



Quantum state tomography with time-continuous measurements: reconstruction with resource limitations

Areeya Chantasri¹ · Shengshi Pang ·
Teerawat Chalermputitarak · Andrew N. Jordan

Received: 4 April 2019 / Accepted: 15 May 2019
© Chapman University 2019

Abstract We propose and analyze quantum state estimation (tomography) using continuous quantum measurements with resource limitations, allowing the global state of many qubits to be constructed from only measuring a few. We give a proof-of-principle investigation demonstrating successful tomographic reconstruction of an arbitrary initial quantum state for three different situations: single qubit, remote qubit, and two interacting qubits. The tomographic reconstruction utilizes only a continuous weak probe of a single qubit observable, a fixed coupling Hamiltonian, together with single-qubit controls. In the single-qubit case, a combination of the continuous measurement of an observable and a Rabi oscillation is sufficient to find all three unknown qubit state components. For two interacting qubits, where only one observable of the first qubit is measured, the control Hamiltonian can be implemented to transfer all quantum information to the measured observable, via the qubit–qubit interaction and Rabi oscillation controls applied locally on each qubit. We discuss different sets of controls by analyzing the unitary dynamics and the Fisher information matrix of the estimation in the limit of weak measurement, and simulate tomographic results numerically. As a result, we obtained reconstructed state fidelities in excess of 0.98 with a few thousand measurement runs.

Keywords Quantum state tomography · Quantum trajectories · Quantum measurement · Bayesian estimation

A. Chantasri · S. Pang · A. N. Jordan
Department of Physics and Astronomy, University of Rochester, Rochester, NY 14627, USA

A. Chantasri · S. Pang · A. N. Jordan
Center for Coherence and Quantum Optics, University of Rochester, Rochester, NY 14627, USA

A. Chantasri (✉)
Centre for Quantum Computation and Communication Technology (Australian Research Council), Centre for Quantum Dynamics, Griffith University, Nathan 4111, Australia
e-mail: a.chantasri@griffith.edu.au

T. Chalermputitarak
Optical and Quantum Physics Laboratory, Department of Physics, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

A. N. Jordan
Institute for Quantum Studies, Chapman University, 1 University Drive, Orange, CA 92866, USA

1 Introduction

Quantum state tomography is a general method for reconstructing a quantum state based on a collection of measurements made on a quantum system, all starting from the unknown state of interest [27]. State tomography is one of the most important parts in building quantum information processing devices, as it is used in verifying quantum states before, during, or after any of the devices' processes [28]. Standard procedures to implement tomography in the laboratory mainly aim to approximate textbooks projective measurements applied on separate ensembles, in which the measured observables are chosen carefully to assure informational completeness for the unknown quantum states [4, 15, 20, 31]. However, researchers have also explored different measurement and estimation approaches to achieve the same end, for example, using a sequential unsharp measurement [2], or continuous weak measurement [32, 34–37] in combination with estimation techniques such as Bayesian or maximum likelihood methods to estimate initial states of the measurement processes [3, 18].

Of particular interest to this paper are investigations using continuous quantum measurements, of the type pioneered in recent experimental works [5, 17, 25, 26, 33, 38] and based on by now well-established theoretical methods such as the stochastic Schrödinger equation, stochastic master equation [1, 6, 19, 41], and stochastic path integral [7, 8]. We note that the weak continuous measurement is effectively a projective measurement after a long enough data-collection time [16, 41], but the advantage of the former over the later is that the model is more practical in experiments and it allows us to include additional controls to the systems during the measurement, to achieve desired processes [13, 40]. The work by Six et al. [36] theoretically and experimentally demonstrates quantum state tomography using continuous measurements and the maximum likelihood method on Rydberg atoms and superconducting circuit experiments. We highlight in particular the works of the group of I. Deutsch and P. S. Jessen, concerning tomographic reconstruction of initial collective state of an ensemble of atomic spins using continuous measurements [9, 32, 34, 35, 37]. In these works, the measured observables are fixed and the information about other observables of interest is continually mapped onto the measured ones via continuous controlled unitary dynamics. We adapt this idea to measurements on individual qubits, using continuous measurement limited to only a single observable and continuous controls, to extract information about unknown initial states. Measurement backaction must be accounted for to accurately reconstruct the state. Closely related to the problem of quantum state tomography is that of parameter estimation, for which there have been a series of works dedicated to estimating Hamiltonian parameters from data acquired by a continuous measurement sequence [10, 21, 29, 30, 39, 42].

In this paper, we take up the problem of continuous measurement tomography applied to a network of qubits. Our aim is to show that global state reconstruction is possible using a fixed set of control parameters in a continuous way, even when all qubits are not measured. The main advantage of such a technique is that in contrast to projective measurement tomography, where a sequence of different gate operations is required to then read out one of many different observables (each of which is repeated many times), in our continuous measurement construction, global information about the entire state is encoded in each run of the continuous measurement. Therefore, the tradeoff is that there is only a bit of information about all of the observables in each run of the experiment, and the protocol must be repeated many times to reduce the statistical uncertainties below a given level. Nevertheless, the experimental procedure becomes much simpler: it is the same every time in principle because the local controls are fixed, and only the state reconstruction algorithms become more complicated to numerically implement. An important problem then becomes how to choose the local controls and estimation methods to accurately find the initial state.

Several different methods are introduced to solve the above problems. We discuss a method of using commutators between the Hamiltonian of the system's unitary dynamics and the different observables to be estimated to choose appropriate values of the local controls that allow all elements of interest to be accessed by the readout channel. We also present the calculation of the Fisher information matrix about the initial quantum state, from which we can optimize the controls to minimize the statistical uncertainties. Given chosen sets of controls, continuous measurement records are then used to compute the Bayesian and maximum likelihood estimators for the unknown states. We construct a mapping operator for all observed records, proportional to an element of a positive-operator-valued measure (POVM), and use it to compute Bayesian probability for all possible candidate states. This is to

dispense with computing separate quantum trajectories for all candidate states which greatly speeds up the numerical estimation procedure.

We illustrate our methods first in detail for the case of a single qubit, showing how the estimation procedures can be applied using quantum trajectory theory, and giving the tomography results given a local control and single component pseudo-spin readout. We then incorporate a resource limitation, and show the same methods are able to reconstruct the state of a remote qubit that is coupled to the one being read out. Finally, we demonstrate arbitrary full two-qubit state reconstruction using local controls and fixed coupling. We show that by choosing the local controls within a relatively broad range of parameters, good state fidelity with the target state can be easily reached.

The paper is organized as follows. In Sect. 2, we layout the background information for the continuous measurement of a single qubit's observable, and estimation techniques used for quantum state tomography. Our main results are presented in Sect. 3, including the methods to choose local qubit controls in Sect. 3.1, and the analysis of the three examples: single-qubit state tomography in Sect. 3.2, remote-qubit state tomography in Sect. 3.3, and two-qubit tomography in Sect. 3.4, with results from numerical simulations. We conclude our results in Sect. 4 and present detailed calculation of the Fisher information matrix and numerical methods in the Appendices.

2 Background

2.1 Time-continuous measurement of a qubit's observable

The concept of quantum measurement generalized to arbitrary strength and occurring over a finite time [1, 12] is becoming more recognized in the quantum information research community [19, 41]. Not only that it describes more accurately the measurement processes in practical experiments, but it also opens possibilities for adding system controls during the measurement processes, such as feedback, to achieve desired outcomes more efficiently. Generalized quantum measurement theory considers a system of interest interacting with its environment (or detectors) where all or parts of the environment are observed. The evolution of the system depends on the system–environment interactions, as well as on the observed results of the measurement performed on the environment.

In a limit where there is no information about the actual state of the environment, i.e., the environment is not observed or is averaged over, under the Markov assumption, the system's evolution is described by a Lindblad master equation [24],

$$\partial_t \rho = -i[\hat{H}, \rho] + \sum_k \mathcal{D}[\hat{c}_k] \rho, \quad (1)$$

where ρ is a density matrix representing a state of the system and $\hat{H}$ is a Hamiltonian describing unitary dynamics of the system. The Lindblad operator, defined as $\mathcal{D}[\hat{c}] \rho = \hat{c} \rho \hat{c}^\dagger - \frac{1}{2}(\hat{c}^\dagger \hat{c} \rho + \rho \hat{c}^\dagger \hat{c})$, describes the decoherence of the system's state as a result of the interaction with the environment, via different channels denoted by the Lindblad operators $\hat{c}_k$.

In general, there can be many dephasing channels to consider; however, in this work we are interested in the interaction that leads to a measurement of an observable $\hat{Z}$ of a qubit, which might be part of a network of many qubits. Therefore, the system's dynamics only includes unitary dynamics [the first term in Eq. (1)] and measurement backaction conditioning on particular measurement records. The backaction from continuous weak measurement of a single qubit has been studied quite extensively, largely motivated by experiments in quantum optics, and in solid state physics such as electronic states in double quantum dots. More recently, superconducting qubits coupled dispersively to microwave cavities have been investigated [25, 33].

We consider a single measurement record from an experiment. Detectors will typically return a discretized signal with a time step dt to get $R \equiv \{r_1, r_2, \dots, r_n\}$, where r_k is an average signal from t_{k-1} to $t_k = t_{k-1} + dt$ and $n dt = T$,

assuming that dt is long enough for the Markov assumption to be valid, but short enough so that all types of system's dynamics commute with each other to first order of dt . The dynamics for the system state under the measurement can be written in a Kraus form, given an initial state ρ_0 ,

$$\rho(t) = \frac{\mathcal{M}_R \rho_0 \mathcal{M}_R^\dagger}{\text{Tr}[\mathcal{M}_R^\dagger \mathcal{M}_R \rho_0]}, \quad (2)$$

where we have used an operator $\mathcal{M}_R$ mapping the state from an initial time $t = 0$ to any time t . The operator is defined as $\mathcal{M}_R \equiv U(dt)M(r_n) \dots U(dt)M(r_1)$, where $U(dt) = \exp(-i\hat{H}dt)$, describing both the unitary dynamics and measurement backaction from observing the record R . The operator $M(r_k)$ is a measurement operator, which, for our studies, can be calculated using the quantum Bayesian method introduced in [22, 23],

$$M(r_k) = \left(\frac{dt}{2\pi\tau}\right)^{\frac{1}{4}} \exp\left[-\frac{(r_k - \hat{Z})^2 dt}{4\tau}\right], \quad (3)$$

describing an approximate Gaussian distribution of the measurement results. The observable $\hat{Z} \equiv \hat{Z}_1 \otimes \hat{I}_2 \otimes \dots \otimes \hat{I}_m$ represents the z Pauli operator on the first qubit of the system of m qubits, and can be considered as the dephasing channel $\hat{c}_k \propto \hat{Z}$ in Eq. (1). The measurement strength is characterized by a measurement rate $\Gamma \equiv 1/2\tau$, where τ in Eq. (3) stands for a characteristic measurement time, i.e., a signal-integration time to reach unit signal-to-noise ratio [23]. Therefore, given a measurement record R and a known Hamiltonian $\hat{H}$, one can calculate a quantum trajectory $\rho(t)$ from an initial state ρ_0 .

Since the measurement results are probabilistic, we can compute a probability density function for a record R to occur given the initial state ρ_0 ,

$$P(R|\rho_0) = \text{Tr}[\mathcal{M}_R^\dagger \mathcal{M}_R \rho_0], \quad (4)$$

which is the same as the denominator of Eq. (2). In the case that the initial state ρ_0 is unknown, a Bayesian probability density function of a possible initial state can be obtained as: $P(\rho'|R) \propto P(R|\rho')q(\rho')$ given a prior probability function of the unknown state $q(\rho')$. This Bayesian probability function for an unknown initial state is the main quantity used in the state estimation process.

We note that Eq. (2) is equivalent to the stochastic master equation [19, 41] in the limit of $dt \rightarrow 0$. Decoherence, inefficiencies, or extra dephasing effects can be added to the above model, which only results in degrading the quality of the state estimation.

2.2 Introduction to quantum state tomography

In this section, we review some main approaches to quantum state tomography, and briefly discuss about how to quantify the errors of tomography. We will mainly focus on strategies of the linear inversion (LI), the maximum likelihood estimation (MLE) and the Bayesian mean estimation (BME).

2.2.1 Tomography strategies

Quantum states can be characterized by a set of parameters which quantify the weights of different basis operators in the density matrix. The purpose of quantum state tomography is to find these parameters that determine the quantum states. Suppose we want to reconstruct a quantum state in a d dimensional Hilbert space, and the density matrix of the state is ρ . We can expand ρ along an orthogonal matrix basis $\{I, \hat{E}_1, \dots, \hat{E}_{d^2-1}\}$ in the $\mathbb{C}^d \times \mathbb{C}^d$ space,

$$\rho = \frac{1}{d} \left(\hat{I} + \sum_{i=1}^{d^2-1} c_i \hat{E}_i \right), \quad (5)$$

where $\hat{I}$ is the identity operator on the d -dimensional Hilbert space, $\hat{E}_i$'s satisfy $\text{Tr}(\hat{E}_i) = 0$, $\text{Tr}(\hat{E}_i \hat{E}_j) = \delta_{ij}d$ and $\hat{E}_i^\dagger = \hat{E}_i$. It is straightforward to verify that

$$c_i = \text{Tr}(\rho \hat{E}_i), \quad (6)$$

is a component of the state ρ corresponding to an observable $\hat{E}_i$.

