

This is the accepted manuscript made available via CHORUS. The article has been published as:

Recovering noise-free quantum observables

Matthew Otten and Stephen K. Gray

Phys. Rev. A **99**, 012338 — Published 22 January 2019

DOI: [10.1103/PhysRevA.99.012338](https://doi.org/10.1103/PhysRevA.99.012338)

Recovering noise-free quantum observables

Matthew Otten and Stephen K. Gray

Center for Nanoscale Materials, Argonne National Laboratory, Lemont, Illinois, 60439

(Dated: December 20, 2018)

We introduce a technique for recovering noise-free observables in noisy quantum systems by combining the results of many slightly different experiments. Our approach is applicable to a variety of quantum systems but we illustrate it with applications to quantum information and quantum sensing. The approach corresponds to repeating the same quantum evolution many times with known variations on the underlying systems' error properties, e.g. the spontaneous emission and dephasing times, T_1 and T_2 . As opposed to standard quantum error correction methods, which have an overhead in the number of qubits (many physical qubits must be added for each logical qubit) our method has only an overhead in number of evaluations, allowing the overhead to, in principle, be hidden via parallelization. We show that the effective spontaneous emission, T_1 , and dephasing, T_2 , times can be increased using this method in both simulation and experiments on an actual quantum computer. We also show how to correct more complicated entangled states and how Ramsey fringes relevant to quantum sensing can be significantly extended in time.

I. INTRODUCTION

Quantum information science is rapidly evolving due to advances in quantum computing, communication, and sensing. Quantum computing, for example, has potential for exponential speedup in areas such as prime number factoring [1] and quantum chemistry [2], and quantum sensing has potential to be far more sensitive than classical sensing [3]. In most cases, however, decoherence or information loss can impede progress. Quantum error correction can extend quantum information lifetime by encoding a single logical qubit into many physical qubits [4] but introduces space and time overheads owing to additional physical qubits and gate operations. Furthermore, it is unclear how standard quantum error correction could be used in complicated quantum sensors.

Here, we describe a technique for recovering observables from a quantum evolution by repeating the evolution with slightly different noise characteristics and combining those results to obtain an estimate of the noise-free answer without need of additional quantum hardware. Our approach bears similarities with interesting work by Gambetta and coworkers [5, 6] as well as Benjamin and coworkers [7, 8], involving one global noise parameter (based on tunable error properties of gates) and Richardson extrapolation. It represents a multi-dimensional generalization not reliant on Richardson extrapolation that can also use the underlying error properties of quantum systems in cases where there is not a tunable, global noise source and is thus applicable to a wider range of quantum systems. We have recently used such a simple error model to develop a different error correction scheme that requires the ability to significantly reduce error on individual qubits via, e.g., quantum error correction [9]. Our approach here does not rely on quantum error correction, but requires many, slightly different runs to be completed; these can be done in time (repeating the evolution many times in sequence) or in parallel (many separate systems simultaneously undergo evolution). Our approach can be used in quantum algorithms, quantum

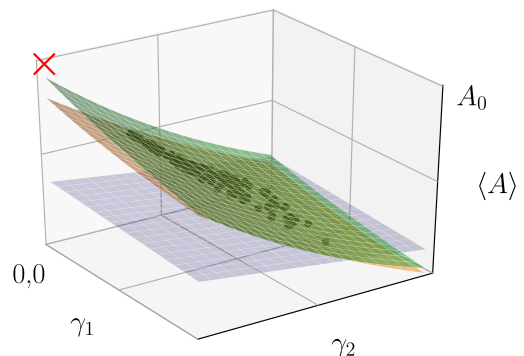


FIG. 1: Example of a ‘hypersurface’ fit to many experiments with slightly different noise parameters, γ_1 and γ_2 . Black points are measurements of an observable with different noise rates. The ‘X’ is the noise-free result. Blue (lower), orange (middle), and green (upper) surfaces are first, third, and fourth order fits, respectively. Many observable measurements are outside the region displayed.

sensing, and general quantum experiments where decoherence times are too short to obtain high-quality signals. The only requirement is that the evolution can be repeated with different, well-characterized noise sources.

II. METHOD

Consider a quantum system with one or more subsystems (e.g., qubits), each undergoing one or more noise processes. Let the set of all noise rates be $\{\gamma_i\}$. This system repeatedly undergoes a given evolution, e.g. a sequence of quantum gates or interaction with a magnetic

field, with varying values of noise rates. Combining all results, we construct a hypersurface embedded in a space where one axis represents the measurement results and the other axes represent the noise rate parameters. The form of the hypersurface is obtained via the Taylor expansion of the quantum system's evolution operator (see Appendix A) and yields an estimate of the noise-free observable, as well as information about the effect of each noise rate.

