Dynamical uncertainty propagation with noisy quantum parameters

Mogens Dalgaard,1 Carrie A. Weidner,1 and Felix Motzoi2*

1Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, 8000 Arhus C, Denmark
2Forschungszentrum Jülich, Institute of Quantum Control (PGI-8), D-52425 Jülich, Germany

(Dated: July 27, 2021)

Many quantum technologies rely on high-precision dynamics, which raises the question of how these are influenced by the experimental uncertainties that are always present in real-life settings. A standard approach in the literature to assess this is Monte Carlo sampling, which suffers from two major drawbacks. First, it is computationally expensive. Second, it does not reveal the effect that each individual uncertainty parameter has on the state of the system. In this work, we evade both these drawbacks by incorporating propagation of uncertainty directly into simulations of quantum dynamics, thereby obtaining a method that is faster than Monte Carlo simulations and directly provides information on how each uncertainty parameter influence the system dynamics. Additionally, we compare our method to experimental results obtained using the IBM quantum computers.

Technologies based on quantum information may lead to groundbreaking progress within numerical search [1], cryptography [2], simulation [3, 4], optimization [5, 6], and machine learning [7, 8]. Realization of these technologies requires improvements in our ability to measure, design, build, and realistically model subparts of these quantum systems [9]. The latter necessitates the incorporation of the experimental uncertainties that are always present in real quantum systems into numerical simulations of such systems.

Propagation of uncertainty (or error) is a standard tool for understanding how experimental uncertainties transform in calculations [10]. In this work we incorporate propagation of uncertainty into quantum dynamical simulation and characterization by drawing on methods from quantum control theory [11–14]. In doing so we demonstrate a significant speed up relative to Monte Carlo sampling, which is the current standard in the literature (see, e.g., Refs. [15–28] for a series of recent papers).

Moreover, our method allows for an in-depth analysis of how experimental uncertainties independently or cooperatively influence the dynamical observables in the system. This is valuable for investigating the consistency between models and experimental results, as well as the characterization of noise, uncertainty levels, and the influence of different experimental parameters. As one example, within circuit QED we have, in recent years, seen improvements in one- and two-qubit gate operations due to better understanding and calibration of the experimental systems [29–33]. Careful study of these systems’ uncertainties and their influence on quantum dynamics is necessary to push this frontier even further.

The methods proposed in this paper could also help in the construction of composite control sequences that are specifically designed to perform robustly against fluctuations in the system parameters [34–37], as well as enhancing quantum control and optimization protocols [38, 39, 40], 41]. That is, the sensitivity to parameter uncertainties can be modified as needed for tomographic reconstructions [42, 43], system identification [44, 45], and quantum metrology [46, 47]. In addition, the presented method could also aid physicists in designing quantum computation architectures [52, 57]. These need to perform stably under smaller fluctuations in the system parameters, and as such, a detailed analysis of their performance under realistic and expected uncertainties could help assess the robustness of a given architecture.

We term our approach Quantum Uncertainty Propagation In Dynamics (QUPID).

This paper is organized as follows: in Section I we introduce QUPID, in Section II we compare to Monte Carlo simulations, and in Section III we compare to experimental results. Further, in Section IV we investigate the scaling of QUPID and finally, in Section V we conclude the paper.

I. THEORY

Let $Z(t)$ denote a function that depends on the evolution of the system, such as the expectation value of a Hermitian operator $Z(t) = \langle Z \rangle$, or the Haar-averaged gate fidelity $Z(t) = \int d\varphi \langle \psi | U(t) U_{\text{target}} | \psi \rangle$. Now let the Hamiltonian operator and other system evolution variables depend on a set of normally distributed parameters $\theta = (\theta_1, \theta_2, \ldots, \theta_M)$ with mean $\bar{\theta} = (\bar{\theta}_1, \bar{\theta}_2, \ldots, \bar{\theta}_M)$ and covariances $\sigma_{ij} = E[(\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j)]$, which we assume are relatively small [58]. Here $\sigma_{ij}^2 = \sigma_{ij}$ denotes the variance. Thus, $Z$ is a function of the parameters $\theta$, which we may Taylor expand around the mean values $\bar{\theta}$,

$$Z(\theta, t) \approx Z(\bar{\theta}, t) + \sum_j \frac{\partial Z(\theta, t)}{\partial \theta_j} (\theta_j - \bar{\theta}_j)$$

$$+ \frac{1}{2} \sum_{i,j} \frac{\partial^2 Z(\theta, t)}{\partial \theta_i \partial \theta_j} (\theta_i - \bar{\theta}_i)(\theta_j - \bar{\theta}_j),$$  