Using a linear inversion method [27], by measuring each observable $\hat{E}_i$, we can obtain an estimate of c_i from,

$$\check{c}_{i,\text{LI}} = \frac{\sum_j N_{ij} \lambda_{ij}}{\sum_j N_{ij}}, \quad (7)$$

where λ_{ij} are the eigenvalues of $\hat{E}_i$, and N_{ij} are the number of times that the measurement result turns out to be λ_{ij} when measuring $\hat{E}_i$. We note that we use a variable with an upside-down hat “ $\check{a}$ ” to represent an estimator of that variable a . Due to the constraint, $\text{Tr}(\rho) = 1$, we only need to measure $d^2 - 1$ of $\hat{E}_i$'s. However, while this linear inversion is simple and efficient, a major defect of this approach is that the estimate of the density matrix it generates is not always a legitimate one for quantum systems. This is because the unknown coefficients c_i 's are estimated separately and the fluctuation in the estimates of the coefficients may make the estimate of the density matrix non-physical. To resolve this issue, we need to find ways to estimate the density matrix of a quantum system within the positivity constraint, instead of estimating the coefficients of the density matrix individually.

The maximum likelihood estimation (MLE) and the Bayesian mean estimation (BME) can resolve the positivity issue. The key idea of these two approaches is to assign prior probabilities to different possible density matrices of the quantum system, so that invalid candidate states can be ruled out.

Let us suppose that we have measurement data denoted by X , which can include many measurement records obtained from measuring an ensemble of unknown quantum systems, and a likelihood function of an unknown quantum state ρ' given the observed record X is denoted by $\mathcal{L}(\rho'; X) \equiv P(X|\rho')$. The MLE [18] is to choose the density matrix which maximizes the likelihood function as the estimate of the unknown state. Mathematically, the estimate of the density matrix by MLE is given by,

$$\check{\rho}_{\text{MLE}} = \arg \max_{\rho'} \mathcal{L}(\rho'; X), \quad (8)$$

noting that the prior distribution for MLE is implicitly the uniform distribution over all legitimate candidate states where the likelihood function is maximized.

In the case if prior probability distribution $q(\rho')$ is non-uniform over the valid state space of ρ' , we can compute the most probable Bayesian estimator (MPBE) for the tomography,

$$\check{\rho}_{\text{MPBE}} = \arg \max_{\rho'} q(\rho') \mathcal{L}(\rho'; X), \quad (9)$$

which differs from the MLE in explicitly taking a non-uniform prior distribution $q(\rho')$ into account.

The benefit of MLE is that, if a large amount of data is given, the MLE can asymptotically saturate the Cramér–Rao bound for parameter estimation (which will be introduced in the next subsection). Therefore, in principle, the MLE method can reconstruct the unknown density matrix with an optimal precision. However, the MLE also suffers from some drawbacks. For example, the estimate may be rank-deficient, sensitive to initial choices of states used in maximization algorithms, and easily trapped in local minimas.

On the other hand, the BME [3] approach, reconstructing the density matrix by averaging overall possible states weighted with posterior probabilities, is capable to overcome the defects of the MLE method. Given that the prior probability distribution for a candidate state ρ' is $q(\rho')$, the posterior distribution is,

$$P(\rho'|X) = \frac{q(\rho') \mathcal{L}(\rho'; X)}{\int q(\rho') \mathcal{L}(\rho'; X) d\rho'}, \quad (10)$$

and the BME is then defined as

$$\check{\rho}_{\text{BME}} = \int P(\rho'|X) \rho' d\rho'. \quad (11)$$

We note that the advantage of BME is that it can provide full-rank estimate of the density matrix as well as simple ways to calculate error bars of the estimate (see next section). Moreover, one can optimize the accuracy of the state tomography measured by some operational divergence [3]. However, this BME method requires sampling the whole space of the candidate states rather than only near the maximum as in Eq. (8), which is generally a computationally intensive task. Moreover, in specific problems where there is a prior knowledge of an unknown state, such as the unknown density matrix is of a lower rank (e.g., a pure state), the BME and MLE methods presented above can still be applied by specifying the prior distribution as needed.

2.2.2 Benchmarking the performance of tomography

To characterize the performance of a quantum state tomography scheme, we need to define error measures that properly quantify any deviations of the estimated states from their corresponding true states. From the point of view of parameter estimation, quantum state tomography is an estimation problem, where the unknown parameters to be estimated are the coefficients of quantum states of interest. Therefore, we consider three types of error measures: the mean square error for the individual components, quantum state fidelity, and the trace distance.

The density matrix of a quantum system is characterized by parameters c_i 's, in Eq. (5). We denote their estimates by $\check{c}_i$'s and write the true and estimated parameters in vector forms as, $\mathbf{c} = \{c_1, \dots, c_{d^2-1}\}$ and $\check{\mathbf{c}} = \{\check{c}_1, \dots, \check{c}_{d^2-1}\}$. The mean square errors of the estimators $\check{c}_i$'s can be described by an error covariance matrix,

$$\text{MSE} = \mathbb{E}[(\check{\mathbf{c}} - \mathbf{c})^T (\check{\mathbf{c}} - \mathbf{c})], \quad (12)$$

where $\mathbb{E}[\dots]$ represents an average over all possible tomography results. If the estimation is unbiased, then $\mathbb{E}[\check{\mathbf{c}}] = \mathbf{c}$ and MSE becomes a covariance matrix of $\check{\mathbf{c}}$,

$$\text{Cov} = \mathbb{E}[(\check{\mathbf{c}} - \mathbb{E}[\check{\mathbf{c}}])^T (\check{\mathbf{c}} - \mathbb{E}[\check{\mathbf{c}}])]. \quad (13)$$

In estimation theory, the covariance matrix of unbiased estimators of deterministic parameters has a lower bound set by the Cramér–Rao bound [11],

$$\text{Cov} \geq \frac{1}{N} \mathbf{F}^{-1}, \quad (14)$$

where N is a number of repeated independent measurements, and $\mathbf{F}$ is the Fisher information matrix. Following the previous section, let us use X to denote the measurement result where its statistics are described by a probability distribution $P(X|\rho)$. The Fisher information matrix is defined as

$$\mathbf{F} = \int \frac{\nabla^T P(X|\rho) \nabla P(X|\rho)}{P(X|\rho)} dX, \quad (15)$$

where ∇ is a vector of partial derivative operators with respect to the parameters $\mathbf{c}$, i.e., $\nabla = \{\frac{\partial}{\partial c_1}, \dots, \frac{\partial}{\partial c_{d^2-1}}\}$.

We note that the matrix $\mathbf{F}$ is not always invertible as required in Eq. (14), unless it is a full-rank matrix. The necessary and sufficient condition for $\mathbf{F}$ to be full-rank is that there exists a $(d^2 - 1) \times (d^2 - 1)$ full-rank matrix of this form,

$$\mathcal{G}_{\mathcal{X}} = \begin{bmatrix} \frac{\partial P(x_1|\rho)}{\partial c_1} & \dots & \frac{\partial P(x_1|\rho)}{\partial c_{d^2-1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial P(x_{d^2-1}|\rho)}{\partial c_1} & \dots & \frac{\partial P(x_{d^2-1}|\rho)}{\partial c_{d^2-1}} \end{bmatrix}, \quad (16)$$

where $P(x_i|\rho)$ is a probability distribution describing the likelihood of getting a measurement outcome $x_i \in X$ and the set of outcomes $\mathcal{X} = \{x_1, x_2, \dots, x_{d^2-1}\}$ can be any $d^2 - 1$ values in the possible range of X . This is a necessary and sufficient condition as one can show that $\mathbf{F} = \mathcal{G}_{\mathcal{X}}^\dagger \mathcal{G}_{\mathcal{X}} = \mathcal{G}_{\mathcal{X}}^\dagger \mathcal{G}_{\mathcal{X}} + \mathcal{G}_{\mathcal{X}^c}^\dagger \mathcal{G}_{\mathcal{X}^c}$, where $\mathcal{G}_{\mathcal{X}}$ and $\mathcal{G}_{\mathcal{X}^c}$ are defined similar to (16) for x_i in the whole set of X (all possible values of X) and $\mathcal{X}^c$ (a complement set of $\mathcal{X}$), respectively.

The covariance matrix defined in Eq. (13) characterizes the estimation errors of individual parameters and the correlation between them; however, it cannot give an overall benchmark for the quality of the quantum state reconstruction. We therefore need to use quantities that can measure quantum state distances to characterize the overall error of quantum state tomography. Two typical measures for the distance between two quantum states are the fidelity and the trace distance [28].

The quantum state fidelity between an estimate $\check{\rho}$ and a true density matrix ρ is defined as,

$$F(\check{\rho}, \rho) = \text{Tr} \sqrt{\sqrt{\rho} \check{\rho} \sqrt{\rho}}, \quad (17)$$

whereas, the trace distance between $\check{\rho}$ and ρ is

$$D(\check{\rho}, \rho) = \frac{1}{2} \text{Tr} |\check{\rho} - \rho|. \quad (18)$$

For a single-qubit state, the trace distance is proportional to the Euclidean distance between the Bloch vectors of the two states. But for a general case, an important relation between the fidelity and the trace distance is,

$$1 - F(\check{\rho}, \rho) \leq D(\check{\rho}, \rho) \leq \sqrt{1 - F(\check{\rho}, \rho)^2}, \quad (19)$$

which implies that, if $F(\check{\rho}, \rho)$ is close to unity, $D(\check{\rho}, \rho)$ is close to zero, and vice versa. The two measures are consistent in characterizing any discrepancies between two neighboring quantum states.

Therefore, if there are N independent measurements, each gives an estimate state $\check{\rho}$, we can define an average fidelity $\mathbb{E}[F(\check{\rho}, \rho)]$ as well as an average infidelity $1 - \mathbb{E}[F(\check{\rho}, \rho)]$, the same way we have defined the covariance matrix Eq. (12) [also Eq. (13) for an unbiased estimation]. For N is large and the estimation is unbiased, the average fidelity is related to the covariance matrix as,

$$\mathbb{E}[F(\check{\rho}, \rho)] = 1 + \sum_{ij} \text{Cov}_{ij} \text{Tr} \chi_{ij}, \quad (20)$$

where the indices i, j are for all elements of the covariance matrix Cov in Eq. (13) and χ_{ij} are operators that only depends on the true state ρ (see details of the derivation in “Appendix 1”). The trace of χ_{ij} is,

$$\text{Tr} \chi_{ij} = \sum_{\lambda_k + \lambda_l \neq 0} \frac{\lambda_k \text{Re} \left(\hat{E}_i^{(lk)} \hat{E}_j^{(kl)} \right)}{(\lambda_k + \lambda_l)^2}, \quad (21)$$

where the indices l, k are for expansions of the state matrix ρ in its eigenbases, i.e., $\rho = \sum_k \lambda_k |k\rangle \langle k|$ (same for the index l). It is worth noting that, as the covariance matrix scales as $1/N$, the average fidelity and average infidelity also scale as $1/N$.

Moreover, we also note that for the Bayesian mean estimation Eq. (11), one can also define a Bayesian covariance matrix, for a single Bayesian estimator, based on the posterior probability [3]. This Bayesian covariance matrix can be used as a measure of the uncertainty of the BME and is given by,

$$\text{Cov}_{\text{BME}} = \int P(\rho' | X) (\mathbf{c}' - \bar{\mathbf{c}})^T (\mathbf{c}' - \bar{\mathbf{c}}) d\rho', \quad (22)$$

where $\mathbf{c}'$ is a component vector of the candidate state ρ' , and $\bar{\mathbf{c}}$ is a Bayesian mean defined as,

$$\bar{\mathbf{c}} = \check{\mathbf{c}}_{\text{BME}} = \int P(\rho' | X) \mathbf{c}' d\rho'. \quad (23)$$

It is straightforward to verify that a component of this mean vector $\bar{\mathbf{c}}$ is related to the Bayesian mean estimate of the quantum state Eq. (11) via the relation: $\bar{c}_i = \text{Tr}(\check{\rho}_{\text{BME}} \hat{E}_i)$.

It should also be noted that the above error measures are dependent on the actual true states of the quantum system. Therefore, to get a state-independent benchmark for the state tomography, the errors should be averaged over all possible true states. Moreover, the difference between the Bayesian error (22) and the statistical error (which the Fisher information, the fidelity, and the trace distance rely on) is that the former measures the uncertainty between different candidate states in the Bayesian posterior distribution, while the statistical error measures the uncertainty between different estimates that arises from the fluctuation in different sets of measurement results.