For example, consider a single qubit with only amplitude damping noise. We repeatedly apply an evolution, each time with differing damping rates. Let $\gamma_1^{[j]}$ be the damping rate for the j th repetition and $\langle A \rangle^{[j]}$ be the corresponding measured observable. For a third order in γ_1 observable model, solving (via, e.g., standard least-squares)

$$\begin{bmatrix} 1 & \gamma_1^{[1]} & (\gamma_1^{[1]})^2 & (\gamma_1^{[1]})^3 \\ 1 & \gamma_1^{[2]} & (\gamma_1^{[2]})^2 & (\gamma_1^{[2]})^3 \\ 1 & \gamma_1^{[3]} & (\gamma_1^{[3]})^2 & (\gamma_1^{[3]})^3 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ A_{11} \\ A_{111} \end{bmatrix} = \begin{bmatrix} \langle A \rangle^{[1]} \\ \langle A \rangle^{[2]} \\ \langle A \rangle^{[3]} \\ \vdots \end{bmatrix}, \quad (1)$$

for $(A_0, A_1, A_{11}, A_{111})$ yields $\langle A \rangle(\gamma_1) \approx A_0 + A_1\gamma_1 + A_{11}\gamma_1^2 + A_{111}\gamma_1^3$. Here the hypersurface is just a cubic curve with intercept A_0 being the desired noise-free value. The formalism extends to higher orders and to many noise parameters, potentially from many qubits. For example, using a single qubit with a spontaneous emission rate, γ_1 and pure dephasing rate, γ_2 , the j th row of our matrix is $[1 \ \gamma_1^{[j]} \ \gamma_2^{[j]} \ (\gamma_1^{[j]})^2 \ \gamma_1^{[j]}\gamma_2^{[j]} \ (\gamma_2^{[j]})^2]$ if truncated to second order. Figure 1 displays hypersurfaces (now 2D surfaces) for a system with two noise parameters. A red 'X' is the noise-free solution. As model order increases, the surfaces better fit the data and extrapolate closer to the noise-free limit.

III. RESULTS

A. Relaxation Time

We first demonstrate the method for a single qubit using a simple relaxation time experiment. We excite the qubit into the $|1\rangle$ state, wait some time, and then measure what state it is in; repeating many times yields the probability of remaining in (or fractional population of) the excited state. Due to the amplitude damping noise, the population will decay to zero with characteristic time $T_1 = 1/\gamma_1$. We first show how to recover the population in simulation, where we select random γ_1 values uniformly in a range representing T_1 times between $5 \mu\text{s}$ and $15 \mu\text{s}$. All simulations are numerical solutions of the Lindblad master equation [10, 11] and utilize the high-performance open quantum systems solver Quac [12]. The results are shown in Fig. 2a, where we present best, worst, and average evolutions over 450 repetitions, and recovered

populations using Eq. (1), up to tenth order. The procedure is applied at specific times, with knowledge only of the measured observables at that time. By recovering at many different times t , we obtain the full evolution. Every order shown is better than the best run and increasing order increases recovery quality. For example, at $60 \mu\text{s}$ the average population is 0.0045, i.e. the state is almost all decayed away. First order recovery gives a population of only 0.017. Tenth order recovery gives a population of 0.90, nearly the noise-free result. If $T_1(n)$ is the time the recovered evolution has population $1/e$ at order n , we see $T_1(n) \approx (n+1)T_1(n=0)$.

We also perform the relaxation time experiment on Rigetti's eight qubit chip, Agave [13, 14], a superconducting qubit quantum computer with a ring topology. Each of the eight qubits has slightly different T_1 and T_2^* times, all approximately $10 \mu\text{s}$. Furthermore, these noise characteristics drift in time [6, 15]. These features provide the necessary variation in noise parameters for our method. We first excite a single qubit using a Pauli-X gate, wait some time, and measure the qubit state. This is repeated for many different wait times. Each experiment at a given wait time is averaged over 10^5 shots, giving an average population. The T_1 time and associated decay rate, γ_1 , is extracted by fitting an exponential to the data, weighted by the standard error of the mean of each measurement. This process is repeated for each qubit in the quantum computer, a few minutes are allowed to pass, and then the full cycle is repeated, starting from the first qubit. This generates many different repetitions with varying noise parameters. The fitted γ_1 parameters are used for the recovery. To obtain the error bars for the experimental data, the least-squares fitting procedure was replaced with a weighted least-squares fitting procedure, taking into account the standard error of the mean in the determination of the populations at each separate wait time. This process does not, however, take into account the error on the fitted γ_1 values. In principle, this could also be included by using total least-squares or orthogonal distance regression [16], but this is a more involved procedure.

Just as in simulation, each wait time is recovered separately. A total of 45 different single qubit repetitions are run with results shown in Fig. 2b. Both first and second order recoveries result in a higher population than the average, similar to the simulations. The recovery is limited by noise sources not included in the model, such as readout noise. Neither first nor second order recovery reaches unity; instead, achieving ≈ 0.85 , roughly the average readout fidelity (0.84) of the qubits used in the experiment. The inset of Fig. 2b also shows a comparison between our Lindblad model for the qubit and the experimental data. We ran 45 different repetitions using experimentally determined T_1 times and shifted the data down by 0.18 to account for the limited readout fidelity in the experimental runs. The model average and experimental averages line up very well; the 1st order recovery of the model and experimental data only agree

in the short times. Since our recovery method can only recover from noise sources that we experimentally determine, this implies that there could be other, important noise sources that were not included in the recovery. These could include state preparation and measurement errors, which were only roughly included by a shift of the data. Other sources of error could include crosstalk [13] or $1/f$ noise [17], which is generally non-Markovian in nature and would not have been included in our simple noise model. The difference between the model and the experimental data implies that the simple T_1 , T_2 noise model is not entirely sufficient for describing the noise of these qubits. Our recovery method allows one to include more noise sources, as long as they are quantifiable, and can help identify which noise sources have weak or strong effects on the resulting observable. In our experiments, different qubits from the same device are used at different times, minimizing potential differences between each repetition. Combining repetitions from disparate architectures, e.g., ion-trap and superconducting qubits, would require non-trivial additions to the model.