where truncating the Taylor expansion to second order is justified by assuming small covariances. Evaluating the mean value reveals
In order to be able to calculate the dynamical uncertainties in system observables with the above equations, it is necessary to be able to efficiently propagate the equations of motion for the derivatives of these observables. To do this, we draw on established methods from the field of quantum optimal control theory [11–14]. We have so far tacitly assumed that \( Z \) was a functional of the current state of the system \( \chi(t) \), which solves the dynamic equation of motion \( \frac{\partial}{\partial t} \chi(t) = A(t) \chi(t) \). This could be, e.g., the Schrödinger equation \( A(t) = -i \hat{H}(t) \) with \( \hbar = 1 \) and \( \chi \) being a single quantum state, a unitary operator, or a collection of states, or a density matrix with evolution governed by the von Neumann or the Lindblad master equation. We elaborate further on this in the Supplemental Material.

We discretize the time evolution for a total duration \( T \) in \( N \) equidistant steps \( \Delta t \) such that \( \Delta t = T/N \). This allows us to numerically solve the equation of motion using the time evolution operator \( \chi_{n+1} = U_n \chi_n \), where \( \chi(t_n) = \chi_n \) and \( U_n = e^{A(t_n+\Delta t/2)} \) is the mid-point interpolation of the truncated Magnus series [29], leading to a global second order-error in \( \chi(T) \) that can be made arbitrarily small by suitable choice of \( \Delta t \). The derivatives of \( Z \) ultimately depend on the derivative of \( \chi \), which we may obtain by using the chain rule

\[
\frac{\partial}{\partial \theta_j} \chi_{n+1} = \left( \frac{\partial}{\partial \theta_j} U_n \right) \chi_n + U_n \frac{\partial}{\partial \theta_j} \chi_n, \tag{4}
\]

In analysis of the dynamics.

\[
\mathbb{E}[Z(\theta, t)] \simeq Z(\bar{\theta}, t) + \frac{1}{2} \sum_{i,j} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \bar{\theta}_i \partial \theta_j} \sigma_{i,j}. \tag{2}
\]

From Eq. (2) we can see directly how uncertainties in the system dynamics may lead to shifts away from the expected value \( Z(\bar{\theta}, t) \) at different times. For example, this term could cause a reduction in the predicted average gate fidelity. This is especially relevant to analytical protocols for parameter estimation, where stochasticity in the parameter may itself modify the mean estimation. Moreover, the intuitive approach of assessing robustness via first derivatives will fail to capture the expected loss in average fidelity, as its effect will simply average to zero.

In the case that the uncertainties are statistically independent (\( \sigma_{i,j} = 0 \) if \( i \neq j \)), we can further see that the individual parameter uncertainties are additive, allowing them to be calculated individually by, e.g., Monte Carlo simulations, or more efficiently, as we show below, by propagation of uncertainty through the dynamics.

A similar calculation for the variance (see Supplemental Material) yields

\[
\text{Var}[Z(\theta, t)] \simeq \sum_{i,j} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \bar{\theta}_i \partial \theta_j} \sigma_{i,j} + \frac{1}{2} \sum_{i,j,k,l} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \bar{\theta}_i \partial \theta_j} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \bar{\theta}_k \partial \theta_l} \sigma_{i,k} \sigma_{j,l}. \tag{3}
\]

We now sketch how to obtain the derivatives that appear in Eq. (2) and (3) through analysis of the dynamics.
and similarly we obtain
\[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \chi_{n+1} = \left( \frac{\partial^2}{\partial \theta_i \partial \theta_j} U_n \right) \chi_n + \left( \frac{\partial}{\partial \theta_j} U_n \right) \frac{\partial}{\partial \theta_i} \chi_n + \left( \frac{\partial}{\partial \theta_i} U_n \right) \frac{\partial}{\partial \theta_j} \chi_n + U_n \frac{\partial^2}{\partial \theta_i \partial \theta_j} \chi_n. \] (5)

To calculate the derivatives for each time slice one should solve
\[ \frac{d}{d\eta} e^{\chi(\eta)} = \int_0^1 e^{\alpha \chi(\eta)} \frac{d\chi(\eta)}{d\eta} e^{(1-\alpha)\chi(\eta)} d\alpha. \] (6)
which has a known solution in terms of the eigenvalues of \( \chi \) [13, 14, 61]. Further steps in the derivation are given in the Supplemental Material.