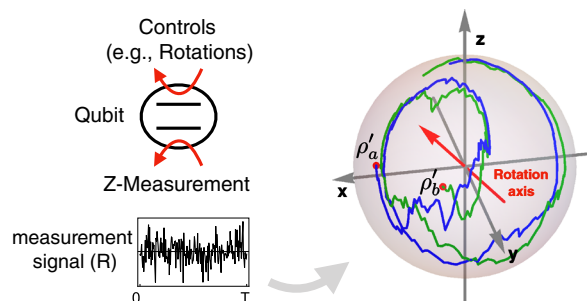


Fig. 1 Schematic diagram showing a single-qubit tomography and trial qubit trajectories on a Bloch sphere. Left: a two-level system (qubit) under a continuous measurement of the $\hat{Z}$ observable and a controlled rotation. The continuous measurement for a duration of time T gives a fluctuating record R , which is used in estimating the qubit's unknown initial state. Right: a Bloch sphere with two-qubit trajectories (blue and green), generated using the same record from the left, but with different guessed initial states (shown as two red dots on the Bloch sphere). The two trajectories give two different values of the probability density (likelihood) function, $P(R|\rho'_a)$ and $P(R|\rho'_b)$ (color figure online)

3 Quantum state tomography with continuous measurement

In the conventional linear inversion tomography, a number of different measurements for tomography of a full-rank state goes as $p = d^2 - 1$ where d is the dimension of the quantum state. An ensemble of N copies of quantum systems presumably prepared in the same state is divided into p sets, to which projective measurements of p observables are applied. Here, we propose a quantum state tomography using a continuous weak (finite strength) probe of only a single observable, in combination with a fixed unitary control, where an exact same measurement setup is applied to all N copies to obtain N time-continuous measurement records. Each of the measurement records contains information of all observables of the unknown initial state, and we can use the methods presented in the previous section (e.g., BME and MLE) to compute an estimator for the unknown state (see Fig. 1).

With our tomography strategy, one can perform the same experimental protocol to all N copies of the system, giving a set of measurement records $\{R_j\} = \{R_1, \dots, R_N\}$. Since the N records are independent of each other, the likelihood function for the measurement results is simply a product of Eq. (4) for all records R_j ,

$$\mathcal{L}(\rho'; \{R_j\}) = P(\{R_j\}|\rho') = \prod_j^N \text{Tr}[\mathcal{M}_{R_j} \rho' \mathcal{M}_{R_j}^\dagger], \quad (24)$$

where we have used ρ' as a candidate for the unknown initial state. The evolution from the unitary control dynamics is embedded in the operator $\mathcal{M}_R$, thus the next question is: how to choose the control so that the measurement results are informationally complete for the tomography of the unknown state. In the following subsections, we will elaborate on two techniques used in determining the controls, and present our investigation using three examples: a single-qubit tomography, a remote-qubit tomography, and a two-qubit tomography, all depending only on the measurement of an observable $\hat{Z}$ of a qubit, and Rabi oscillations applied locally to each of the qubits, with fixed angles and rotation frequencies.

3.1 Proposed methods to determine qubit controls

In this subsection, we present two approaches for determining the form of the system's control Hamiltonian that offers informationally complete measurement results. Since we require the unitary controls to map all system's observable information to the measured observable at some point during the measurement time, we need the unitary dynamics to dominate the measurement backaction. The first approach, only concerned about the unitary dynamics

of the system, is to use commutation relationships between all qubits' observables and possible terms in the control Hamiltonian to determine whether all observable information is accessible. The second approach, on the other hand, is to compute the Fisher information matrix of the unknown state parameters, for which the calculation includes the measurement backaction (with a weak measurement limit) to first order of the infinitesimal time. Depending on the complexity of the system's dynamics, both approaches in combination with numerical simulations using different parameter regimes can be used to find near-optimal controls (optimal in the state fidelity) for quantum state tomography.

For the first approach, considering the unitary dynamics after a finite time Δt , the system evolves from an initial state ρ_0 as,

$$\begin{aligned} \rho(\Delta t) &= e^{-i\hat{H}\Delta t} \rho_0 e^{+i\hat{H}\Delta t} \\ &= \frac{1}{d} \sum_j c_j^0 \hat{E}_j - i \left[\hat{H}, \frac{1}{d} \sum_j c_j^0 \hat{E}_j \right] \Delta t \\ &\quad - \frac{\Delta t^2}{2} \left[\hat{H}, \left[\hat{H}, \frac{1}{d} \sum_j c_j^0 \hat{E}_j \right] \right] + \dots, \end{aligned} \quad (25)$$

where we have substituted the state expansion defined in Eq. (5) for the initial state ρ_0 . The $\dots$ refers to terms with higher orders of Δt which contain three or more orders of commutators between the initial state ρ_0 and the system's Hamiltonian $\hat{H}$. Equation (25) shows that the evolved quantum state $\rho(\Delta t)$ is a function of the components c_j^0 of the initial state, and the evolved state is being measured continuously in time via the observable $\hat{Z}$ of the system.

If we expand the left-hand side of Eq. (25) as $\rho(\Delta t) = (1/d) \sum_k c_k^{\Delta t} \hat{E}_k$ and only collect terms that are the components for the measured observable $\hat{E}_k = \hat{Z}$ from both sides of the equation, we find a relationship,

$$c_{\hat{Z}}^{\Delta t} = c_{\hat{Z}}^0 + \Delta t \sum_{j'} \alpha_{j'} c_{j'}^0 + \Delta t^2 \sum_{j''} \alpha_{j''} c_{j''}^0 + \dots \quad (26)$$

where $c_{\hat{Z}}^{\Delta t} = \text{Tr}[\rho(\Delta t)\hat{Z}]$. Since $c_{\hat{Z}}^{\Delta t}$ is what determines the statistics of measurement results, an important task now is to find what are the non-vanishing components $c_{j'}^0, c_{j''}^0, \dots$, on the right-hand side of Eq. (26). These non-vanishing components indicate available information of the corresponding observables ($\hat{E}_{j'}, \hat{E}_{j''}, \dots$) of the initial state obtained from measuring $\hat{Z}$ at time Δt . From Eq. (25), we find that these observables need to satisfy the following relationships,

$$[\hat{H}, \hat{E}_{j'}] \propto \hat{Z}, \quad (27a)$$

$$[\hat{H}, [\hat{H}, \hat{E}_{j''}]] \propto \hat{Z}, \quad (27b)$$

$$[\hat{H}, \dots [\hat{H}, \hat{E}_{j''\dots}]] \propto \hat{Z}, \quad (27c)$$

at some point in the nested commutators, for their initial state components to effect measurement results (noting that the last term is for all higher order terms). We will implement these constraints to examples in the following subsections by constructing tables of commutators between system's observables and terms in the control Hamiltonian $\hat{H}$. The goal is to choose the control Hamiltonian so that every single observable $\hat{E}_k$ of the system satisfies one of the relationships in Eq. (27).

The commutator approach, however, only determines whether the statistics of measurement results can be a function of the observable components of the unknown initial state, and thus it cannot give a quantitative figure to show how much the information is actually obtainable from the records. We then use the second approach: to

calculate the Fisher information matrix Eq. (15) taking the system's unknown components c_j^0 as parameters to be estimated.

To derive the Fisher information matrix, we use the probability distribution $P(R|\rho_0) = \text{Tr}[\mathcal{M}_R \rho_0 \mathcal{M}_R^\dagger]$ of a measurement result R and replace the integral $\int dX$ in (15) with $\int dr_1 \dots dr_n$. However, as the Kraus operator $\mathcal{M}_R$ is a product of noncommuting matrices $U(dt)$ and $M(r_k)$ for $k = 1, \dots, n$, we need to simplify the calculation by expanding the measurement operator Eq. (3) for small dt ,

$$M(r_k) \approx \left(\frac{dt}{2\pi\tau} \right)^{1/4} e^{-\frac{(2r_k^2+1)dt}{4\tau}} \left(\hat{I} + \frac{r_k dt}{2\tau} \hat{Z} \right), \quad (28)$$

where $\hat{I} \equiv \hat{I}_1 \otimes \hat{I}_2 \dots \otimes \hat{I}_m$ is an identity matrix for a system of m qubits, and substitute the expansion to compute the probability distribution. The Fisher information matrix is a square matrix of order $(d^2 - 1)$, which is the maximum number of degrees of freedom of an unknown d -dimensional quantum state. The (j, j') -th element of the matrix is given by,

$$F_{j,j'} = \int \dots \int \frac{\partial_j P(R|\rho_0) \partial_{j'} P(R|\rho_0)}{P(R|\rho_0)} dr_1 \dots dr_n, \quad (29)$$

where j (and j') are indices for different components c_j^0 (and $c_{j'}^0$) of the system state ρ_0 and the partial derivatives are defined as $\partial_j \equiv \partial/\partial c_j^0$. This Fisher Information matrix elements will be functions of all parameters related to the system's dynamics and can be optimized numerically. We show in Appendix 1 the full derivation of the matrix for the two-qubit case, but the calculation can be generalized to the single-qubit tomography and more. The analytical results we have are expansions to first order in $1/\tau$, therefore, the results are applicable to the weak measurement limit $\tau \gg T$. We will use these two approaches of analyses to help us choose dynamical controls for the three examples in the following subsections.

3.2 Single-qubit state tomography

The first application of our approach is quantum state tomography for an unknown single-qubit state, where only the $\hat{Z}$ observable of the qubit can be measured and the qubit control is a simple Rabi oscillation along a fixed axis with constant frequency. We will analyze this scenario in a rather fine detail, since it can help us gain intuition about the information transfer from the Rabi controls, and how an estimation error varies in time as well as a number of measurement records (number of copies of unknown states). These intuitions can be applied to the applications for multiple qubits in Sects. 3.3 and 3.4.

In this single-qubit case, the control Hamiltonian is given by,

$$\hat{H} = \frac{\Omega}{2} (\mathbf{n} \cdot \hat{\sigma}), \quad (30)$$

where Ω is the oscillation rate, $\mathbf{n}$ is a unit vector describing the axis of rotation, and $\hat{\sigma} = \{\hat{X}, \hat{Y}, \hat{Z}\}$ is a vector of Pauli operators. Therefore, there are only two degrees of freedom to vary: the axis of rotation and the rotation rate.

We first use the analysis shown in Eqs. (25) and (26) and construct a commutator table in Fig. 2a. The table shows all possible commutators between observables of interest (on the first column) and terms that could exist in the control Hamiltonian (the first row). If one chooses the qubit's oscillation to be around the x -axis, i.e., $\hat{H} \propto \hat{X}$, then we find that a commutator $[\hat{H}, \hat{E}_{j'}] \propto \hat{Z}$ in (25) has only one possibility which is $\hat{E}_{j'} = \hat{Y}$ [shown as the intersection of dashed boxes in Fig. 2a]. This can be read as the information about $\hat{Y}$ can be transferred to the measured $\hat{Z}$ via the $\hat{X}$ rotation. For the second-order term, we find that $[\hat{X}, [\hat{X}, \hat{E}_{j''}]] \propto \hat{Z}$ is satisfied only when $\hat{E}_{j''} = \hat{Z}$, so it tells us nothing new. Thus, the measured component $c_{\hat{Z}}^{dt}$ in (26) can only be a function of initial state components of $\hat{Z}$ and $\hat{Y}$, but not $\hat{X}$. To obtain knowledge of all three qubit components, we require the Rabi control to be a combination of at least two of the three observables. For example, applying a control as a combination of

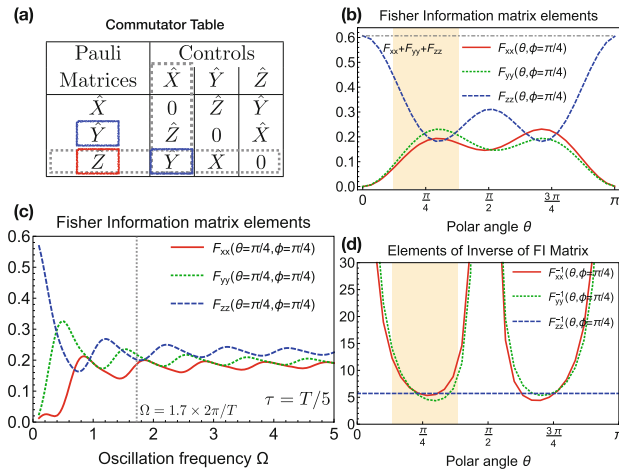


Fig. 2 Analysis of possible controls for the single-qubit state tomography. **a** The commutator table shows results of all commutators between three observables on the first column and three possible terms in the system's (control) Hamiltonian on the first row. We omit any complex proportional factors of the commutators. The dotted boxes are to highlight the control ($\hat{H} \propto \hat{X}$) and the measured observable ($\hat{Z}$), while the solid boxes are for the observables that can be estimated given the particular control. **b, d** Diagonal elements of the Fisher information matrix and its inverse matrix are plotted with different polar angles θ of the Rabi oscillation axis. The rotation rate is fixed at $\Omega = 2\pi(1.7)/T$ where T is the duration of the experiment. The yellow bands show the range of θ one can use for the qubit tomography. **c** The diagonal elements of the Fisher information matrix with varying Ω , keeping $\theta = \phi = \pi/4$ fixed. The units of Ω are $2\pi/T$ and we used $\tau = T/5$ for all plots

$\hat{X}$ and $\hat{Z}$ will lead to the transfer of information of $\hat{X}$ and $\hat{Y}$ to the observed component of $\hat{Z}$, which we can write arrow diagrams for as,

$$\begin{aligned} \hat{X} &\xrightarrow{\hat{Z}} \hat{Y} \xrightarrow{\hat{X}} \hat{Z}, \\ \hat{Y} &\xrightarrow{\hat{X}} \hat{Z}. \end{aligned}$$

We note that the arrow refers to an information transfer, and observables above an arrow refers to terms in the Hamiltonian that are responsible for such transfer. Therefore, we can obtain the three qubit coordinates information from only measuring $\hat{Z}$ as long as the Rabi rotation axis is not aligned completely with $\hat{Z}$ (or $\hat{X}$, or $\hat{Y}$).