As higher orders are used in the recovery, the error bars also get larger; arbitrary order cannot be used with a limited number of measurements. Sampling error ultimately limits the highest order that can be used in the recovery procedure. This is studied in detail in Ref. 8, where it is shown that higher orders in a single parameter fit moves the mean but also increases the variance. Rather than using a new repetition to achieve a higher order fit, that data could have been used to reduce the size of the error bars in a lower order fit.

With only one noise parameter the method is similar to Richardson extrapolation, which is able to extrapolate to the zero-noise limit in superconducting qubits [5, 6]; in this case the methods differ only in choice of points and fitting strategy, but the source of the variation in noise differs greatly. The Richardson extrapolation technique assumes a single global noise source, which has been implemented by scaling the length of pulses while running the quantum algorithm on the same set of qubits [6] or otherwise increasing the noise [7]. The hypersurface method allows for any number of noise sources and is a multi-dimensional generalization utilizing natural variations in qubit properties. To show this, we use Ramsey interferometry with no background magnetic field, a common technique for measuring T_2 times [18]. This involves applying a $\pi/2$ X rotation to $|0\rangle$, waiting some time, and then applying another $\pi/2$ X rotation. Without noise and in a rotating frame, the final state would be $|1\rangle$. As opposed to the relaxation time experiment, both amplitude damping and dephasing noise affect the result. Figure 3a shows simulated results of this experiment with recovery up to 8th order. Spontaneous emission T_1 and pure dephasing T_2^* times are independently, randomly chosen between $5 \mu\text{s}$ and $15 \mu\text{s}$ with 450 simulations being run. Without recovery, the excited state population associated with a given $\gamma_1 = 1/T_1$ and $\gamma_2 = 1/T_2^*$ exponentially decays with rate $1/T_2 = \gamma_1/2 + \gamma_2$. As with the

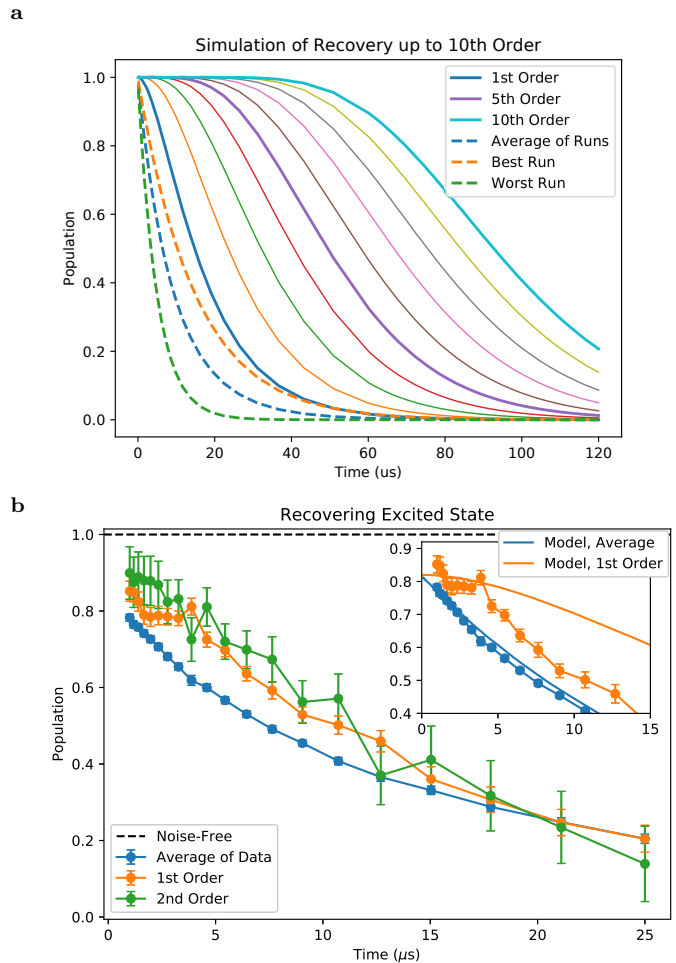


FIG. 2: Population recovery in a relaxation time experiment. (a) Simulated data from 450 simulations with random T_1 times. (b) Experimental results on Rigetti's 8 qubit quantum computer [13]. 45 repetitions are made at each time with experimentally determined T_1 times. The inset shows the comparison between our noise model run with the experimentally determined T_1 times and the experimental data.

relaxation time experiments, first order recovery yields a better evolution for most points compared with even the best run of all the qubits. As ever higher orders are considered, a unity excited state population is recovered for longer periods of time.

B. Dephasing Time

We also carry out Ramsey experiments on Rigetti's eight qubit Agave quantum computer. To obtain correct noise rates, γ_i , for a given repetition we first characterize T_1 using the relaxation time experiment discussed above. We then perform a Ramsey interferometry experiment as previously described. The results of this experiment are

fit to an exponential to obtain T_2 . This allows determination of the pure dephasing rate via $\gamma_2 = 1/T_2 - 1/(2T_1)$. Each wait time in both the T_1 and T_2 determinations is averaged over 10^5 shots. With both rates determined, the hypersurface method, Eq. (1), is used to recover the excited state population at each wait time, using 12 repetitions (Fig. 3b). The first order correction clearly recovers a fraction of the missing population and approaches the readout fidelity limit at early times.