II. COMPARISON TO MONTE CARLO SIMULATIONS

We consider here a two-level Hamiltonian given by \((\hbar = 1)\)
\[ H(t) = \Delta \hat{Z} + \Omega_{\text{max}} u(t) \hat{X}, \] (7)
with detuning \( \Delta = 0 \), maximum amplitude \( \Omega_{\text{max}} \), Pauli operators \( \hat{X} \) and \( \hat{Z} \), and dimensionless control function \(-1 \leq u(t) \leq +1\). We consider a state transfer |0⟩ → |1⟩ where the dynamical evolution is subject to uncertainties in the system parameters \( \Omega_{\text{max}} \) and \( \Delta \). The transfer itself is induced by Gaussian π-pulse (Fig. 1(a)). We time-evolve the system for a duration \( T = \theta / \Omega_{\text{max}} \) with independent uncertainties \( \sigma(\Omega_{\text{max}}) / \Omega_{\text{max}} = \sigma(\Delta) / \Omega_{\text{max}} = 3.0 \times 10^{-2} \) using 10,000 Monte Carlo simulations. We calculate the average trajectory over the Pauli observable \( \langle \hat{Z} \rangle \) and define the confidence interval as twice the standard deviation \( \pm 2\sigma(\langle \hat{Z} \rangle) \) from the average trajectory, which we truncate between ±1, as this is a physical limitation on \( \langle \hat{Z} \rangle \) (Fig. 1(a)). We now compare the predictions of QUPID, Eqs. (2) and (3), to the Monte Carlo simulations by zooming in on the end of the trajectory in Fig. 1(b). Here we see how uncertainties in the system parameters induce a shift in the mean value away from the ideal, uncertainty-less case \( \langle \hat{Z} \rangle = -1 \). In addition, we see an excellent agreement between the Monte Carlo simulations and uncertainty propagation predictions. On a standard laptop computer, the Monte Carlo simulations take \( \approx 30 \) minutes, whereas QUPID takes \( \approx 0.2 \) seconds, thereby providing a significant speed up.

An additional benefit of QUPID over Monte Carlo methods is that Eq. (2) allows us to directly track how each uncertainty parameter influences the system dynamics. To this end, we may define the error contribution from Eq. (2) for each parameter as \( \frac{1}{2} \left( \frac{\partial^2}{\partial \theta_i \partial \theta_j} \sigma(\theta)^2 \right) \) with \( \theta = \Omega_{\text{max}} \Delta \). We plot in Fig. 1(c) the error contribution for each parameter along the trajectory. Interestingly, we see that around half-way through the trajectory, the error contributions from the amplitude and detuning are opposite in sign and thus cancel each other out to some extent; such fortuitous cancellations can be analyzed and exploited when designing pulse sequences using our method. We also see at the end of the trajectory that the two error contributions converge on different specific values, i.e. uncertainty in the amplitude and detuning constitute respectively around 59% and 41% of the total error. This analysis is useful in determining, for example, which parameters are most sensitive in an experiment and thus where efforts must be concentrated when improving the precision of experimental systems.

III. EXPERIMENTAL COMPARISON

We compare the predictions of QUPID to a single-qubit experiment via pulse-level control of IBM’s ibmq_armonk single-qubit system accessible via Qiskit pulse [62–64]. Such pulse-level control has previously been used to optimize the fidelity of cross-resonance gates [28], the CV-gate [65], and to build a quantum compiler implementing basis gates and qudit operations [66].

First, we determine Gaussian and composite BB1 pulse sequences that drive π/2 (Hadamard) and π (bit-flip) transitions. After calibration (see Supplemental Material), we artificially add amplitude noise to each of the aforementioned pulse sequences. The amplitude noise is sampled from a Gaussian distribution centered around the optimal amplitudes with a width of \( \epsilon = \sigma(\Omega_{\text{max}}) / \Omega_{\text{max}} \) for \( \Theta = \pi / 2 \) and \( \pi \) [67]. We compare the predictions of our theory to the experimental results in Fig. 2 with \( \epsilon = 0.05 \) for Gaussian and BB1 π and π/2 pulses. The data and simulation match well throughout the duration of the pulses for both cases, although our theory cannot fully account for measured BB1 errors at the final time due to discrimination/shot noise (see blue error bars) being larger than the BB1 final error, which has intrinsic robustness up to sixth order in amplitude error [34, 36]. To determine how the error scales with \( \epsilon \), we compare the experimental data and theoretical predictions for the final expectation value \( \langle \hat{Z} \rangle \) of a Gaussian π/2 pulse in the inset of Fig. 2(c), where we again obtain excellent agreement between theory and experiment. Thus, when compared with experimental data, our theory can determine at what point various experimental uncertainties dominate the error, which can be useful for experimental debugging, theoretical model-building, or designing new systems. Moreover, by calibrating the exact value of \( \epsilon \) in the model, one can readily use Eq. (2) to remove the correction term coming from uncertainty and thereby estimate the true model parameters, e.g. for Hamiltonian learning or quantum sensing.