Once the observables in the control Hamiltonian are determined, we then compute the Fisher information matrix (29) for the weak measurement limit to see finer effects of the angle and frequency of the rotation. We show in Fig. 2b–d the diagonal elements F_{xx} , F_{yy} and F_{zz} for different angles of the Rabi axis and oscillation frequencies. In the panels (b) and (d), we vary the polar angle θ of the Rabi axis, keeping the azimuthal angle fixed at $\phi = \pi/4$. As θ grows, the information of z decreases from its maximum value, while F_{xx} and F_{yy} increase and reach roughly the same level as F_{zz} at around $\theta \sim \pi/4 \pm \pi/8$. Figure 2c shows diagonal elements for the inverse of the Fisher information matrix $C \equiv F^{-1}$ which approximate the lower bound for the variances of the estimated parameters.

We then choose the oscillation axis to be at an angle $\theta = \phi = \pi/4$ and see how the Fisher information changes with varying Rabi oscillation rate Ω in Fig. 2c. Increasing the frequency results in changing information access from only $\hat{Z}$ to all $\hat{X}$, $\hat{Y}$, and $\hat{Z}$, and the trend is practically unchanged after Ω reaches $\sim 1.0 \times 2\pi/T$. We note that changing τ (inversely proportional to the measurement strength) only effects the amount of total information gain during the course of measurement time T , but does not effect the relative division of information between the three parameters.

To investigate the quality of the state tomography, we numerically simulate quantum trajectories along with their measurement records, and analyze the Bayesian estimators and the maximum likelihood estimators for different measurement and control parameters (see Appendix 5 for more detail on the numerical simulation). We consider the

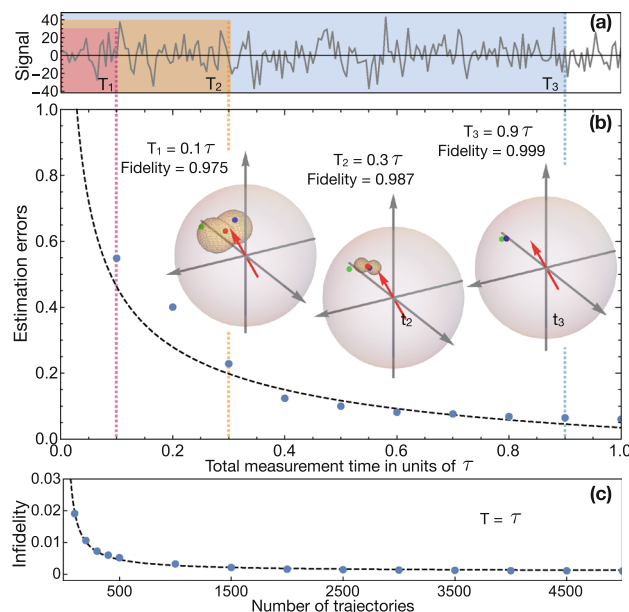


Fig. 3 Estimation errors of the unknown single-qubit state. **a** An example of a measurement record as a function of time. **(b** data points) The errors of the estimate calculated from the square root of the trace of the Bayesian covariance matrix Eq. (22) for different total measurement (data collection) times with $N = 5000$. The fitting curve is $\sim 1/\sqrt{T}$. The three Bloch spheres, at $T_1 = 0.1\tau$, $T_2 = 0.3\tau$ and $T_3 = 0.9\tau$, show the Bayesian estimators (Red), the maximum likelihood estimators (Blue), the unknown true state (Green), and the error ellipsoids from Eq. (31). **c** The infidelity between the Bayesian estimators and the true state (averaged over 300 repetitive sets) is shown to scale as $1/N$ (the fitting dashed curve) with $T = \tau$ (color figure online)

root mean square error (RMSE) calculated from the trace of a Bayesian covariance matrix, Eq. (22). This covariance matrix C can be used to construct an illustrative *error ellipsoid* for the Bayesian mean estimate. The ellipsoid is described by,

$$r^2 = \mathbf{n}^\dagger \cdot C \cdot \mathbf{n}, \quad (31)$$

where r is the radius of the ellipsoid and is dependent on the direction of a unit vector $\mathbf{n}$ pointing out from the Bayesian estimator in the Bloch sphere.

We show in Fig. 3 the RMSEs, the error ellipsoids, and the infidelity of the Bayesian estimators for different measurement times T and numbers of measurement records N . As the total time increases, the estimation quality gets better (RMSE decreases as $1/\sqrt{T}$) and the size of the error ellipsoid shrinks. In the bottom panel, we show the infidelity (averaged over 300 sets for statistically stable results) for different N confirming the $1/N$ scaling as predicted in Sect. 2.2.2. From these results, we learn that the total measurement time T should be at least a few τ given that N is large enough ($N > 2000$), so that we can be sure the information about the initial states are collected via the measurement to the desired accuracy. We stress that making the total measurement time T much longer than that will not give further information about the initial state, because the measurement signal becomes uncorrelated with the initial state.

We also perform the numerical simulation for varying Rabi oscillation frequencies, to confirm the trend seen in Fig. 2c for the Rabi angle $\theta = \phi = \pi/4$. As expected, see Fig. 4a, at zero frequency $\Omega = 0$, the $\hat{Z}$ component of the unknown qubit is the only estimated parameter with a very small error bar. As the oscillation frequency reaches $\Omega \sim 0.5 \times 2\pi/T$, the estimation quality for $\hat{X}$ and $\hat{Y}$ components attains the similar level of precision as the $\hat{Z}$ component. One interesting feature is that the estimation error stays unchanged for $\Omega \gtrsim 1.0 \times 2\pi/T$, similar to what was found in Fig. 2c. This can be explained that, because the amount of total information is fixed by T , τ and N , the role of the rotation is simply to propagate information of all other qubit coordinates to the

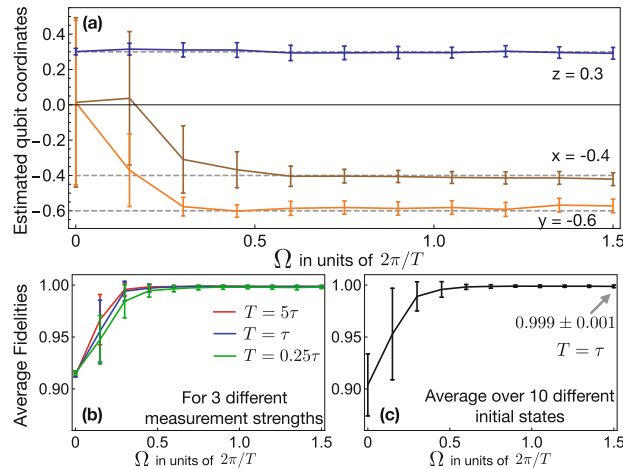


Fig. 4 Estimated single-qubit coordinates and average fidelities for a fixed Rabi axis ($\theta = \phi = \pi/4$) and different values of the Rabi oscillation rate. **a** The Bayesian estimates of x , y , and z coordinates of the unknown true state $\rho_0 = (1/2)(\hat{I} - 0.4\hat{X} - 0.6\hat{Y} + 0.3\hat{Z})$ using $N = 5000$. The error bars of the estimates are calculated from the square root of the trace of the Bayesian covariance matrix Eq. (22). **b** Average fidelities for 100 simulated Bayesian estimators, for different values of τ and Ω , where each estimator is computed from $N = 5000$ measurement records. The test initial state is the same as in **a**. **c** Average fidelities for different initial states chosen randomly (see Appendix 6). The errors can be understood as coming from within the same initial state (as seen in **b** for example) and from different initial states. At $\Omega = 2\pi(1.5)/T$, the average fidelity reads 0.999 ± 0.001

measured observable. Once the oscillation rate reaches a threshold $\sim 1.0 \times 2\pi/T$, the estimation quality among all observables stays about the same. In Fig. 4b, we show the average fidelities for different values of Ω and τ . As Ω increases, the average fidelities increase from the value around 0.91, where only the z coordinate is estimated, to the value close to 1 at a high oscillation rate. The change in τ results in competing effects between the measurement rate and the oscillation rate, i.e., given fixed T , the stronger measurement (red) reaches a high value of state fidelity faster than the weak measurement. However, we note that if the measurement rate is too strong, i.e., much stronger than the oscillation rate $\Omega \ll 1/\tau$, then the measured information will only be about the $\hat{Z}$ component, and the fidelity will be bound at $F(\rho_0, (1/2)(\hat{I} + 0.3\hat{Z})) = 0.918$ for this particular ρ_0 . In Fig. 4c, we show that the quality of the estimate is independent of the initial states, where we simulate the estimation using 10 randomly chosen initial states.

3.3 Remote-qubit state tomography

As the second application of our method, we consider quantum state tomography for an unknown single qubit that cannot be measured directly. Instead, the qubit information can only be accessed through an interaction with another qubit of which only its $\hat{Z}$ is measured. Let us assume that the interaction between the two qubits is described by a $\hat{X} \otimes \hat{X}$ (capacitive) coupling term, and the possible controls are Rabi drives applied locally on each qubit [see Fig. 5a], with fixed axes and oscillation rates. The two-qubit Hamiltonian for this remote-qubit example (also for the two-qubit example in the next section) is then given by,

$$\hat{H} = \frac{g}{2} \hat{X} \otimes \hat{X} + \frac{\Omega_a}{2} (\mathbf{n}_a \cdot \hat{\boldsymbol{\sigma}}) \otimes \hat{I} + \frac{\Omega_u}{2} \hat{I} \otimes (\mathbf{n}_u \cdot \hat{\boldsymbol{\sigma}}), \quad (32)$$

where the first term describes the coupling between the two qubits with a coupling strength g , and the rest is the local Rabi oscillations with frequencies Ω_a and Ω_u , for the *ancilla* qubit and the *unknown* qubit, respectively. In this remote-qubit protocol, we assume the first qubit [left one in Fig. 5a] is a measured qubit with a known prepared state, and the second (right) qubit is an unknown qubit.

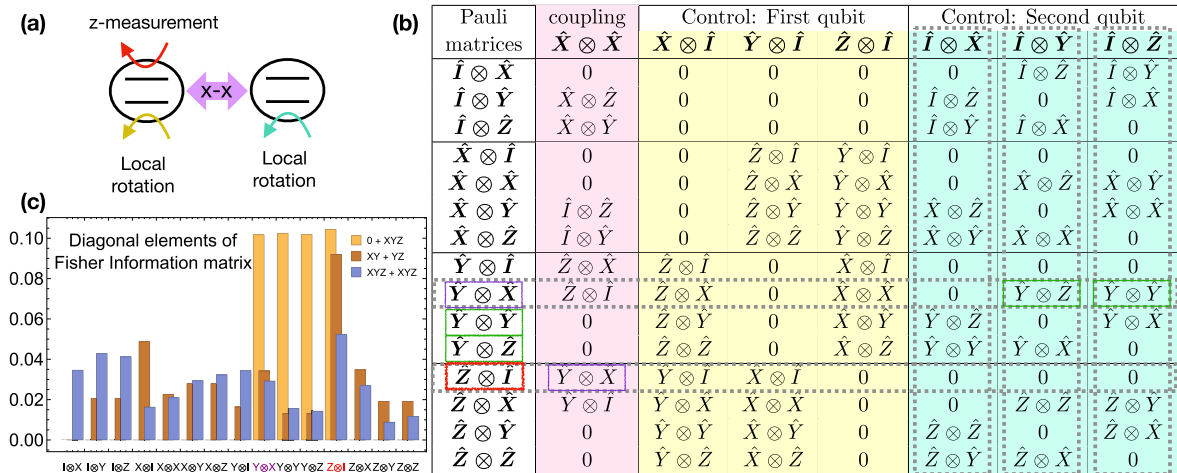


Fig. 5 Analysis of controls for the remote-qubit and two-qubit state tomography. **a** Schematic diagrams for the two-qubit setup: qubits interact with each other via $\hat{X} \otimes \hat{X}$ coupling and Rabi controls are applied locally on each of the qubits. For the remote-qubit tomography, the first qubit is known and acts as an ancilla qubit, while the second qubit is unknown. For the two-qubit tomography, both qubits are unknown. **b** A commutator table shows the commutation relationships between the 15 observables on the first column and 7 terms (terms that can be included in the Hamiltonian of the qubits, Eq. (32)) in the first row. Any proportional complex constants are omitted. The colored solid boxes show four estimated observables given that the control is in the 0+XYZ setting (see text). **c** Fisher information matrix elements (diagonal elements only) calculated using $g = \Omega = 1.5 \times 2\pi/T$ and $\tau = T/5$. The colored axis labels in panel (c) are to emphasize the measured observable $\hat{Z} \otimes \hat{I}$ and the observable $\hat{Y} \otimes \hat{X}$ of which its information gets transfer first via the qubit-qubit coupling

In comparison to the single-qubit tomography in the previous subsection, one can think of the unknown single qubit in this case is measured via its $\hat{X}$ observable (instead of $\hat{Z}$). This is because the coupling term effectively provides a von Neumann measurement of $\hat{X}_u$ of the unknown qubit, using $\hat{X}_a$ of the ancilla qubit as a pointer. The $\hat{X}$ component of the unknown qubit determines the rotation speed around the x -axis of the ancilla qubit, which can then be mapped to the measured $\hat{Z}$ component of the ancilla qubit. Therefore, a straightforward option for qubit controls is to set the Rabi axis for the unknown qubit at $\theta_u = \phi_u = \pi/4$, as we found in the single-qubit example, and then initialize the ancilla qubit state at $y_a^0 = \text{Tr}(\rho_0 \hat{Y} \otimes \hat{I}) = 1$ with no Rabi drive $\Omega_a = 0$.