Characterization of the noise rates and a good understanding of the noise sources are imperative for our method. For example, because of how γ_2 is determined, poor determinations of T_1 and T_2 can sometimes lead to negative γ_2 ; such unphysical repetitions are excluded from the recovery analysis. Furthermore, each observable measurement $\langle A \rangle^{[j]}$ is assumed to be measured at a set of determined noise rates $\gamma^{[j]}$. In superconducting qubit systems, noise characteristics fluctuate over time scales in the few minute range [6]. The total time to determine $\langle A \rangle^{[j]}$ (which involves the average over shots) must be significantly smaller than this time scale. In this work, determination of T_1 and T_2 takes ≈ 1 s, which is less than the time scale of the fluctuations. Given that the time scale of the longest gate on the Rigetti Agave quantum computer [13] is on the order of 100 ns (and including the time to reinitialize the quantum computer after a measurement is made), the total gate depth for more complicated algorithms will be limited and T_1 and T_2 would need to be characterized before each repetition to ensure that the hypersurface equations have the correct noise sources. As qubit quality increases, we expect the noise rates will become both much lower and much more stable, allowing for longer circuits to be used in this method. In both experiments on the Rigetti quantum computer, the correct unity excited state population could not be recovered. Instead, the method recovered the readout fidelity limit.

The Ramsey experiments show the strength of this method, compared to Richardson extrapolation with a global noise parameter. These latter methods [5, 7] make use of gate noise tuning allowing them to be used very efficiently in gate-based quantum computation. However, in quantum systems with no ‘gates’ (e.g. quantum sensors, quantum memories, and general quantum experiments) or if increasing all noise sources (e.g., T_1 and T_2^*) at the same rate is not easy, our approach can still be used. In a situation where measurement of a single pulse is important, as is the case for many sensing applications, being able to use ‘space’, by having more sensors participating in the time-sensitive sensing, can increase sensitivity dramatically. In other situations, such as measuring a background field, being able to use either space or time offers flexibility in design of high-sensitivity devices.

By utilizing a single global noise source, Richardson extrapolation requires only a few evaluations, even for thousands of qubits. Furthermore, using gate noise, rather than the underlying physical qubit noise, the methods of refs. [5, 7] can be effectively applied to near-term,

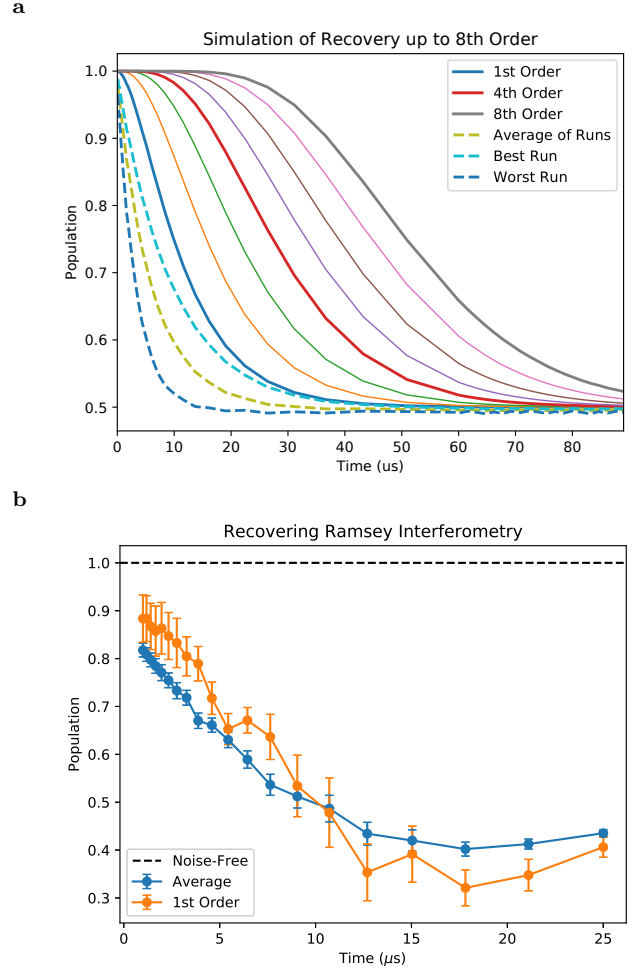


FIG. 3: Population recovery in Ramsey interferometry with no background magnetic field. (a) Simulated data from 450 simulations with random T_1 and T_2^* times. (b) Experimental results on Rigetti’s 8 qubit quantum computer [13]; 12 repetitions are included at each time, with experimentally determined T_1 and T_2^* times.

noisy quantum computers [6]. The hypersurface method, in contrast, requires more evaluations as the number of quantum subsystems increases. The resource scaling is defined by the truncation order of the Taylor series. The number of unknowns is $\approx m^l$, where m is the number of noise sources and l is the truncation order (see Appendix B for exact result). Even with this polynomial scaling, the number of unknowns can still become large with high order l or a large number of noise sources m . In this case, an under-determined system, where the number of repetitions is smaller than the number of unknowns, could be solved using numerical techniques such as regularization [19], where we would introduce constraints on the fit based on our knowledge of the physics of the error channels. Nonetheless for systems where scaling a global noise source is infeasible, the hypersurface method offers error mitigation with tunable resource cost and ef-