IV. SCALING AND SPEED UP

Here, we consider how QUPID scales with the number of parameters compared to Monte Carlo simulations. In the case of independent uncertainties, estimating the
mean via Eq. (2) scales linearly with the number of parameters, while it scales quadratically in the more general case of dependent uncertainties. If we compare to Monte Carlo simulations, the averaging error over $N$ samples is given by the standard deviation of the mean $s = \sigma/\sqrt{N}$, leading to a $N^{-1/2}$ scaling. We may consider the scaling of the standard deviation by considering the square root of Eq. (6). In the general case of dependent uncertainties, Eq. (6) has a fourth order scaling with the number of parameters, which implies that the standard deviation would scale quadratically. In the case of independent parameters, this reduces to a linear scaling. In other words, the scaling of the two approaches with respect to the number of uncertainty parameters remains, in principle, the same.

However, in practice, many of the mixed derivatives in Eq. (3) tend to be zero, leading to an effective sub-linear and sub-quadratic scaling of Monte Carlo simulations with independent and dependent uncertainties, respectively. For this reason, we should compare the precision of the two methods over a wide range of uncertainty parameters. We consider the spin-star system illustrated in the top right of Fig. 3. The system consists of 5 qubits with interactions between the nearest and the next-nearest spins, where we model an external, global $X$-control for all qubits. The Hamiltonian for the spin-star system is given by

$$H = u(t) \sum_j \hat{X}_j - \sum_{\langle i,j \rangle} \left( J^{(i,j)}_x \hat{X}_i \hat{X}_j + J^{(i,j)}_y \hat{Y}_i \hat{Y}_j + J^{(i,j)}_z \hat{Z}_i \hat{Z}_j \right) - \sum_{\langle \langle i,j \rangle \rangle} \left( g^{(i,j)}_x \hat{X}_i \hat{X}_j + g^{(i,j)}_y \hat{Y}_i \hat{Y}_j + g^{(i,j)}_z \hat{Z}_i \hat{Z}_j \right), \quad (8)$$

where $\langle \rangle$ denotes the sum over neighboring qubits, while $\langle \langle \rangle \rangle$ denotes next-neighboring qubits. $J = J^{(i,j)}_k$ denotes the nearest coupling and we choose the next-nearest coupling $g = g^{(i,j)}_k = J/10$ for $k = x, y, z$ and all indices $(i, j)$. With this model we have 30 different parameters. We start in the initial state $|\psi_0\rangle = |0, 0, 0, 0, 0\rangle$ and evolve with a control pulse of the form $u(t) = \phi_1(t) + \phi_3(t) - \phi_2(t)$, where $\phi_n(t) = \sin(n\pi t/T)$ for a total duration $T J = 10$. In the absence of uncertainties, this leads to a final state $|\psi\rangle = |\psi(T)\rangle$ that populates a wide range of different Fock states in the Hilbert space.

We investigate how uncertainties in the system parameters induces population in states other than $|\psi\rangle$, which can be described by the fidelity $F = \langle \psi | \hat{P} | \psi \rangle$, with $P = |\tilde{\psi}\rangle \langle \tilde{\psi}|$ denoting the projector onto the expected final state. Again, we assume independent normally-distributed uncertainties, with relative amplitude $\sigma(J)/J = \sigma(g)/g = 0.01$, which we model in a ran-
We model uncertainties in 1, 10, 20, and 30 parameters chosen at random, where we perform 100,000 Monte Carlo simulations for each choice. To understand the scaling of the Monte Carlo method, we compare the difference between the fidelity average over the 100,000 samples ($F_{\text{true}}$) versus only averaging over a subset of the simulations ranging from 1 to 10,000 ($F_{\text{approx}}$).

For each given subset, we draw a random number of fidelities from the entire set of simulated fidelities. This procedure is repeated 10,000 times in order to calculate the mean estimation error i.e., average over $|F_{\text{true}} - F_{\text{approx}}|$. The results are depicted in Fig. 3 for different numbers of uncertainty parameters. Note the lines scale as $N^{-1/2}$, consistent with the standard deviation of the mean, but at different heights due to different standard deviations over the fidelity distributions.

We may compare the performance of QUPID to Monte Carlo simulations in both accuracy and computational expense. The latter is achieved by comparing the number of fidelity evaluations required to calculate the derivatives using a finite difference scheme (see Supplemental Material). Since evaluating the fidelities is by far the most expensive step in each approach, this gives an accurate comparison. We plot the results of QUPID in Fig. 3 as individual data points.