We can see that this straightforward protocol works by looking at the commutator table and the Fisher information matrix elements shown in Fig. 5b, c, respectively. If we choose the Rabi axis for the unknown qubit to be in the direction described by $\theta_u = \phi_u = \pi/4$ and no oscillation on the ancilla qubit, then the Hamiltonian is $\hat{H} = (g/2)\hat{X} \otimes \hat{X} + (\Omega_u/4)(\hat{I} \otimes \hat{X} + \hat{I} \otimes \hat{Y} + \sqrt{2}\hat{I} \otimes \hat{Z})$. Using the analysis described in Eqs. (25) and (26), also explained in Fig. 5 and its caption, we find that the information of state components that can be transferred to the measured observable $\hat{Z} \otimes \hat{I}$ via this unitary dynamics are: $\hat{Y} \otimes \hat{X}$, $\hat{Y} \otimes \hat{Y}$, and $\hat{Y} \otimes \hat{Z}$. We can write diagrams to explain the information transfer as,

$$\hat{Y} \otimes \hat{X} \xrightarrow{\hat{X} \otimes \hat{X}} \hat{Z} \otimes \hat{I}, \quad (33)$$

$$\hat{Y} \otimes \hat{Y} \xrightarrow{\hat{I} \otimes \hat{Z}} \hat{Y} \otimes \hat{X} \xrightarrow{\hat{X} \otimes \hat{X}} \hat{Z} \otimes \hat{I}, \quad (34)$$

$$\hat{Y} \otimes \hat{Z} \xrightarrow{\hat{I} \otimes \hat{Y}} \hat{Y} \otimes \hat{X} \xrightarrow{\hat{X} \otimes \hat{X}} \hat{Z} \otimes \hat{I}, \quad (35)$$

which can be interpreted as: the $\hat{Y} \otimes \hat{X}$ component information is transferred to the measured $\hat{Z} \otimes \hat{I}$ component via the $\hat{X} \otimes \hat{X}$ coupling, and the component information of $\hat{Y} \otimes \hat{Y}$ (or $\hat{Y} \otimes \hat{Z}$) is transferred to $\hat{Z} \otimes \hat{I}$ via the control term $\hat{I} \otimes \hat{Z}$ (or $\hat{I} \otimes \hat{Y}$) as well as the coupling term. These three observables $\hat{Y} \otimes \hat{X}$, $\hat{Y} \otimes \hat{Y}$ and $\hat{Y} \otimes \hat{Z}$ are sufficient for the

estimation of three coordinates of the unknown qubit, given that the y coordinate of the ancilla state is initialized in a non-zero value $y_a^0 \neq 0$. The analysis of the Fisher information matrix in the weak measurement limit also gives non-zero elements for the three observables (in addition to the measured $\hat{Z} \otimes \hat{I}$). This is shown as the yellow bars for “ $0 + XYZ$ ” setting in Fig. 5c. For convenience, we use the code “ $A + B$ ” to represent a set of Rabi controls, where A is a combination of X , Y , and Z describing a Rabi axis of the first qubit with non-zero components in $\hat{X}$, $\hat{Y}$, and $\hat{Z}$, and the same goes for B for the second qubit’s rotation axis. If $A = 0$, it refers to no Rabi oscillation on the first qubit.

We numerically simulate the tomography process to test the quality of the estimation using the proposed $0 + XYZ$ setting. Similar to the previous single-qubit example, we compute a Bayesian estimated state from the probabilities of the trial states given the measurement signals (see “Appendix C” for the detail on the numerical simulation). To analyze the quality of the estimation, we simulate 100 estimators (each estimator is computed from $N = 5000$ measurement records) for each set of the control parameters of interest and compute RMSEs from the 100 estimators compared with a given true state.

The results are shown in Fig. 6a, presenting the RMSEs for the estimation of the 16 two-qubit density matrix components. Large values of RMSE mean that the elements are poorly estimated, or simply that little information about them was obtained from the measurement records. The vanishing errors, on the other hand, represent absolute accuracy, which in this case happens to the elements that are known to have zero values. For example, choosing $y_a^0 = 1$ (meaning that $x_a^0 = z_a^0 = 0$), we know with certainty that the elements other than $\hat{I} \otimes \hat{X}$, $\hat{I} \otimes \hat{Y}$, $\hat{I} \otimes \hat{Z}$, $\hat{Y} \otimes \hat{X}$, $\hat{Y} \otimes \hat{Y}$, and $\hat{Y} \otimes \hat{Z}$ of the initial state have zero values. The take-away messages from the numerical data shown in Fig. 6a–c are that: (1) the $0 + XYZ$ setting works for the remote-qubit state estimation especially when $y_a^0 = 1$ for the ancilla state, (2) the other two settings “ $XY + YZ$ ” and “ $XYZ + XYZ$ ” used in the full two-qubit tomography (discussed later in the next section) can also be used to estimate the remote unknown qubit with comparable precision independent of the ancilla initial state, and (3) the parameters g , Ω are best in the range of 1.0–1.5 in units of $2\pi/T$.

We also show average fidelity for different unknown states in Fig. 6d, where the ten randomly chosen unknown states are listed in the “Appendix D”. The values of average fidelity are all really close to 1, where some of them are in the range of 0.99. The latter corresponds to the states that have high purity (see first and seventh data points in the plot, which are the states with purity 1 and 0.9899). This slightly low fidelity for high-purity states is a result of a boundary effect, where the actual true state lies on the boundary of the state space used in the Bayesian estimation. Therefore, with the uniform prior probability $q(\rho)$ over the entire full-rank density matrix space, a Bayesian estimate will always be inside the boundary causing an error in the estimate that scales with T and N slower than the true states that are inside the state space. However, as mentioned in Sect. 2.2, we can improve the quality of the estimate using the prior probability to be a uniform distribution only over the pure-state subspace. We show in Fig. 6d, with a red dot as a Bayesian estimate using this pure-state-only prior probability.

3.4 Two-qubit state tomography

As the third application, we consider a full two-qubit tomography where all of the 15 observables of the coupled two qubits can be accessed by only continuously measuring the observable $\hat{Z} \otimes \hat{I}$ and applying the fixed qubit controls. From the commutator table in Fig. 5b, one can convince oneself that a “ $XY + YZ$ ” control, where $\mathbf{n}_a = (1/\sqrt{2})(\hat{X} + \hat{Y})$ and $\mathbf{n}_u = (1/\sqrt{2})(\hat{Y} + \hat{Z})$, and a more overall “ $XYZ + XYZ$ ” control, where $\mathbf{n}_a = (1/2)(\hat{X} + \hat{Y} + \sqrt{2}\hat{Z})$ and $\mathbf{n}_u = (1/\sqrt{2})(\hat{X} + \hat{Y} + \sqrt{2}\hat{Z})$, lead to accessing information of all 15 observables. The Fisher information matrix analysis for the two settings are shown in Fig. 5c. In this subsection, we use these sets of controls to examine their performance in estimating full unknown two-qubit states, including pure states, mixed states, and non-separable states.

In contrast to the three dimensional real-number space of the Bloch sphere representing single-qubit states in the previous two examples, the state space of the two-qubit system, is a space of 15 real numbers with finite ranges. This large state space makes the Bayesian estimation a non-preferable method, as one needs to calculate the posterior probability density from Eq (24) for each of candidate states. To get a similar accuracy as in the single-qubit case,

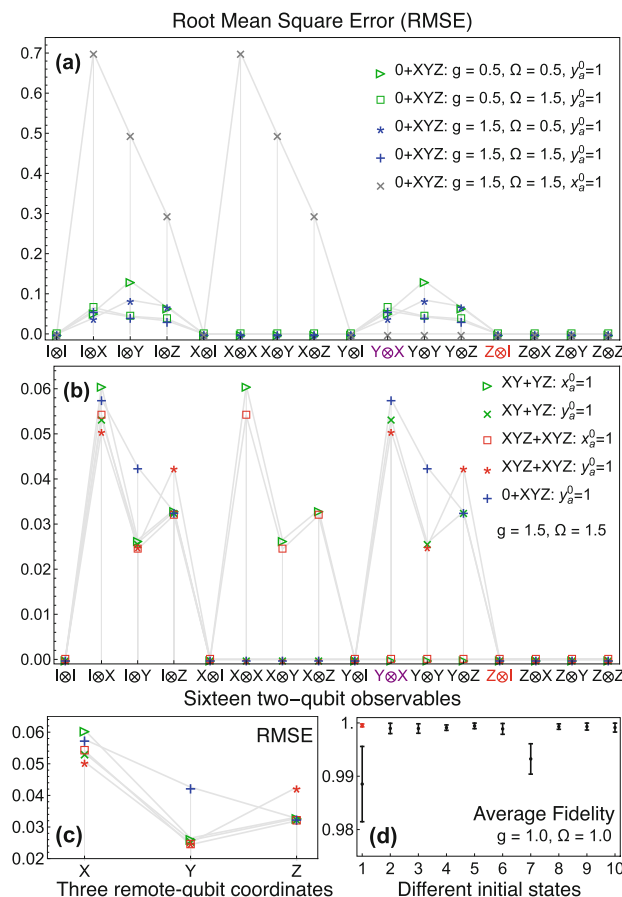


Fig. 6 Root mean square errors for the remote-qubit tomography. **a** The errors of the estimated 16 remote-and-ancilla-qubit coordinates for the control settings where the Rabi oscillation is only applied on the remote qubit, with different values of g , $\Omega_u = \Omega$, and the initial states of the ancilla qubit (denoted by x_a^0, y_a^0). **b, c** The estimation errors for successful control settings using the coupling strength $g = 1.5$ and the qubit oscillation frequency $\Omega_a = \Omega_u = \Omega = 1.5$. **b** shows the errors of all 16 elements, while **c** shows only the errors of estimated remote-qubit coordinates. Note that the vertical scale of **a** is an order of magnitude lower than **(b)**. The abbreviations $0 + XYZ$, $XY + YZ$, etc., denote the controls and are defined in the text. Data points in **a–c** are numerically simulated with the same unknown remote-qubit state $\rho_u = (1/2)(\hat{I} + 0.7\hat{X} - 0.5\hat{Y} + 0.3\hat{Z})$ using $N = 5000$ and the same measurement strength $\tau = T/5$. We test the remote-qubit state estimation with different initial states using the $XY + YZ$ setting with $g = \Omega = 1.0$ and show average fidelities with their error bars in **d** (see text for the red data point). The units of g and Ω are $2\pi/T$

we need much more than 10^8 trial two-qubit states distributed uniformly with a specific measure (or with some prior distribution), only to compute one Bayesian estimator (see more detail in Appendix C). Therefore, to minimize the computation overhead, we instead maximize the likelihood function Eq. (24) over the valid state space using a search algorithm. The estimator in this two-qubit tomography protocol is then the maximum likelihood estimator defined in Eq. (8).

To guarantee valid physical quantum states of the MLE, we need to specify a real-number space to be used with a numerical search algorithm. For an expansion of components of a two-qubit state given by,

$$\rho = \frac{1}{4} \sum_{i,j} c_{ij} \hat{\sigma}_i \otimes \hat{\sigma}_j,$$

where the indices are $i, j = 0, 1, 2, 3$, and the Pauli matrices $\hat{\sigma}_j$ for $j = 0, 1, 2$, and 3 , represent $\hat{I}$, $\hat{X}$, $\hat{Y}$, and $\hat{Z}$, respectively. We then have $c_{00} = 1$ and the rest are real numbers. Following the calculation shown in [14], if we write the coefficients c_{ij} in a matrix form,

$$D = \begin{pmatrix} 1 & c_{01} & c_{02} & c_{03} \\ c_{10} & c_{11} & c_{12} & c_{13} \\ c_{20} & c_{21} & c_{22} & c_{23} \\ c_{30} & c_{31} & c_{32} & c_{33} \end{pmatrix} \equiv \begin{pmatrix} 1 & \vec{v}^\dagger \\ \vec{u} & B \end{pmatrix}, \quad (36)$$

where we have defined the vector $\vec{v}$, $\vec{u}$, and the matrix B on the right-hand side, the positivity condition for the two-qubit state is given by three constraints,

$$0 \leq 4 - \|D\|^2, \quad (37)$$

$$0 \leq 2(\vec{u}^\dagger B \vec{v} - \det B) - (\|D\|^2 - 2), \quad (38)$$

$$0 \leq 8(\vec{u}^\dagger B \vec{v} - \det B) + (\|D\|^2 - 2)^2 + 8\vec{u}^\dagger \tilde{B} \vec{v} - 4(\|u\|^2 \|v\|^2 + \|\vec{u}^\dagger B\|^2 + \|B \vec{v}\|^2 + \|\tilde{B}\|^2), \quad (39)$$

where $\tilde{B}$ is the cofactor matrix of B , and $\|A\| = \text{Tr}[A^\dagger A]$ using the Hilbert–Schmidt inner product. We note that, for the numerical search, we use the Differential Evolution algorithm to find a state that maximizes Eq. (24), with 50 initial random searches to reach a reasonable convergence (more detail of the numerical method in Appendix 5).

