To represent ACP maps compactly, for each gate i there is a ACP map Ω_i defined as

$$\Omega_i = (I \otimes J) \circ \left(\sum_k G_i(k) \otimes |k\rangle\langle k| \right) \tag{3.1}$$

Relative to the standard system-bath approach to modelling non-Markovian behavior, this model is extremely limited. The 'environment' is not only modelled classically, but also by a finite number of states K (and we will be considering cases where K is small, in particular $K = 2$). The environment dynamics is completely Markovian, and is not influenced by the system at all. Finally, all logic gates G_i will be conditioned on the same environment.

3.3. Modeling common Non-Markovian errors

The simplicity of this model belies its power. As we will demonstrate, even the simplest $K = 2$ model is capable of quantitatively describing many features of QIPs in the presence of non-Markovian noise sources. In this section of this report, we provide several examples demonstrating that the simple ACP model we propose can describe a number of realistic non-Markovian noise effects on QIPs.

3.3.1. Low-frequency noise

Low frequency noise can cause correlated errors within the timeframe of a single quantum circuit, but over the course of a full experiment noise may change dramatically. Assume the noise is a low-frequency, two-level fluctuator with states $\pm$ that causes the qubit frequency to change by $\pm\delta$,

respectively. For a gate set that includes I, X, and Y, the corresponding ACP map matrices are (as block matrices):

$$\mathcal{G}_I = \left(\begin{array}{c|c} Z(+\delta) & 0 \\ \hline 0 & Z(-\delta) \end{array}\right) \tag{3.2}$$

$$\mathcal{G}_X = \left(\begin{array}{c|c} X & 0 \\ \hline 0 & X \end{array}\right) \tag{3.3}$$

$$\mathcal{G}_Y = \left(\begin{array}{c|c} Y & 0 \\ \hline 0 & Y \end{array}\right) \tag{3.4}$$

Here we have assumed that only the identity gate is appreciably sensitive to the qubit frequency error. The initial state of the bath should be $(1,1)^{\mathsf{T}}/2$ and the measurement should average equally over the bath states. We can similarly model noise with shorter time correlations. For symmetric two level systems characterized by a transition rate γ, this leads to ACP maps with off-diagonal blocks that are non-zero, such as

$$\mathcal{G}_I = \left(\begin{array}{c|c} (1-\gamma)Z(+\delta) & \gamma Z(-\delta) \\ \hline \gamma Z(+\delta) & (1-\gamma)Z(-\delta) \end{array}\right) \tag{3.5}$$

3.3.2. Serial context dependence

Quantum information processors execute quantum circuits as a series of precisely timed electromagnetic control pulses. Ideally, the pulses implementing a particular quantum operation will stop before the next operation is to begin. In many cases, however, a quantum operation will be negatively impacted by the operation that came before. A common source of such errors is the finite ring-down of microwave cavities, but it can also arise in trapped ion systems or because of overly optimized control pulses. Pulse predistortion can even appear to violate causality, leading to gates performance that depends *following* operation.

The ACP map formalism is well suited to modeling this type of error. As an example, we will consider a single qubit with I, X, and Y gates. The identity will experience additional decoherence if the preceding gate is an X or a Y, but will work well if the preceding gate is another identity. The classical state space can then serve as a memory recording whether the last gate was an identity or not. In block matrix form, the associated matrices are:

$$\mathcal{G}_I = \left(\begin{array}{c|c} I_{\text{good}} & I_{\text{bad}} \\ \hline 0 & 0 \end{array}\right) \tag{3.6}$$

$$\mathcal{G}_X = \left(\begin{array}{c|c} 0 & X \\ \hline X & 0 \end{array}\right) \tag{3.7}$$

$$\mathcal{G}_Y = \left(\begin{array}{c|c} 0 & Y \\ \hline Y & 0 \end{array}\right) \tag{3.8}$$

Here $I_{\text{good/bad}}$ are the process matrices associated to the good identity gate and the bad gate, respectively. The initial state of the bath should be $(1,0)^{\mathsf{T}}$, and, as usual, the measurement should average equally over the final bath states.

3.4. Discussion

ACP map models are attractive, compact, and intuitive representations for memory effects in quantum information processors. We are currently working to incorporate these models in pyGSTi, and to test our ability to fit these models using data from real hardware.