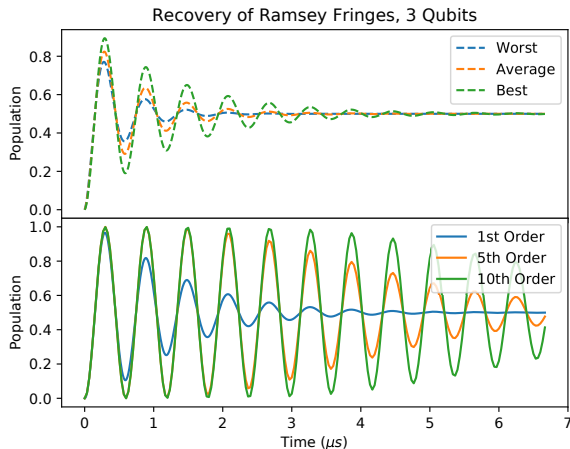


FIG. 4: Ramsey fringe recovery from using 350 simulations of three entangled qubits with random T_2^* times.

fectiveness. The multiparameter Taylor expansion shown in Appendix A gives some insight into the bound on the remaining error in the observable after applying our method; the remaining error, for truncation order l , is of the order of the sum of all products of all error terms, γ_i , such that there are $l + 1$ different γ terms in each product. Ref. 5 derives a formal bound in the single parameter case; a multiparameter generalization of this would provide a more stringent bound for our method.

C. Ramsey Fringes

As a final example, we simulate Ramsey interferometry with a background magnetic field, a common technique used in quantum sensing of magnetic fields [3]. Both single qubit and entangled arrays of qubits can be used as sensors [20]. The system is first prepared in a superposition state such as $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ for a single qubit and a GHZ state, $\frac{1}{\sqrt{2}}(|00\dots 0\rangle + |11\dots 1\rangle)$, for an entangled array of qubits [3]. Single qubit superposition states can be prepared with a Hadamard gate; GHZ states can be prepared with a Hadamard gate and a sequence of CNOT gates. The system then evolves in the presence of a background magnetic field, causing it to pick up a phase. The inverse entangling operation is then applied, transferring the phase onto a single qubit, which is then measured. The phase accumulated over the course of the interaction depends on magnetic field strength and interaction time. If the interaction time is much longer than the coherence time, the useful information decays away and the Ramsey fringes will not be visible. We simulate this experiment using both a single qubit and three entangled qubits. For each qubit, we select a random T_2^* times in the range of $0.5\mu\text{s}$ to $1.5\mu\text{s}$, consistent with parameters for nitrogen-vacancy centers [21]. The

characteristic T_1 time is large enough to be ignored. We set the background magnetic field to $10\mu\text{T}$ and use a total of 350 samples for both single qubit and entangled qubit simulations. Figure 4 shows the recovery of Ramsey fringes from a three qubit GHZ state; the single qubit case is plotted in Appendix C. In the three qubit case we now have three noise parameters, one for each qubit. The higher order recovery involves many cross terms, and the resulting hypersurface is three-dimensional. Even in this maximally entangled state, Ramsey fringes are still recovered long after most of the individual fringes have decayed away.

IV. CONCLUSION

The hypersurface method can be used to recover from any incoherent errors. In this paper, we have demonstrated recovery from both amplitude damping and pure dephasing. Our method could also be applied to any other incoherent error which can be parameterized by a simple noise rate which is zero (or small) for the noise free case. This includes multi-qubit noise and incoherent noise from application of quantum gates. Determination of these more complex noise rates at the time an algorithm is run is necessary; it is still an open question how these might be measured efficiently. Coherent errors, such as over-rotation during the application of a gate, cannot be clearly translated into the hypersurface method, as they are not describable by a simple noise rate.

We presented a method to recover arbitrary quantum observables by repeatedly measuring the observables with differing noise rates and fitting a hypersurface to the repetitions. Including more and more repetitions and increasing hypersurface order, an increasingly good approximation of noise-free observables of a general quantum system can be recovered. For many quantum systems, T_1 and T_2^* noise are dominant noise sources and there are many techniques for characterizing them. Our method recovers a good approximation of the noise-free evolution in these cases. As shown in this Letter, this method has applications in quantum computing and quantum sensing. Further study on ways to minimize the number of repetitions needed for higher-order recovery of a large number of coupled quantum systems is necessary to control the overhead of the method.

ACKNOWLEDGMENTS

This work was performed at the Center for Nanoscale Materials, a U.S. Department of Energy Office of Science User Facility, and supported by the U.S. Department of Energy, Office of Science, under Contract No. DE-AC02-06CH11357. We gratefully acknowledge the computing resources provided on Bebop, a high-performance computing cluster operated by the Laboratory Computing

Resource Center at Argonne National Laboratory. We thank Rigetti Computing for use of the Agave quantum computer, and Guenevere Prawiroatmodjo for help in setting up the experiments on Agave.