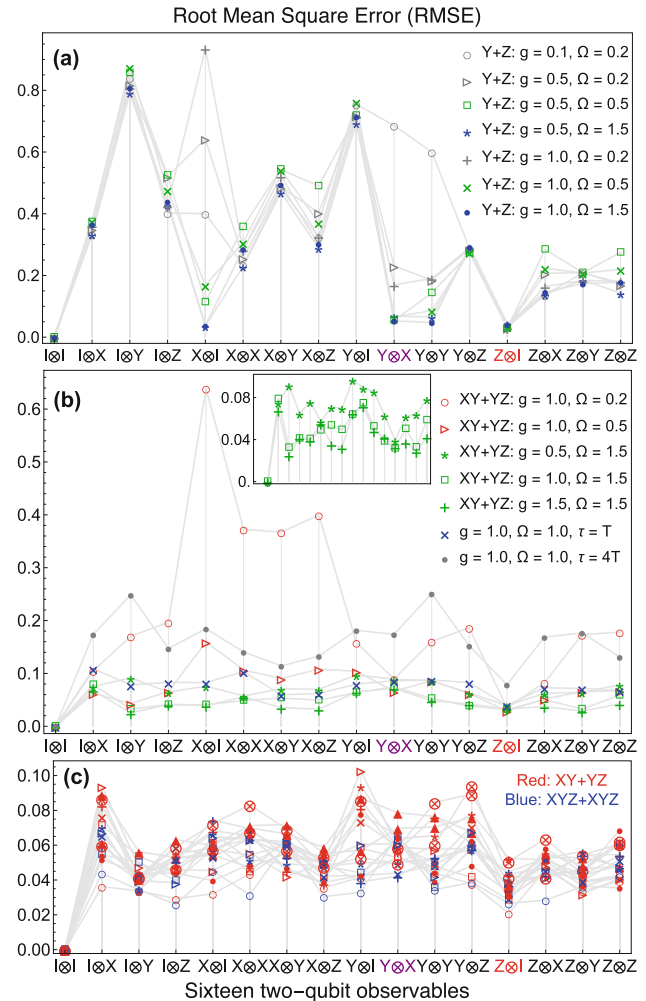
For comparison purpose, we first show numerical results for the two-qubit state tomography for a poorly chosen set of controls using a control setting “ $Y + Z$,” where the Rabi oscillations of the first and second qubits are around the y -axis and z -axis, respectively. The RMSEs for this setting are shown in Fig. 7a for different values of control parameters g and Ω . Only three components for $\hat{X} \otimes \hat{I}$, $\hat{Y} \otimes \hat{X}$, and $\hat{Y} \otimes \hat{Y}$ that show significant improvement in the errors, which are reduced to less than 0.1 while increasing the parameters g and Ω to be around $\sim 1.0 - 1.5$, in units of $2\pi/T$. This agrees with a prediction from the two-qubit commutator table in Fig. 5b, where one can deduce that the unitary dynamics allow access to the information of components for only four observables: $\hat{X} \otimes \hat{I}$, $\hat{Y} \otimes \hat{X}$, $\hat{Y} \otimes \hat{Y}$, and $\hat{Z} \otimes \hat{I}$.

To estimate all the 15 elements of the unknown two-qubit state concurrently, we use the control settings: $XY + YZ$ or $XYZ + XYZ$, mentioned in the beginning of this section. Since the $XY + YZ$ control is simpler to numerically simulate than the latter, we choose this control as our preferred setting; however, the results are similar for the $XYZ + XYZ$ control. We show in Fig. 7b the estimation errors for different control parameter values. As the coupling strength and the oscillation frequency grow, the estimation errors for all 15 elements are significantly reduced to values less than 0.1, at around $g \sim 1.0 - 1.5$ and $\Omega \sim 1.0 - 1.5$ in units of $2\pi/T$. In the plot, we also show the effect of the measurement strength by changing τ keeping the controls fixed, as seen in the last two data sets of the figure legend. As expected, weakening the measurement strength simply lifts up the baseline of the errors of all 15 elements, which can be interpreted that the relative information distributed among them is approximately the same, and only the total information of all parameters changes.

From these numerical results in Fig. 7b, we find that the quality of the estimation is well saturated at around $g \sim \Omega \sim 1.0 - 1.5$ in units of $2\pi/T$, similar ranges to what we have found in the previous single-qubit and remote-qubit cases. We then perform the numerical tests for different randomly chosen unknown states, containing product states and non-separable state mixtures (e.g., Werner states) with varied purities, using $g = \Omega = 1.5 \times 2\pi/T$ for both $XY + YZ$ and $XYZ + XYZ$ settings. As shown in Fig. 7c, the RMSEs for nine different initial states per one setting are roughly the same, varying in the small range of 0.02–0.1. We also calculate the average fidelity for all these eighteen cases giving a value of 0.984 with an error bar of 0.003. We note that these results are from maximum likelihood estimators using only $N = 4000$, and thus the fidelity should be better if one have larger numbers of N used in the estimation.

We note that both $XY + YZ$ and $XYZ + XYZ$ settings can also be used for the remote-qubit tomography, since the information about the unknown qubit coordinates is embedded in the components of all twelve observables:

Fig. 7 Root mean square errors for the full two-qubit state tomography. **a** The estimation errors for the $Y+Z$ control setting with different values of g and Ω . The initial state is fixed at a product state between $\rho_1 = (1/2)(\hat{I} - 0.4\hat{X} - 0.75\hat{Y} + 0.5\hat{Z})$ and $\rho_2 = (1/2)(\hat{I} + 0.6\hat{X} - 0.5\hat{Y} + 0.6\hat{Z})$, where $\tau = T/5$, and $N = 4000$. The RMSEs are calculated from 50 sets of maximum likelihood estimators for each set of parameters. **b** The estimation errors for the $XY+YZ$ control, with the same initial unknown state and number of sets as in **a**. We use $\tau = T/5$ except for last two sets of data, where we use $\tau = T$ and $\tau = 4T$. The inset shows small improvement from increasing g and keeping $\Omega = 1.5 \times 2\pi/T$. **c** The errors for the two successful settings, $XY + YZ$ and $XYZ + XYZ$, for 9 different initial states for each setting, using $g = \Omega = 1.5 \times 2\pi/T$ and $\tau = T/5$. The average fidelity of all 18 sets is 0.984 ± 0.003



$\hat{I} \otimes \hat{X}, \hat{I} \otimes \hat{Y}, \hat{I} \otimes \hat{Z}, \dots, \hat{Z} \otimes \hat{Y}, \hat{Z} \otimes \hat{Z}$, given that the initial state of the ancilla qubit is known. We show the results of numerical simulation in Fig. 6b. Both settings perform equally well independent of the initial ancilla states ($x_a^0 = 1, y_a^0 = 1$, or a mixture of non-zero x_a^0 and y_a^0).

4 Discussion and conclusion

We have analyzed our proposed method for quantum state tomography using continuous weak measurement where a system of qubits with an unknown state is measured with a resource limitation. We consider in particular a system of interacting qubits that can be measured only through one of its observables, and the qubit controls are simple Rabi oscillations applied locally on each qubit. We find measurement and control settings, consisting of values of the measurement strength, the oscillation axes, the oscillation rates, and the coupling rates, that allows access to the information of all unknown components of qubit observables in one measurement run.

The proposed tomographic scheme is analyzed using three proof-of-principle applications: (a) single qubit, (b) remote qubit, and (c) full two-qubit tomography, with varying control settings and estimation methods (Bayesian mean and maximum likelihood estimation). The control settings are chosen based on the analysis of the commutation relations as well as the Fisher information matrix of the estimated components. In the single-qubit tomography, a

continuous probe of the $\hat{Z}$ observable and a Rabi oscillation with an axis of rotation off-parallel from the measured z -axis are sufficient to extract information of the three qubit coordinates. For the remote-qubit and the full two-qubit tomography, a continuous measurement of the $\hat{Z} \otimes \hat{I}$ observable is sufficient to extract the qubits' information if it is combined with specific local qubit controls. The axes of local qubit controls we found most efficient are: (a) the rotation axis of the measured qubit is aligned diagonally in $x - y$ plane and the rotation axis of the unmeasured qubit is aligned diagonally in $y - z$ plane (denoted as $XY + YZ$ setting), and (b) the rotation axes of both qubits are off-parallel from all x, y, z axes (denoted as $XYZ + XYZ$ setting).

Once the axes of the oscillations are fixed, from our analysis, we found that the rotation rates need to be at least $1.0 \times 2\pi/T$, to assure uniformly extracted information of all qubits' components. In other words, the oscillation rates should be fast enough that a net rotated angle (ignoring the measurement backaction) covers at least a full 2π rotation during the total measurement time T . The coupling rate of the two qubits is found to be optimal at the same value as the local Rabi oscillation rate. Moreover, the measurement strength should be strong enough that all information is collected from the measurement, but not too strong that it results in solely extracting the information of the measured observable and less information about other observables. That is, the total measurement time should be about one or two times the characteristic measurement time τ .

As a summary, we have demonstrated that quantum state tomography can be achieved using only limited measurement and fixed control resources, if the control settings are chosen wisely. The reason why this scheme works can be explained as follows. The unitary dynamics and the measurement backaction play the role of a random sampler, effectively mapping the measured observable (in this case is the $\hat{Z}$ observable of a qubit) to a series of random observables that can sample the unknown state in a variety of different bases continuously in time. We stress the relative simplicity of this method, which should be practical for implementation in the laboratory. As an outlook for future work, it is natural to investigate the optimality of the methods, and their state estimation performance in comparison to other conventional methods, such as the linear inversion, taking into account the practical finite-strength quantum measurement. Moreover, the issue of optimality of these tomographic schemes also poses many interesting geometrical questions whether they can achieve a uniform measurement sampling of a quantum state.

Acknowledgements This work was supported by US Army Research Office Grant no. W911NF-18-1-0178, and by the National Science Foundation grant DMR-1809343. AC acknowledges support by the Australian Research Council Centre of Excellence CE170100012. TC acknowledges the Sri Trang Thong scholarship for the undergraduate program at Mahidol University and for the summer research at University of Rochester.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Average fidelity and covariance relation

In this Appendix section, we present a detailed calculation for the fidelity and the average fidelity in relation to the covariance matrix of the estimated state coefficients. As the number of independent measurement results N is large, the error in the estimates $\check{c}_i$'s of the coefficients c_i 's are small. In this case, the fidelity $F(\check{\rho}, \rho)$ can be computed from the covariance of $\check{c}_i$'s. We first write an estimated state in terms of the error of the state coefficients,

$$\check{\rho} = \rho + \sum_i \Delta c_i \hat{E}_i, \quad (40)$$

where $\Delta c_i = \check{c}_i - c_i$. Then, when $\Delta c_i \ll 1$, we have

$$\begin{aligned} \sqrt{\check{\rho}} \check{\rho} \sqrt{\check{\rho}} &= \sqrt{\rho^2 + \sum_i \Delta c_i \sqrt{\rho} \hat{E}_i \sqrt{\rho}} \\ &\approx \rho + \sum_i \Delta c_i \vartheta_i + \sum_{ij} \Delta c_i \Delta c_j \chi_{ij}, \end{aligned} \quad (41)$$

where ϑ_i 's and χ_{ij} 's are operators that satisfy the following Lyapunov equations

$$\begin{aligned}\vartheta_i \rho + \rho \vartheta_i &= \sqrt{\rho} \hat{E}_i \sqrt{\rho}, \quad \forall i, j, \\ \chi_{ij} \rho + \rho \chi_{ij} &= \vartheta_i \vartheta_j + \vartheta_j \vartheta_i.\end{aligned}\quad (42)$$

The fidelity between ρ and $\check{\rho}$ becomes

$$F(\check{\rho}, \rho) = 1 + \sum_i \Delta c_i \text{Tr} \vartheta_i + \sum_{ij} \Delta c_i \Delta c_j \text{Tr} \chi_{ij}. \quad (43)$$

If the estimation is unbiased, the average of Δc_i is zero, therefore we have the average fidelity

$$\mathbb{E}[F(\check{\rho}, \rho)] = 1 + \sum_{ij} \mathbb{E}[\Delta c_i \Delta c_j] \text{Tr} \chi_{ij}, \quad (44)$$

noting that $\mathbb{E}[\Delta c_i \Delta c_j]$ is exactly the (i, j) -th element of the covariance matrix Cov. This equation leads to Eq. (20) in the main text.