We may read the speed up of QUPID relative to Monte Carlo simulations in an equal-accuracy comparison by considering the distance from the individual data points in Fig. 3 to the lines along the first axis. For clarity we have indicated the factor of speed up with arrows on the figure, which ranges from 6818 for one parameter to 109 for 30 parameters. Note that the relative speed up grows smaller with the number of parameters due to the sub-linear scaling of Monte Carlo. This implies that for a large enough number of uncertainty parameters, Monte Carlo would be preferable. However, we have still demonstrated significant improvements with up to 30 parameters, which would comprise the vast majority of cases typically considered in the literature [15, 18–28].

V. OUTLOOK

In this work, we have demonstrated QUPID, a method for analytically incorporating and propagating uncertainties during quantum dynamical evolution. For a reasonable number of parameters (at least up to 30), this method performs dramatically better than the standard, cutting-edge methods that are commonly used in the literature, namely Monte Carlo sampling. This analytical form has the additional advantages of (a) being able to separate the role of different error sources and (b) understanding how these uncertainties develop over time. This can be useful for theoretical protocol improvement but also for experimental design and debugging. That is, with an understanding of the dominant uncertainty sources in a given protocol, experimentalists can better understand the level of precision that must be reached in their system and which parameters are most likely to cause imprecisions. Likewise, these analytical forms allow the design of protocols (e.g. pulse sequences) that may either decrease or increase the effect of parameter uncertainty, which is useful for increased robustness or improved sensing, respectively.

ACKNOWLEDGMENTS

We acknowledge the use of IBM Quantum services for this work. The views expressed are those of the authors, and do not reflect the official policy or position of IBM or the IBM Quantum team. This work was funded by the Carlsberg Foundation and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC 2004/1 – 390534769. The numerical results presented in this work were obtained at the Centre for Scientific Computing, Aarhus, phys.au.dk/forskning/cseaa.
SUPPLEMENTARY MATERIAL

Calculation of variance

We start by evaluating the general expression of the variance based on the results derived in the main text.

\[
\text{Var}[Z(\theta)] = \mathbb{E}\left[ (Z(\theta, t) - \mathbb{E}[Z(\theta, t)])^2 \right]
\]

\[
= \mathbb{E}\left[ \sum_{i,j} \frac{\partial Z(\bar{\theta}, t)}{\partial \theta_i} \frac{\partial Z(\bar{\theta}, t)}{\partial \theta_j} (\theta_i - \mu_i)(\theta_j - \mu_j) \right]
\]

\[
+ \sum_{i,j,k} \frac{\partial Z(\bar{\theta}, t)}{\partial \theta_k} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \theta_i \partial \theta_j} (\theta_k - \mu_k)(\theta_i - \mu_i)(\theta_j - \mu_j)
\]

\[
- \sum_{i,j,k} \frac{\partial Z(\bar{\theta}, t)}{\partial \theta_k} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \theta_i \partial \theta_j} \sigma_{i,j}(\theta_k - \mu_k)
\]

\[
+ \frac{1}{4} \sum_{i,j,k,l} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \theta_i \partial \theta_j} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \theta_k \partial \theta_l} (\theta_i - \mu_i)(\theta_j - \mu_j)(\theta_k - \mu_k)(\theta_l - \mu_l)
\]

\[
- \frac{1}{2} \sum_{i,j,k,l} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \theta_i \partial \theta_j} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \theta_k \partial \theta_l} \sigma_{k,l}(\theta_i - \mu_i)(\theta_j - \mu_j) + \frac{1}{4} \sum_{i,j,k,l} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \theta_i \partial \theta_j} \frac{\partial^2 Z(\bar{\theta}, t)}{\partial \theta_k \partial \theta_l} \sigma_{i,j} \sigma_{k,l} \right].
\] (9)

In order to evaluate the above we use Wick’s probability theorem (also known as Isserlis' theorem), which states that the mean of product of zero-mean normal distributed variables \( \{Y_j\}_j \) is given by

\[
\mathbb{E}\left[ Y_1 Y_2 \ldots Y_n \right] = \sum_{p \in P^2_n} \prod_{(i,j) \in p} \sigma_{i,j}.
\] (10)

Here the sum is over all distinct partitions \( P^2_n \) of \( \{Y_j\}_j \) into pairs. Since this cannot be done for an odd number \( n \), the above gives zero for all odd \( n \). From this we e.g. have

\[
\mathbb{E}\left[ Y_1 Y_2 Y_3 \right] = 0; \quad \mathbb{E}\left[ Y_1 Y_2 Y_3 Y_4 \right] = \sigma_{1,2}\sigma_{3,4} + \sigma_{1,3}\sigma_{2,4} + \sigma_{1,4}\sigma_{2,3}.
\] (11)

Applying Wick’s probability theorem to Eq. (9) and using that the order of differentiation does not matter we obtain the expression for the variance as stated in the main text.