Appendix A: Derivation

The Lindblad master equation for a general density matrix $\rho(t)$ evolved from an initial state ρ_0 is defined as (throughout, $\hbar = 1$)

$$\frac{d\rho}{dt} = -i[H, \rho_0] + L(\rho_0), \quad (\text{A1})$$

where H is the Hamiltonian of the system, describing coherent evolution and L is the Lindblad superoperator, describing incoherent evolution, such as noise processes. Both H and L can, generally, be time dependent. ‘Vectorizing’ the density matrix [22], ρ , allows us to write a general solution of the Lindblad master equation as

$$\tilde{\rho}(t) = \exp(-i\tilde{H}t + \tilde{L}t)\tilde{\rho}_0, \quad (\text{A2})$$

where $\tilde{\rho}$ is the vectorized density matrix, \tilde{H} is the vectorized Hamiltonian, and \tilde{L} is the vectorized Lindblad superoperator. See Ref. 22 for more details about the process of vectorization. Let $\tilde{U}(t) = \exp(-i\tilde{H}t + \tilde{L}t)$, the superoperator propagator. We can decompose the superoperator propagator in the solution, equation (A2), using the Trotter decomposition [23]:

$$\tilde{U}(t) = \lim_{n \rightarrow \infty} \left(\left(\exp(-i\tilde{H}\frac{t}{n}) \exp(\tilde{L}\frac{t}{n}) \right)^n \right) \quad (\text{A3})$$

$$\begin{aligned} \tilde{U}(t) &\approx \left(U(1 + \gamma_1 V_1 + \gamma_2 V_2) \right)^3 \\ &\approx UUU + \gamma_1(UV_1UU + UUV_1U + UUV_1U) + \gamma_2(UV_2UU + UUV_2U + UUV_2U) + \\ &\quad + \gamma_1^2(UV_1UV_1U + UUV_1UV_1 + UV_1UUV_1) + \gamma_2^2(UV_2UV_2U + UUV_2UV_2 + UV_2UUV_2) \\ &\quad + \gamma_1\gamma_2(UV_1UV_2U + UUV_1UV_2 + UV_1UUV_2 + UV_2UV_1U + UUV_2UV_1 + UV_2UUV_1). \end{aligned} \quad (\text{A6})$$

The generalization to higher order Trotterizations is clear; the expanded sum will have many more terms (due to each term having a smaller timestep, $\frac{t}{n}$), but terms can be grouped by their γ_j prefactors. For higher order Taylor expansions of $\exp \tilde{L}$, terms can still be grouped by their γ_j prefactors. For example, take the γ_1^2 term from equation (A6). With a second order Taylor expansion, the γ_1^2 terms now contain contributions from V_i^2 :

$$\begin{aligned} [\gamma_1^2 \text{ terms}] &= UV_1UV_1U + UUV_1UV_1 \\ &\quad + UV_1UUV_1 + UV_1^2UU \\ &\quad + UUV_1^2U + UUV_1^2U \end{aligned} \quad (\text{A7})$$

and take the Taylor expansion of the superoperator propagator of the Lindblad

$$\tilde{U}(t) = \lim_{n \rightarrow \infty} \left(\left(\exp(-i\tilde{H}\frac{t}{n}) \sum_{m=0}^{\infty} \frac{(\tilde{L}\frac{t}{n})^m}{m!} \right)^n \right). \quad (\text{A4})$$

We now write the Lindblad superoperator as a sum of many different Lindblad superoperators, each with its own rate: $\tilde{L} = \sum_j \gamma_j \tilde{L}_j$ and plug this into equation (A4), giving

$$\tilde{U}(t) = \lim_{n \rightarrow \infty} \left(\left(\exp(-i\tilde{H}\frac{t}{n}) \sum_{m=0}^{\infty} \frac{(\sum_j \gamma_j \tilde{L}_j \frac{t}{n})^m}{m!} \right)^n \right). \quad (\text{A5})$$

Equation (A5) an infinite sum over m , a finite sum over i , and is raised to an infinite power, n . Though this equation has infinite terms, we can collect all terms that have the same γ prefactors. For example, the collection of all terms with only γ_j will include all terms from the product that have one first order element from the Taylor series expansion of \tilde{L} . Terms with γ_j^2 will include products with two first order elements, as well as products with one second order element. To provide a concrete example, we truncate the Trotterization at third order, the Taylor expansion at first order, and include two noise terms. Let $U = \exp(-i\tilde{H}\frac{t}{3})$ and $V_j = \tilde{L}_j \frac{t}{3}$. With our truncation, we rewrite equation (A5) as

Combining the generalizations to both higher order Trotterization and higher order Taylor expansion is relatively straightforward; the number of terms grows precipitously, but they can always be gathered by their γ_j prefactors. Collecting all the terms for both the infinite limits of Trotterization and the Taylor expansion leads

to

$$\begin{aligned} \tilde{U}(t) = & \lim_{n \rightarrow \infty} \left(\left(\exp(-i\tilde{H} \frac{t}{n}) \right)^n \right) \\ & + \sum_j \gamma_j [\gamma_j \text{ terms}] \\ & + \sum_j \sum_k \gamma_j \gamma_k [\gamma_j \gamma_k \text{ terms}] + \dots, \end{aligned} \quad (\text{A8})$$

where we have used $[\gamma_j \text{ terms}]$ to represent the (infinite) collection of terms all with γ_j as a prefactor. This represents the general (exact) evolution operator for our system. Our density matrix at time t can now be written

$$\begin{aligned} \tilde{\rho}(t) = & \lim_{n \rightarrow \infty} \left(\left(\exp(-i\tilde{H} \frac{t}{n}) \right)^n \right) \tilde{\rho}_0 \\ & + \sum_j \gamma_j [\gamma_j \text{ terms}] \tilde{\rho}_0 \\ & + \sum_j \sum_k \gamma_j \gamma_k [\gamma_j \gamma_k \text{ terms}] \tilde{\rho}_0 + \dots. \end{aligned} \quad (\text{A9})$$