We would like to show how one can solve Eq. (42). We first expand the density matrix ρ as $\rho = \sum_k \lambda_k |k\rangle\langle k|$ (and $\rho = \sum_l \lambda_l |l\rangle\langle l|$), and take the (l, k) -th element of the first equation of (42). This gives

$$\vartheta_i^{(lk)} = \frac{\sqrt{\lambda_k \lambda_l} \hat{E}_i^{(lk)}}{\lambda_k + \lambda_l}, \quad \forall \lambda_k + \lambda_l \neq 0, \quad (45)$$

where $\vartheta_i^{(lk)} = \langle l | \vartheta_i | k \rangle$ and $\hat{E}_i^{(lk)} = \langle l | \hat{E}_i | k \rangle$, and

$$\vartheta_i = \sum_{\lambda_k + \lambda_l \neq 0} \frac{\sqrt{\lambda_k \lambda_l} \hat{E}_i^{(lk)}}{\lambda_k + \lambda_l} |l\rangle\langle k|. \quad (46)$$

Similarly, by substituting Eq. (46) into the second equation of (42), we have

$$\chi_{ij} = \sum_{\substack{\lambda_l + \lambda_m \neq 0 \\ \lambda_k + \lambda_l \neq 0 \\ \lambda_k + \lambda_m \neq 0}} \frac{\lambda_k \sqrt{\lambda_l \lambda_m} (\hat{E}_i^{(lk)} \hat{E}_j^{(km)} + \hat{E}_j^{(lk)} \hat{E}_i^{(km)})}{(\lambda_l + \lambda_m)(\lambda_k + \lambda_l)(\lambda_k + \lambda_m)} |l\rangle\langle m|,$$

which gives the trace,

$$\text{Tr} \chi_{ij} = \sum_{\lambda_k + \lambda_l \neq 0} \frac{\lambda_k \text{Re}(\hat{E}_i^{(lk)} \hat{E}_j^{(kl)})}{(\lambda_k + \lambda_l)^2}, \quad (47)$$

as shown in Eq. (20) in the main text.

Analytic calculation for the Fisher information matrix

In this section, we present a detailed calculation of the Fisher information matrix Eq. (29) for estimating a two-qubit unknown state in the limit of weak measurement. As defined in the main text, the weak measurement operator for a short time dt can be written as

$$\begin{aligned}M(r_k) &= \left(\frac{dt}{2\pi\tau} \right)^{\frac{1}{4}} \exp \left[-\frac{(r_k - \hat{Z})^2 dt}{4\tau} \right], \\ &= e^{-\frac{dt}{4\tau}} \sqrt{f(r_k)} \exp \left(\frac{r_k dt}{2\tau} \hat{Z} \right),\end{aligned}\quad (48)$$

where $\hat{\mathbf{Z}} = \hat{\mathbf{Z}} \otimes \hat{\mathbf{I}}$ and

$$f(r_k) = \sqrt{\frac{dt}{2\pi\tau}} \exp\left(-\frac{r_k^2}{2\tau} dt\right), \quad (49)$$

is a Gaussian distribution of r_k with mean zero and variance τ/dt . The trajectory Kraus operator is

$$\mathcal{M}_R = M(r_n)U(dt) \dots M(r_1)U(dt), \quad (50)$$

where we denote the total time of the measurement as $T = n dt$.

When dt is very small, we can expand each $M(r_k)$ to the first order of dt ,

$$M(r_k) \approx e^{-\frac{dt}{4\tau}} \sqrt{f(r_k)} \left(\hat{\mathbf{I}} + \frac{r_k dt}{2\tau} \hat{\mathbf{Z}} \right), \quad (51)$$

where $\hat{\mathbf{I}} = I \otimes I$. Note that we only need the first order of dt in Eq. (51), even though we want to keep higher order terms in $1/\tau$ in $\mathcal{M}_R$ later. This is because higher order terms in dt in (51) will have measure zero when we perform integrations over time (see the following equations). The time integration needs a limit taking $dt \rightarrow 0$ and $n \rightarrow \infty$ for any finite T . We substitute the measurement operators and the unitary operators in Eq. (50), keeping only to first order in dt . We then replace any time summations $\sum_k dt$ with the integral $\int dt$ to obtain the operator $\mathcal{M}_R$ in a time-ordering integral form:

$$\begin{aligned} \mathcal{M}_R = & e^{-\frac{T}{4\tau}} \sqrt{\prod_k f(r_{t_k})} \left[U(T) + \sum_{k=1}^{+\infty} \frac{1}{(2\tau)^k k!} \right. \\ & \mathcal{T} \int_0^T \dots \int_0^T r_{t_1} \dots r_{t_k} U(t - t_k) \hat{\mathbf{Z}} U(t_k - t_{k-1}) \\ & \left. \hat{\mathbf{Z}} \dots U(t_2 - t_1) \hat{\mathbf{Z}} U(t_1) dt_1 \dots dt_k \right], \end{aligned} \quad (52)$$

where $\mathcal{T}$ is the time-ordering operator and r_{t_k} is a measurement result at time t_k . If we denote a time-dependent operator

$$\hat{\mathbf{Z}}(t) = U^\dagger(t) \hat{\mathbf{Z}} U(t), \quad (53)$$

then $\mathcal{M}_R$ can be simplified to

$$\begin{aligned} \mathcal{M}_R = & e^{-\frac{T}{4\tau}} \sqrt{\prod_k f(r_{t_k})} U(T) \left[\hat{\mathbf{I}} + \sum_{k=1}^{+\infty} \frac{1}{(2\tau)^k k!} \right. \\ & \left. \times \mathcal{T} \int_0^T \dots \int_0^T r_{t_1} \dots r_{t_k} \hat{\mathbf{Z}}(t_k) \dots \hat{\mathbf{Z}}(t_1) dt_1 \dots dt_k \right], \end{aligned} \quad (54)$$

and we can write a POVM element corresponding to $\mathcal{M}_R$ as

$$\mathcal{E}_R = \mathcal{M}_R^\dagger \mathcal{M}_R. \quad (55)$$

In the weak measurement limit, we will keep the terms up to the order $1/\tau$ throughout the rest of this appendix. For the POVM $\mathcal{E}_R$, we expand to first order in $1/\tau$ and get,

$$\mathcal{E}_R = e^{-\frac{T}{2\tau}} \prod_k f(r_{t_k}) \left[\hat{\mathbf{I}} + \frac{1}{\tau} \int_0^T r_t \hat{\mathbf{Z}}(t) dt \right], \quad (56)$$

where r_t is the measurement result at time t . Given an initial state of the two qubits

$$\rho_0 = \frac{1}{4} \sum_{i,j=0,x,y,z} c_{ij} \hat{\sigma}_i \otimes \hat{\sigma}_j, \quad (57)$$

where $\hat{\sigma}_0 = \hat{I}$ and $c_{00} = 1$, the probability for the measurement record R can be computed from the POVM element, $P(R|\rho_0) = \text{Tr}(\mathcal{E}_R \rho_0)$,

$$= e^{-\frac{T}{2\tau}} \prod_k f(r_{t_k}) \left[1 + \frac{1}{\tau} \int_0^T r_t \langle \hat{Z}(t) \rangle_0 dt \right], \quad (58)$$

where $\langle \cdot \rangle_0$ denotes the expectation value with the initial state ρ_0 . We calculate the derivative of $P(R|\rho_0)$ with respect to the parameter c_{ij} giving,

$$\partial_{ij} P(R|\rho_0) = \frac{e^{-\frac{T}{2\tau}}}{4\tau} \prod_k f(r_{t_k}) \int_0^T r_t \alpha_{ij}(t) dt, \quad (59)$$

where we have defined,

$$\alpha_{ij}(t) = \text{Tr}(\hat{Z}(t) \hat{\sigma}_i \otimes \hat{\sigma}_j). \quad (60)$$

We then use the above equations to compute the $(ij, i'j')$ -th element of Fisher information matrix for estimating a two-qubit unknown state,

$$F_{ij,i'j'} = \int \dots \int \frac{\partial_{ij} P(R|\rho_0) \partial_{i'j'} P(R|\rho_0)}{P(R|\rho_0)} dr_{t_1} \dots dr_{t_n}. \quad (61)$$

Keeping to first order in $1/\tau$, we have

$$\frac{1}{P(R|\rho_0)} = \frac{e^{\frac{T}{2\tau}}}{\prod_k f(r_{t_k})} \left[1 - \frac{1}{\tau} \int_0^T r_t \langle \hat{Z}(t) \rangle_0 dt \right], \quad (62)$$

and

$$\begin{aligned} & \frac{\partial_{ij} P(R|\rho_0) \partial_{i'j'} P(R|\rho_0)}{P(R|\rho_0)} \\ &= \frac{e^{-\frac{T}{2\tau}}}{16\tau^2} \prod_k f(r_{t_k}) \int_0^T \int_0^T r_{t_1} r_{t_2} \alpha_{ij}(t_1) \alpha_{i'j'}(t_2) dt_1 dt_2. \end{aligned} \quad (63)$$

Substituting the above two equations to the Fisher information matrix element Eq. (61), we then perform the integrals of $r_1, r_2, \dots, r_n$ with $f(r_k)$ being a Gaussian distribution with zero mean and variance τ/dt . Using the properties of integrals with Gaussian functions, we have

$$\int \dots \int \prod_k f(r_{t_k}) r_{t_1} r_{t_2} dr_{t_1} \dots dr_{t_n} = \tau \delta(t_1 - t_2). \quad (64)$$

By applying Eqs. (64)–(63), we obtain

$$\int \dots \int \prod_k f(r_{t_k}) r_{t_1} r_{t_2} \alpha_{ij}(t_1) \alpha_{i'j'}(t_2) dr_{t_1} \dots dr_{t_n} = \tau \alpha_{ij}(t_1) \alpha_{i'j'}(t_2) \delta(t_1 - t_2), \quad (65)$$

and, the $(ij, i'j')$ -th element of Fisher information matrix

$$F_{ij,i'j'} = \frac{e^{-\frac{T}{2\tau}}}{16\tau} \int_0^T \alpha_{ij}(t) \alpha_{i'j'}(t) dt, \quad (66)$$

where we have replaced t_1 or t_2 by t .

5 Methods for numerical simulation

Numerical simulation is used to investigate the state estimation quality for different measurement and control settings. There are two main parts in the numerical calculation. The first part is to generate N measurement records

using an initial state ρ_0 chosen at random, regardless of its purity or levels of entanglement. Starting with the given initial state, each record consists of measurement results at all time steps $r_1, r_2, \dots, r_n$ are randomly generated with probability distributions $P(r_k|\rho_k)$ for $k = 1, 2, \dots, n$, where the quantum states ρ_k are calculated for each time step with an update equation,

$$\rho(t_{k+1}) = \frac{U(dt)M(r_k)\rho(t_k)M(r_k)^\dagger U(dt)^\dagger}{\text{Tr}[U(dt)M(r_k)\rho(t_k)M(r_k)^\dagger U(dt)^\dagger]}, \quad (67)$$

where $M(r_k)$ is defined in Eq. (3) in the main text and $U(dt) = \exp(-i\hat{H}dt)$. We use the time step $dt = 0.01$ and the total measurement time is fixed at $T = 2$ for all results presented in the main text. The second part of the simulation is to use the generated N records in an estimation of the initial state ρ_0 assuming that it is unknown. Using the N records denoted as $R_1, R_2, \dots, R_N$, we construct evolution operators $\mathcal{M}_{R_j}$ for $j = 1, \dots, N$ defined in Eq. (2) and compute a function $P(\{R\}|\rho')$ Eq. (24) for an arbitrary state ρ' . This is the likelihood function used in estimating the unknown state.

For the single-qubit state tomography, we focus more on the Bayesian estimator because of its robustness in comparison to the maximum likelihood estimation. The probability function $P(\{R\}|\rho')$ is then applied to 10^4 different trial states ρ' which are equally distributed on the Bloch sphere. We use a MATLAB code named “RandomDensityMatrix” from QETLAB to generate single-qubit random density matrices distributed uniformly according to Hilbert–Schmidt measure [http://www.qetlab.com]. The reason we chose 10^4 trial states is so that an average distance between any two trial state is given as $\Delta\rho = \left(\frac{4}{3}\frac{\pi}{10000}\right)^{1/3} = 0.074 \leq 0.1$. Once the probability function is calculated for every trial state, we find a Bayesian probability from,

$$P(\rho'|\{R\}) = P(\{R\}|\rho')P(\rho'), \quad (68)$$

where, in the case of the uniform distribution, we simply replace $P(\rho')$ by 1. This Bayesian probability is used to calculate the Bayesian estimator Eq. (11) and the maximum likelihood estimator Eq. (8). The error of the estimation is computed from the Bayesian covariance matrix Eq. (22) using the true state ρ_0 . For the results in Fig. 4b, c, we generated 100 sets of the estimators, from different sets of N records, so that the results are more reliable.

For the remote-qubit state tomography, we use the similar method as in the single-qubit case. The trial states in this case are 4×4 density matrices, which are tensor products of the ancilla initial state and the single-qubit trial state ρ' (same 10^4 set as in the single-qubit case). For the results shown in Fig. 6, we compute the root mean square errors from 100 sets of the estimators, considering each element of the density matrix as an estimated parameter. We could use the errors from the Bayesian covariance matrix similar to the single-qubit case, but using this type of errors makes the transition to the two-qubit tomography discussion easier to understand. The results for single-qubit and remote-qubit cases are generated using Python and MATLAB codes.