Numerical calculation of derivatives

In this work we seek to solve the equation of motion for a quantum system on the form

\[
\frac{\partial}{\partial t} \chi(t) = A(t) \chi(t),
\] (12)

which could be the Schrödinger equation with \( A(t) = -iH(t) \) and \( \chi(t) \) being a single state \( \chi(t) = |\psi(t)\rangle \), a unitary \( \chi(t) = U(t) \), or a collection of \( K \) states with the \( n \)th column of \( \chi \) being the \( n \)th state \( \chi(t) = \left[ |\psi_1(t)\rangle, |\psi_2(t)\rangle, \ldots, |\psi_K(t)\rangle \right] \). Eq. (12) could also be the von Neumann equation or the master equation on the Lindblad form, where \( A(t) = L(t) \) denotes the Liouvillean and \( \chi = |\rho\rangle \) the vectorized form of the density matrix i.e. \( |\rho\rangle_{n+(m-1)d} = \rho_{n,m} \), with \( d \) denoting the Hilbert space dimension. In either case we may numerically solve the equation of motion by propagating the state in sufficiently small time-steps \( \Delta t \) via the time evolution operator elaborated on in the main text.

In order to evaluate the derivatives, we assume the time propagator on the form \( A(t) = A_0(t) + \sum u_n A_n(t) \), where \( u_n \) is assumed a constant with time. This could either be a system-parameter \( u_n = \theta_n \) or a function of the system parameters. In the latter case we may evaluate the derivative via the chain rule \( \frac{\partial}{\partial \theta_j} = \frac{\partial u_n}{\partial \theta_j} \frac{\partial}{\partial u_n} \). Here both the first \( u \frac{\partial}{\partial u} \) and second order derivatives \( u \frac{\partial^2}{\partial u^2} \) has previously been calculated in literature but the results are reproduced here for completeness. Note this approach also allow us model uncertainties in the initial state of the
system $\chi_0 = \chi(t_0)$, which could e.g. stem from initialization errors. Note the following works for Hermitian dynamics such as the Schrödinger equation $A = -iH$. For non-Hermitian dynamics the second order derivative is calculable by other means [71]. In the following we assume, for simplicity, that $A = -iH$.

The derivative of the time evolution operator may be evaluated in the eigenbasis of $H_n = H(t_n + \Delta t/2)$ [13, 61]

$$\left\langle h \left| \frac{\partial U_n}{\partial \theta_j} \right| k \right\rangle = \left\langle h \left| \frac{\partial H_n}{\partial \theta_j} \right| k \right\rangle I(h, k).$$

(13)

Here $|h\rangle$ and $|k\rangle$ denote eigenstates of $H_n$. We have further defined

$$I(h, k) = \begin{cases} -i\Delta t e^{-iE_h \Delta t}, & \text{if } E_h = E_k \\ e^{iE_h \Delta t} - e^{-iE_k \Delta t}, & \text{if } E_h \neq E_k. \end{cases}$$

(14)

Here $E_h$ denotes the eigenenergy of $|h\rangle$. In Ref. [14] the second order derivative was also calculated

$$\left\langle h \left| \frac{\partial U_n}{\partial \theta_j \partial \theta_j} \right| k \right\rangle = \sum_l \left( \frac{\partial H_n^{(h,l)}}{\partial \theta_i} \frac{\partial H_n^{(l,k)}}{\partial \theta_j} + \frac{\partial H_n^{(h,l)}}{\partial \theta_j} \frac{\partial H_n^{(l,k)}}{\partial \theta_i} \right) I(k, l, h),$$

(15)

where we use the notation $\frac{\partial H_n^{(h,l)}}{\partial \theta_i} = \langle h | \frac{\partial H_n}{\partial \theta_i} | l \rangle$. Here $I(k, l, h)$ can be expressed through the already calculated elements $I(h, k)$, see Eq. (B9) in Ref. [14].

The method outlined here requires propagating a single state, if we assume $M$ system-parameters, we also need $M$ first order derivatives, and $M^2$ second order derivatives. However, since the order of differentiation does not matter, this reduces to $\frac{1}{2}M(M-1)$ second order derivatives. In most applications, this is significantly lower than Monte Carlo sampling, which typically requires several hundreds if not thousands of samples.