The first term of this expansion represents the noise-free result. Other terms represent the effects of noise on the evolution. Since this method directly corrects the density matrix, ρ , it follows that it also corrects an arbitrary observable,

$$\langle A \rangle = A_0 + \sum_j \gamma_j A_j + \sum_j \sum_k \gamma_j \gamma_k A_{jk} + \dots, \quad (\text{A10})$$

where A_0 is the noise-free observable value and A_j is the effect of noise rate j on the observable. We define the ‘order’ of the combined Trotterization and Taylor expansion by the number of γ_j terms included. The first order terms, for example, are those with a single γ_j prefactor, and include all possible ways that one (infinitesimal) error evolution can be included. The second order terms include all possible ways that two (infinitesimal) error evolutions can be included, and so on. We do not *a priori* know what the values of the effects of noise on the observable (such as A_j) for any order are; however, we can characterize γ_j for a given experiment. By taking a sequence of experiments, varying γ_j , we can reconstruct the unknown evolution terms by fitting a hypersurface to the points. The coefficients of the hypersurface represent the effects (or, for the zeroth order term, the lack of effects) of noise to a given order on the density matrix (or an arbitrary observable). Equation (A10) generally has effects up the infinite order; to make it tractable, we truncate at some given order. As more orders are included, the fit becomes more accurate, and, therefore, a better approximation of the noise-free result is obtained.

Appendix B: Number of Terms in Expansion

Naively, the number of terms in a given order l would be m^l , where m is the number of noise terms (which

could be the number of qubits or a small factor times the number of qubits). Since γ_j is a scalar, all γ_j will commute, allowing us to fuse terms with the same set of γ . For instance, given a second order expansion with two noise terms, we would general have terms with prefactors $\gamma_1 \gamma_1, \gamma_1 \gamma_2, \gamma_2 \gamma_1$, and $\gamma_2 \gamma_2$, but since $\gamma_1 \gamma_2 = \gamma_2 \gamma_1$, we can reduce the number of fitted parameters by combining the $\tilde{L}_1 \tilde{L}_2$ and $\tilde{L}_2 \tilde{L}_1$ terms. For a given order l and number of noise terms m , the number of parameters n for that order is the number of multinomial coefficients, which is given by the formula

$$n = \binom{l+m-1}{m-1}. \quad (\text{B1})$$

For an expansion truncated at order l , the total number of parameters, for all orders, is the sum of equation (B1) for each order up to, and including, l . To provide an explicit example of the polynomial scaling, we will take the example of four qubits with one noise term each ($m = 4$) and truncate at first order ($l = 1$). In this case, we have to sum the terms from order 0 and 1, giving $n = \binom{3}{3} + \binom{4}{3} = 5 \approx 4^1$ unknowns. Following equation (1) in the main text, and using the same number of repetitions as unknowns, our hypersurface equations become

$$\begin{bmatrix} 1 & \gamma_1^{[1]} & \gamma_2^{[1]} & \gamma_3^{[1]} & \gamma_4^{[1]} \\ 1 & \gamma_1^{[2]} & \gamma_2^{[2]} & \gamma_3^{[2]} & \gamma_4^{[2]} \\ 1 & \gamma_1^{[3]} & \gamma_2^{[3]} & \gamma_3^{[3]} & \gamma_4^{[3]} \\ 1 & \gamma_1^{[4]} & \gamma_2^{[4]} & \gamma_3^{[4]} & \gamma_4^{[4]} \\ 1 & \gamma_1^{[5]} & \gamma_2^{[5]} & \gamma_3^{[5]} & \gamma_4^{[5]} \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = \begin{bmatrix} \langle A \rangle^{[1]} \\ \langle A \rangle^{[2]} \\ \langle A \rangle^{[3]} \\ \langle A \rangle^{[4]} \\ \langle A \rangle^{[5]} \end{bmatrix}. \quad (\text{B2})$$

As long as the set of noise parameters for each repetition are different, any values can be chosen. One specific choice is to have all of the noise parameters take some small value γ_i^L and to increase the noise parameter of each qubit independently to some high value γ_i^H . In this case, the hypersurface equations become

$$\begin{bmatrix} 1 & \gamma_1^L & \gamma_2^L & \gamma_3^L & \gamma_4^L \\ 1 & \gamma_1^H & \gamma_2^L & \gamma_3^L & \gamma_4^L \\ 1 & \gamma_1^L & \gamma_2^H & \gamma_3^L & \gamma_4^L \\ 1 & \gamma_1^L & \gamma_2^L & \gamma_3^H & \gamma_4^L \\ 1 & \gamma_1^L & \gamma_2^L & \gamma_3^L & \gamma_4^H \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = \begin{bmatrix} \langle A \rangle^{[1]} \\ \langle A \rangle^{[2]} \\ \langle A \rangle^{[3]} \\ \langle A \rangle^{[4]} \\ \langle A \rangle^{[5]} \end{bmatrix}. \quad (\text{B3})$$

This case, of independently changing each qubits’ noise properties one at a time, represents an independent linear extrapolation along each dimension of the hyperspace. The hypersurface equations are much more flexible than this, however. If all of the noise parameters changed between runs, the solution would still be able to fit linear extrapolations along each dimension from the simultaneously changed data. This could be useful for solving the system in an underdetermined fashion, where the number of repetitions is smaller than the number of unknowns.