REFERENCES

[1] E. Nielsen, J. K. Gamble, K. Rudinger, T. Scholten, K. Young, and R. Blume-Kohout, "Gate set tomography," Sept. 2020.

[2] R. Blume-Kohout, J. K. Gamble, E. Nielsen, K. Rudinger, J. Mizrahi, K. Fortier, and P. Maunz, "Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography," *Nat. Commun.*, vol. 8, Feb. 2017.

[3] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor," *Nat. Commun.*, vol. 5, p. 4213, July 2014.

[4] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," Nov. 2014.

[5] D. Aharonov, A. Kitaev, and J. Preskill, "Fault-tolerant quantum computation with long-range correlated noise," *Phys. Rev. Lett.*, vol. 96, p. 050504, Feb. 2006.

[6] D. Gottesman, "An introduction to quantum error correction and Fault-Tolerant quantum computation," Apr. 2009.

[7] P. Aliferis and B. M. Terhal, "Fault-Tolerant quantum computation for local leakage faults," Nov. 2005.

[8] T. J. Proctor, P. A. Knott, and J. A. Dunningham, "Multiparameter estimation in networked quantum sensors," *Phys. Rev. Lett.*, vol. 120, p. 080501, Feb. 2018.

[9] T. J. Proctor, P. A. Knott, and J. A. Dunningham, "Networked quantum sensing," Feb. 2017.

[10] T. Proctor, M. Revelle, E. Nielsen, K. Rudinger, D. Lobser, P. Maunz, R. Blume-Kohout, and K. Young, "Detecting, tracking, and eliminating drift in quantum information processors," July 2019.

[11] B. J. Dalton, "Liouville space theory of sequential quantum processes. i. general theory," *J. Phys. A Math. Gen.*, vol. 15, p. 2157, July 1982.

[12] J. P. Dehollain, J. T. Muhonen, R. Blume-Kohout, K. M. Rudinger, J. K. Gamble, E. Nielsen, A. Laucht, S. Simmons, R. Kalra, A. S. Dzurak, and A. Morello, "Optimization of a solid-state electron spin qubit using gate set tomography," *New J. Phys.*, vol. 18, p. 103018, Oct. 2016.

[13] J. Morris, F. A. Pollock, and K. Modi, "Non-Markovian memory in IBMQX4," Feb. 2019.

[14] M. Y. Niu, V. Smelyanskyi, P. Klimov, S. Boixo, R. Barends, J. Kelly, Y. Chen, K. Arya, B. Burkett, D. Bacon, Z. Chen, B. Chiaro, R. Collins, A. Dunsworth, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, T. Huang, E. Jeffrey, D. Landhuis, E. Lucero, A. Megrant, J. Mutus, X. Mi, O. Naaman, M. Neeley, C. Neill, C. Quintana, P. Roushan, J. M. Martinis, and H. Neven, "Learning Non-Markovian quantum noise from Moiré-Enhanced swap spectroscopy with deep evolutionary algorithm," Dec. 2019.

[15] I. A. Luchnikov, S. V. Vintskevich, D. A. Grigoriev, and S. N. Filippov, "Machine learning non-markovian quantum dynamics," Feb. 2019.

[16] R. S. Gupta and M. J. Biercuk, "Machine learning for predictive estimation of qubit dynamics subject to dephasing," *Phys. Rev. Applied*, vol. 9, p. 064042, June 2018.

[17] I. A. Luchnikov, S. V. Vintskevich, D. A. Grigoriev, and S. N. Filippov, "Machine learning of markovian embedding for non-markovian quantum dynamics," *arXiv preprint arXiv:1902.07019*, 2019.

[18] G. A. L. White, C. D. Hill, F. A. Pollock, L. C. L. Hollenberg, and K. Modi, "Experimental non-markovian process characterisation and control on a quantum processor," Apr. 2020.

[19] P. Rebentrost, I. Serban, T. Schulte-Herbrüggen, and F. K. Wilhelm, "Optimal control of a qubit coupled to a non-markovian environment," *Phys. Rev. Lett.*, vol. 102, p. 090401, Mar. 2009.

[20] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, "Non-Markovian quantum processes: Complete framework and efficient characterization," *Phys. Rev. A*, vol. 97, p. 012127, Jan. 2018.

[21] S. Milz, M. S. Kim, F. A. Pollock, and K. Modi, "Completely positive divisibility does not mean markovianity," *Phys. Rev. Lett.*, vol. 123, p. 040401, July 2019.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.