To concretize these ideas, we demonstrate how to calculate the derivatives of an expectation value of a Hermitian operator $\hat{A}$ with respect to a single state $Z(t) = \langle \psi(t) | \hat{A} | \psi(t) \rangle$ as also treated in the main text. This has also previously been calculated in e.g. Ref. [11].

$$\frac{\partial}{\partial \theta_j} Z(t) = \left\langle \frac{\partial}{\partial \theta_j} \psi(t) \left| \hat{A} \right| \psi(t) \right\rangle + \left\langle \psi(t) \left| \hat{A} \right| \frac{\partial}{\partial \theta_j} \psi(t) \right\rangle + \left\langle \psi(t) \left| \frac{\partial \hat{A}}{\partial \theta_j} \right| \psi(t) \right\rangle$$

$$= 2 \Re \left[ \left\langle \psi(t) \left| \hat{A} \right| \frac{\partial}{\partial \theta_j} \psi(t) \right\rangle + \left\langle \psi(t) \left| \frac{\partial \hat{A}}{\partial \theta_j} \right| \psi(t) \right\rangle \right].$$

(16)

Where we in the above equation, have taken into account that $\hat{A}$ itself may also be dependent on the parameters, for instance, the case where $\hat{A} = H$ is the expected energy of the system. The second derivative reads

$$\frac{\partial^2}{\partial \theta_i \partial \theta_j} Z(t) = 2 \Re \left[ \left\langle \frac{\partial}{\partial \theta_i} \psi(t) \left| \hat{A} \right| \frac{\partial}{\partial \theta_j} \psi(t) \right\rangle + \left\langle \psi(t) \left| \hat{A} \right| \frac{\partial^2}{\partial \theta_i \partial \theta_j} \psi(t) \right\rangle + \left\langle \psi(t) \left| \frac{\partial \hat{A}}{\partial \theta_j} \right| \frac{\partial}{\partial \theta_i} \psi(t) \right\rangle \right]$$

$$+ \left\langle \psi(t) \left| \frac{\partial^2 \hat{A}}{\partial \theta_i \partial \theta_j} \right| \psi(t) \right\rangle.$$  

(17)

As an alternative to calculating the derivatives exactly, we also demonstrate how they may be estimated using finite differences [72]. The first derivative using a central difference scheme is given by

$$\frac{\partial Z(\theta_j)}{\partial \theta_j} = \frac{Z(\theta_j + \epsilon) - Z(\theta_j - \epsilon)}{2\epsilon} + \mathcal{O}(\epsilon^2),$$

(18)

where we have used the shorthand notation $Z(\theta, t) = Z(\theta_1, \theta_2, \ldots, \theta_j, \ldots, \theta_M, t) = Z(\theta_j)$. Note, the error of the approximation in Eq. (18) can be made negligibly small by a sufficiently small selection of $\epsilon$, but not too small to induce round-off errors. In a similar manner, the second order derivative reads
\[
\frac{\partial^2 Z(\theta_j)}{\partial \theta_j^2} = \frac{Z(\theta_j + \epsilon) - 2Z(\theta_j) + Z(\theta_j - \epsilon)}{\epsilon^2} + O(\epsilon^2).
\] (19)

Note that Eq. (18) and (19) requires \(1 + 2M\) function evaluations, where \(M\) again denotes the number of system-parameters. That is, for applications where we do not model covariances (i.e. assume independent parameters), and only need the mean value, our method scales linearly with the number of function evaluations.

If we also needed the mixed derivatives, a finite difference scheme is given by

\[
\frac{\partial^2 Z(\theta_j, \theta_i)}{\partial \theta_j \partial \theta_i} = \frac{1}{\epsilon^2} \left[ Z(\theta_j + \epsilon, \theta_i + \epsilon) - Z(\theta_j + \epsilon, \theta_i) - Z(\theta_j, \theta_i + \epsilon) + Z(\theta_j, \theta_i - \epsilon) \right] + O(\epsilon^2).
\] (20)

Note, this only requires two additional function evaluations for each of the \(\frac{1}{2}M(M-1)\) possible combinations. This means that the overall number of evaluations becomes \(1 + M + M^2\). Note that although finite difference schemes are often easier to implement, they are not as numerically stable as calculating the exact derivatives, which does not depend on a choice of \(\epsilon\). Note in the main text, Section IV we compare calculating the mean with independent parameters, and so we use a scaling of \(1 + 2M\).

**Experimental details**

The details of the experiment are as follows: First, we initialize the qubit in the \(|0\rangle\) state. Then we drive the qubit on resonance with a series of Gaussian pulses of constant width (the pulse is truncated at \(\pm 6\sigma\), where \(\sigma = 75\) ns) and varying amplitude, which effectively realizes the two level Hamiltonian in the main text with \(\Delta = 0\) and varying \(\Omega_{\text{max}}\). The qubit is then measured and a discriminator is applied such that we obtain a zero or one depending on if the qubit was in the ground or excited state, respectively. This experiment, and all other experimental runs reported in this manuscript, is repeated 1024 times to obtain statistics, and the resulting curve is fit to a sinusoid and scaled between \(\langle z \rangle = \pm 1\), which allows us to maintain consistency with theory in the face of experimental systematics that result in finite contrast. From this, we can determine the pulse amplitudes \(A_\Theta\) required to obtain a given \(X\) rotation \(\Theta\). This allows us to calibrate the system.