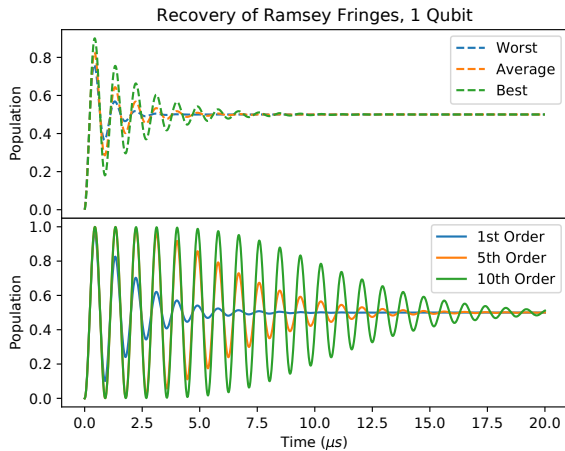


FIG. 5: Single qubit Ramsey interferometry. 350 simulations with random T_2^* times are used.

Appendix C: Recovery of Single Qubit Ramsey Fringes

Figure 5 shows the recovery of Ramsey fringes for a single qubit. In this single qubit case, there is only one noise parameter, γ_2 , as $\gamma_1 \ll \gamma_2$, in contrast to the Ramsey interferometry in the superconducting qubit system, where $\gamma_1 \approx \gamma_2$. We set the background magnetic field to 10 μT and combine the results of 350 experiments. As the order of the recovery is increased, more and more fringes are recovered; at tenth order, the fringes extend are recovered even when the best single qubit run has no clearly visible fringes. Without correction, the fringes decay with the characteristic T_2 time [18].

-
- [1] P. W. Shor, SIAM review **41**, 303 (1999).
 - [2] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, *et al.*, Nature chemistry **2**, 106 (2010).
 - [3] C. L. Degen, F. Reinhard, and P. Cappellaro, Reviews of modern physics **89**, 035002 (2017).
 - [4] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge university press, 2010).
 - [5] K. Temme, S. Bravyi, and J. M. Gambetta, Physical review letters **119**, 180509 (2017).
 - [6] A. Kandala, K. Temme, A. D. Corcoles, A. Mezza-capo, J. M. Chow, and J. M. Gambetta, arXiv preprint arXiv:1805.04492 (2018).
 - [7] Y. Li and S. C. Benjamin, Physical Review X **7**, 021050 (2017).
 - [8] S. Endo, S. C. Benjamin, and Y. Li, Physical Review X **8**, 031027 (2018).
 - [9] M. Otten and S. K. Gray, arXiv preprint arXiv:1804.06969 (2018).
 - [10] M. Otten, R. A. Shah, N. F. Scherer, M. Min, M. Pelton, and S. K. Gray, Physical Review B **92**, 125432 (2015).
 - [11] M. Otten, J. Larson, M. Min, S. M. Wild, M. Pelton, and S. K. Gray, Physical Review A **94**, 022312 (2016).
 - [12] M. Otten, “QuaC: Open quantum systems in C, a time-dependent open quantum systems solver,” <https://github.com/Ott3r/QuaC> (2017).
 - [13] M. Reagor, C. B. Osborn, N. Tezak, A. Staley, G. Prawiroatmodjo, M. Scheer, N. Alidoust, E. A. Sete, N. Didier, M. P. da Silva, *et al.*, Science advances **4**, 3603 (2018).
 - [14] R. S. Smith, M. J. Curtis, and W. J. Zeng, arXiv preprint arXiv:1608.03355 (2016).
 - [15] C. Müller, J. Lisenfeld, A. Shnirman, and S. Poletto, Physical Review B **92**, 035442 (2015).
 - [16] P. T. Boggs and J. E. Rogers, Contemporary Mathematics **112**, 183 (1990).
 - [17] F. Yoshihara, K. Harrabi, A. Niskanen, Y. Nakamura, and J. Tsai, Physical review letters **97**, 167001 (2006).
 - [18] H. Paik, D. Schuster, L. S. Bishop, G. Kirchmair, G. Catelani, A. Sears, B. Johnson, M. Reagor, L. Frunzio, L. Glazman, *et al.*, Physical Review Letters **107**, 240501 (2011).
 - [19] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Vol. 375 (Springer Science & Business Media, 1996).
 - [20] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Physical Review Letters **79**, 3865 (1997).
 - [21] F. Jelezko, T. Gaebel, I. Popa, M. Domhan, A. Gruber, and J. Wrachtrup, Physical Review Letters **93**, 130501 (2004).
 - [22] R. Uzdin, A. Levy, and R. Kosloff, Physical Review X **5**, 031044 (2015).
 - [23] H. F. Trotter, Proceedings of the American Mathematical Society **10**, 545 (1959).