For the two-qubit state tomography, we use a different technique from the previous two examples because the state space of the two qubits is much larger. As stated in the main text, to use the Bayesian estimator for the two-qubit tomography, one needs much more than 10^8 trial states. This number comes from a number of product states of the single-qubit trial states $10^4 \times 10^4$. Instead of using the Bayesian estimator, we then compute the maximum likelihood estimator from maximizing the function $P(\{R\}|\rho')$ over all valid two-qubit states ρ' , as discussed in the main text. We use the built-in Differential Evolution algorithm in Mathematica to search for the estimator, using the constraints in Eq. (37) and 50 searches initialized with random numbers between $-\sqrt{3}$ and $\sqrt{3}$ [for each of the 15 elements, see the first constraint in Eq. (37)]. We compute 50 MLE estimators, each with different sets of $N = 4000$ measurement records, and calculate the RMSEs from the 50 estimators to get the results in Fig. 7.

6 Lists of random initial states of two-qubit tomography simulation

For the single-qubit case, the average fidelity for different initial states in Fig. 4c is computed from these nine randomly chosen states: $\{x_0, y_0, z_0\} = \{-0.4, -0.6, 0.3\}, \{-0.4, 0.6, -0.3\}, \{-0.4, 0.6, 0.3\}, \{0.4, 0.6, -0.3\}, \{-0.7, -0.5, -0.3\}, \{-0.7, -0.5, 0.3\}, \{0.7, -0.5, 0.3\}, \{-0.5, 0.3, 0.8\}, \{-0.5, -0.3, -0.8\}, \{0.5, 0.3, 0.8\}$.

For the remote qubit, the average fidelity for different initial states mentioned in the caption of Fig. 6d is computed from these ten randomly chosen states: $\{x_u^0, y_u^0, z_u^0\} = \{1/\sqrt{2}, 1/\sqrt{2}, 0\}, \{-0.6, -0.4, 0.3\}, \{-0.7, -0.5, -0.3\}, \{0.0, -0.5, 0.3\}, \{0.3, -0.3, 0.3\}, \{0.3, -0.7, 0.5\}, \{0.5, 0.3, 0.8\}, \{0.7, -0.5, 0.1\}, \{0.7, 0.0, 0.0\}, \{0.7, 0.0, 0.3\}$. The fidelities for these initial states (each with 100 repetitions) are 0.98853, 0.998968, 0.998952, 0.999116, 0.999447, 0.99888, 0.993246, 0.999316, 0.999317, 0.999122, giving an average 0.998 ± 0.004 .

In the two-qubit case, we presented in the main text the average fidelities in Fig. 7. The initial states were chosen randomly including mixed states and non-separable states. We define a function $W(p, \rho_1, \rho_2)$,

$$W(p, \rho_1, \rho_2) = (1 - p)\rho_1 \otimes \rho_2 + p|\Phi^+\rangle\langle\Phi^+|, \quad (69)$$

where $|\Phi^+\rangle = 1/\sqrt{2}(|00\rangle + |11\rangle)$ is a Bell state written in a computational basis 00, 01, 10, 11 of the two-qubit state. The nine random initial states are,

$$\begin{aligned} & - W(0.5, \frac{1}{2}(\hat{I} + 0.7\hat{X} - 0.2\hat{Y} + 0.3\hat{Z}), \frac{1}{2}(\hat{I} + 0.6\hat{X} - 0.1\hat{Y} + 0.4\hat{Z})) \\ & - |\Phi^+\rangle\langle\Phi^+| \\ & - W(0.2, \frac{1}{2}(\hat{I} - 0.4\hat{X} - 0.75\hat{Y} + 0.5\hat{Z}), \frac{1}{2}(\hat{I} + 0.6\hat{X} - 0.5\hat{Y} + 0.6\hat{Z})) \\ & - W(0.5, \frac{1}{2}(\hat{I} + 0.2\hat{X} - 0.75\hat{Y} + 0.5\hat{Z}), \frac{1}{2}(\hat{I} + 0.6\hat{X} - 0.5\hat{Y} + 0.4\hat{Z})) \\ & - W(0.8, \frac{1}{2}(\hat{I} + 0.2\hat{X} - 0.75\hat{Y} + 0.5\hat{Z}), \frac{1}{2}(\hat{I} + 0.6\hat{X} - 0.5\hat{Y} + 0.4\hat{Z})) \\ & - \frac{1}{2}(\hat{I} + 0.7\hat{X} - 0.2\hat{Y} + 0.5\hat{Z}) \otimes \frac{1}{2}(\hat{I} + 0.6\hat{X} - 0.5\hat{Y} + 0.4\hat{Z}) \\ & - W(0.5, \frac{1}{2}(\hat{I} + 0.7\hat{X} - 0.2\hat{Y} + 0.5\hat{Z}), \frac{1}{2}(\hat{I} + 0.6\hat{X} - 0.5\hat{Y} + 0.4\hat{Z})) \\ & - W(0.8, \frac{1}{2}(\hat{I} + 0.7\hat{X} - 0.2\hat{Y} + 0.5\hat{Z}), \frac{1}{2}(\hat{I} + 0.6\hat{X} - 0.5\hat{Y} + 0.4\hat{Z})) \\ & - \frac{1}{2}(\hat{I} + 0.7\hat{X} - 0.2\hat{Y} + 0.3\hat{Z}) \otimes \frac{1}{2}(\hat{I} + 0.6\hat{X} - 0.1\hat{Y} + 0.4\hat{Z}) \end{aligned}$$

References

1. Barchielli, A., Gregoratti, M.: Quantum Trajectories and Measurements in Continuous Time. Springer, Berlin (2009)
2. Bassa, H., Goyal, S.K., Choudhary, S.K., Uys, H., Diósi, L., Konrad, T.: Process tomography via sequential measurements on a single quantum system. *Phys. Rev. A* **92**, 032102 (2015)
3. Blume-Kohout, R.: Optimal, reliable estimation of quantum states. *New J. Phys.* **12**, 043034 (2010)
4. de Burgh, M.D., Langford, N.K., Doherty, A.C., Gilchrist, A.: Choice of measurement sets in qubit tomography. *Phys. Rev. A* **78**, 052122 (2008)
5. Campagne-Ibarcq, P., Six, P., Bretheau, L., Sarlette, A., Mirrahimi, M., Rouchon, P., Huard, B.: Observing quantum state diffusion by heterodyne detection of fluorescence. *Phys. Rev. X* **6**, 011002 (2016)
6. Carmichael, H.J.: An Open Systems Approach to Quantum Optics. Springer, Berlin (1993)
7. Chantasri, A., Dressel, J., Jordan, A.N.: Action principle for continuous quantum measurement. *Phys. Rev. A* **88**, 042110 (2013)
8. Chantasri, A., Jordan, A.N.: Stochastic path-integral formalism for continuous quantum measurement. *Phys. Rev. A* **92**, 032125 (2015)
9. Cook, R.L., Riofrío, C.A., Deutsch, I.H.: Single-shot quantum state estimation via a continuous measurement in the strong backaction regime. *Phys. Rev. A* **90**, 032113 (2014)
10. Cortez, L., Chantasri, A., García-Pintos, L.P., Dressel, J., Jordan, A.N.: Rapid estimation of drifting parameters in continuously measured quantum systems. *Phys. Rev. A* **95**, 012314 (2017)
11. Cramér, H.: Mathematical Methods of Statistics. Princeton University Press, Princeton (1946)
12. Diósi, L.: Continuous quantum measurement and itô formalism. *Phys. Lett.* **129**, 419–423 (1988)
13. Doherty, A.C., Habib, S., Jacobs, K., Mabuchi, H., Tan, S.M.: Quantum feedback control and classical control theory. *Phys. Rev. A* **62**, 012105 (2000)
14. Gamel, O.: Entangled bloch spheres: bloch matrix and two-qubit state space. *Phys. Rev. A* **93**, 062320 (2016)

15. Gross, D., Liu, Y.K., Flammia, S.T., Becker, S., Eisert, J.: Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **105**, 150401 (2010)
16. Gross, J.A., Dagniam, N., Ferrie, C., Caves, C.M.: Novelty, efficacy, and significance of weak measurements for quantum tomography. *Phys. Rev. A* **92**, 062133 (2015)
17. Hacohe-Gourgy, S., Martin, L.S., Flurin, E., Ramasesh, V.V., Whaley, K.B., Siddiqi, I.: Quantum dynamics of simultaneously measured non-commuting observables. *Nature* **538**, 491–494 (2016)
18. Hradil, Z., Řeháček, J., Fiurášek, J., Ježek, M.: Maximum-likelihood methods in quantum mechanics. In: *Quantum State Estimation*, pp. 59–112. Springer, Berlin (2004)
19. Jacobs, K.: *Quantum Measurement Theory and its Applications*. Cambridge University Press, Cambridge (2014)
20. James, D.F.V., Kwiat, P.G., Munro, W.J., White, A.G.: Measurement of qubits. *Phys. Rev. A* **64**, 052312 (2001)
21. Kiilerich, A.H., Mølmer, K.: Bayesian parameter estimation by continuous homodyne detection. *Phys. Rev. A* **94**, 032103 (2016)
22. Korotkov, A.N.: Continuous quantum measurement of a double dot. *Phys. Rev. B* **60**, 5737 (1999)
23. Korotkov, A.N.: Selective quantum evolution of a qubit state due to continuous measurement. *Phys. Rev. B* **63**, 115403 (2001)
24. Lindblad, G.: On the generators of quantum dynamical semigroups. *Commun. Math. Phys.* **48**, 119 (1976)
25. Murch, K.W., Weber, S.J., Macklin, C., Siddiqi, I.: Observing single quantum trajectories of a superconducting quantum bit. *Nature* **502**, 211 (2013)
26. Naghiloo, M., Foroozani, N., Tan, D., Jadbabaie, A., Murch, K.W.: Mapping quantum state dynamics in spontaneous emission. *Nat. Commun.* **7**, 11527 (2016)
27. Newton, R.G., Young, B.I.: Measurability of the spin density matrix. *Ann. Phys.* **49**, 393–402 (1968)
28. Nielsen, M.A., Chuang, I.L.: *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge (2000)
29. Ralph, J.F., Jacobs, K., Hill, C.D.: Frequency tracking and parameter estimation for robust quantum state estimation. *Phys. Rev. A* **84**(5), 052119 (2011)
30. Ralph, J.F., Maskell, S., Jacobs, K.: Multiparameter estimation along quantum trajectories with sequential monte carlo methods. *Phys. Rev. A* **96**, 052306 (2017)
31. Renes, J.M., Blume-Kohout, R., Scott, A.J., Caves, C.M.: Symmetric informationally complete quantum measurements. *J. Math. Phys.* **45**(6), 2171–2180 (2004)
32. Riofrío, C.A., Jessen, P.S., Deutsch, I.H.: Quantum tomography of the full hyperfine manifold of atomic spins via continuous measurement on an ensemble. *J. Phys. B Atom. Mol. Opt. Phys.* **44**(15), 154007 (2011)
33. Roch, N., Schwartz, M.E., Motzoi, F., Macklin, C., Vijay, R., Eddins, A.W., Korotkov, A.N., Whaley, K.B., Sarovar, M., Siddiqi, I.: Observation of measurement-induced entanglement and quantum trajectories of remote superconducting qubits. *Phys. Rev. Lett.* **112**, 170501 (2014)
34. Shojaei, E., Jackson, C.S., Riofrío, C.A., Kalev, A., Deutsch, I.H.: Optimal pure-state qubit tomography via sequential weak measurements. *Phys. Rev. Lett.* **121**, 130404 (2018)
35. Silberfarb, A., Jessen, P.S., Deutsch, I.H.: Quantum state reconstruction via continuous measurement. *Phys. Rev. Lett.* **95**, 030402 (2005)
36. Six, P., Campagne-Ibarcq, P., Dotsenko, I., Sarlette, A., Huard, B., Rouchon, P.: Quantum state tomography with noninstantaneous measurements, imperfections, and decoherence. *Phys. Rev. A* **93**, 012109 (2016)
37. Smith, G.A., Silberfarb, A., Deutsch, I.H., Jessen, P.S.: Efficient quantum-state estimation by continuous weak measurement and dynamical control. *Phys. Rev. Lett.* **97**, 180403 (2006)
38. Weber, S.J., Murch, K.W., Kimchi-Schwartz, M.E., Roch, N., Siddiqi, I.: Quantum trajectories of superconducting qubits. *Comptes Rendus Phys.* **17**(7), 766–777 (2016). *Quantum microwaves / Micro-ondes quantiques*
39. Wheatley, T.A., Berry, D.W., Yonezawa, H., Nakane, D., Arao, H., Pope, D.T., Ralph, T.C., Wiseman, H.M., Furusawa, A., Huntington, E.H.: Adaptive optical phase estimation using time-symmetric quantum smoothing. *Phys. Rev. Lett.* **104**, 093601 (2010)
40. Wiseman, H.M.: Quantum theory of continuous feedback. *Phys. Rev. A* **49**, 2133–2150 (1994)
41. Wiseman, H.M., Milburn, G.J.: *Quantum measurement and control*. Cambridge University Press, Cambridge (2009)
42. Yonezawa, H., Nakane, D., Wheatley, T.A., Iwasawa, K., Takeda, S., Arao, H., Ohki, K., Tsumura, K., Berry, D.W., Ralph, T.C., Wiseman, H.M., Huntington, E.H., Furusawa, A.: Quantum-enhanced optical-phase tracking. *Science* **337**(6101), 1514–1517 (2012)