The Gaussian pulses found above can also be used to form composite BB1 pulse sequences [34], which take the form

\[
R_{\text{BB1}}(\Theta) = R_{\pi, \phi(\Theta)} R_{2\pi, 3\phi(\Theta)} R_{\pi, \phi(\Theta)} R_{\phi, 0}
\] (21)

where \(R_{\theta, \phi}\) defines a Gaussian pulse with a phase \(\phi\) that results in a rotation \(\theta\). The optimal phase \(\phi\) for a desired final rotation \(\Theta\) is given by \(\phi(\Theta) = \cos^{-1}(-\Theta/4\pi)\) [34].

[1] L. K. Grover, in Proceedings of the twenty-eighth annual ACM symposium on theory of computing (1996) pp. 212–219.
[2] P. W. Shor, SIAM Review 41, 303 (1999).
[3] J. T. Barreiro, M. Müller, P. Schindler, D. Nigg, T. Monz, M. Chwalla, M. Hennrich, C. F. Roos, P. Zoller, and R. Blatt, Nature 470, 486 (2011).
[4] I. Bloch, J. Dalibard, and S. Nascimbene, Nature Physics 8, 267 (2012).
[5] A. Das and B. K. Chakrabarti, Reviews of Modern Physics 80, 1061 (2008).
[6] C. C. McGeoch, Synthesis Lectures on Quantum Computing 5, 1 (2014).
[7] M. Schulz, I. Sinayskiy, and F. Petruccione, Contemporary Physics 56, 172 (2015).
[8] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature 549, 195 (2017).
[9] A. Acín, I. Bloch, H. Buhrman, T. Calarco, C. Eichler, J. Eisert, D. Esteve, N. Gisin, S. J. Glaser, F. Jelezko, et al., arXiv preprint arXiv:1712.03773 (2017).
[10] J. Taylor, Introduction to Error Analysis, the Study of Uncertainties in Physical Measurements (University Science Books, 1997).
[11] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Journal of magnetic resonance 172, 296 (2005).
[12] F. Motzoi, J. M. Gambetta, S. Merkel, and F. Wilhelm, Phys. Rev. A 84, 022307 (2011).
[13] P. De Fouquieres, S. Schirmer, S. Glaser, and I. Kuprov, Journal of Magnetic Resonance 212, 412 (2011).
[14] M. Daalderop, F. Motzoi, J. H. M. Jensen, and J. Sherson, Physical Review A 102, 042612 (2020).
[15] V. Langrock and D. P. Di Vincenzo, arXiv preprint arXiv:2012.09517 (2020).
[16] J. J. Sørensen, J. S. Nyemann, F. Motzoi, J. Sherson, and T. Vosegaard, The Journal of Chemical Physics 152, 054104 (2020).
[17] X. Ge, H. Ding, H. Rabitz, and R.-B. Wu, Physical Review A 101, 052317 (2020).
[18] R. Vieira and G. Rigolin, Physics Letters A 384, 126536 (2020).
[19] X. Song, J. Ryu, K. Baek, and J. Bang, Journal of the Korean Physical Society 1, 1 (2021).
[20] F. Motzoi, E. Halperin, X. Wang, K. B. Whaley, and S. Schirmer, Physical Review A 94, 032313 (2016).
[21] J. Sherson, F. Motzoi, and A. Buchleitner, Physical review letters 105, 020501 (2010).
[22] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Physical Review A 80, 012304 (2009).
[23] H. Mabuchi, Quantum and Semiclassical Optics: Journal of the European Optical Society Part B 8, 1103 (1996).
[24] S. Schirmer and D. Oi, Laser Physics 20, 1203 (2010).
[25] M. Guţă and N. Yamamoto, IEEE Transactions on Automatic Control 61, 921 (2015).
[26] S. L. Braunstein and C. M. Cavé, Physical Review Letters 72, 3439 (1994).
[27] C. Rigetti and D. V. Houdre, Physical Review A 105, 042319 (2020).
[68] G.-C. Wick, Physical Review 80, 268 (1950).
[69] L. Isserlis, Biometrika 12, 134 (1918).
[70] C. Navarrete-Benlloch, arXiv preprint arXiv:1504.05266 (2015).
[71] D. Goodwin and I. Kuprov, The Journal of chemical physics 144, 204107 (2016).
[72] B. Fornberg, Mathematics of computation 51, 699 (1988).