Prediction and retrodiction for a continuously monitored superconducting qubit

D. Tan, S. J. Weber, I. Siddiqi, K. Mølmer, and K. W. Murch

1 Department of Physics, Washington University, St. Louis, Missouri 63130
2 Quantum Nanoelectronics Laboratory, Department of Physics, University of California, Berkeley CA 94720
3 Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark

(Dated: November 11, 2014)

The quantum state of a superconducting transmon qubit inside a three-dimensional cavity is monitored by transmission of a microwave field through the cavity. The information inferred from the measurement record is incorporated in a density matrix $\rho_t$, which is conditioned on probe results until $t$, and in an auxiliary matrix $E_t$, which is conditioned on probe results obtained after $t$. Here, we obtain these matrices from experimental data and we illustrate their application to predict and retrodict the outcome of weak and strong qubit measurements.

In quantum mechanics, predictions about the outcome of experiments are given by Born’s rule which for a state vector $|\psi_i\rangle$ provides the probability $P(a) = |\langle a|\psi_i\rangle|^2$ that a measurement of an observable $A$ with eigenstates $|a\rangle$ yields one of the eigenvalues $a$. As a consequence of the measurement, the quantum state is projected into the state $|a\rangle$. Yet, after this measurement, further probing of the system is possible, and the probability that the quantum system yields outcome $a$ is subsequently detected in a final state $|\psi_f\rangle$ factors into the product $|\langle a|\psi_i\rangle|^2|\langle a|\psi_f\rangle|^2$. Considering initial and final states raises the issue of post-selection in quantum measurements: What is the probability that the result of the measurement of $A$ was $a$, if we consider only the selected measurement events where the initial state was $|\psi_i\rangle$ and the final state was $|\psi_f\rangle$? The answer is known as the Aharonov-Bergmann-Lebowitz rule [1].

$$P_{ABL}(a) = \frac{P(f, a|\psi_i)}{\sum_{a'} P(f, a'|\psi_i)} = \frac{|\langle a|\psi_i\rangle|\langle a|\psi_f\rangle|^2}{\sum_{a'} |\langle a'|\psi_i\rangle|^2} \tag{1}$$

and it differs from Born’s rule, which takes into account only knowledge about the state prior to the measurement.

While it is natural that full measurement records reveal more information about the state of a physical system at a given time $t$ than data obtained only until that time, the interpretation of the time symmetric influences from the future and from the past measurement events on $P_{ABL}$ has stimulated some debate, see for example [1][3]. Meanwhile, probabilistic state assignments and correlations observed in atomic, optical and solid state experiments have been conveniently understood in relation to post-selection [7][11], and precision probing theories [12][17] have incorporated full measurement records.

In this letter, we consider a superconducting qubit that is subject to continuous monitoring and driven unitary evolution. We make use of the full measurement record and examine how measurements before time $t$ can be used to make predictions, while measurements after time $t$ can be used to make retrodictions about measurements at time $t$. We then consider a recent generalization [2] of Eq.(1) to the case of continuously monitored and evolving mixed states. Our experiments verify the predictions of both projective and weak (weak value) measurements conditioned on full measurement records. These predictions are more confident and nontrivially different from predictions based only on the measurement record up to time $t$.

To analyze non-pure states and partial measurements, we represent our system by a density matrix $\rho$, and measurements by the theory of positive operator-valued measures (POVM) which yields the probability $P(m) = \text{Tr}(\Omega_m \rho \Omega_m^\dagger)$ for outcome $m$, and the associated back action on the quantum state, $\rho \rightarrow \Omega_m \rho \Omega_m^\dagger / P(m)$, where the operators $\Omega_m$ obey $\sum_m \Omega_m \Omega_m^\dagger = I$. When $\Omega_a = |a\rangle\langle a|$ is a projection operator and $\rho = |\psi\rangle\langle \psi|$, the theory of POVMs is in agreement with Born’s rule.

For systems subject to unitary and dissipative time evolution along with continuous monitoring before and after a measurement described by operator $\Omega_m$, one can show [2] that

$$P_p(m) = \frac{\text{Tr}(\Omega_m \rho \Omega_m^\dagger E_t) \text{Tr}(\Omega_m \rho \Omega_m E_t)}{\sum_m \text{Tr}(\Omega_m \rho \Omega_m E_t)^2}, \tag{2}$$

where $\rho_t$ is the system density matrix at time $t$, conditioned on previous measurement outcomes, and propagated forward in time until time $t$, while $E_t$ is a matrix which is propagated backwards in time in a similar manner and accounts for the time evolution and measurements obtained after time $t$. The subscript $p$ denotes “past”, and in [2] it was proposed that, if $t$ is in the past, the pair of matrices $(\rho_t, E_t)$, rather than only $\rho_t$, is the appropriate object to associate with the state of a quantum system at time $t$. We observe that for the case of pure states and projective measurements, $P_p(m)$ in [2] acquires the form of Eq.(1) with $\rho_t = |\psi_i\rangle\langle \psi_i|$ and $E_t = |\psi_f\rangle\langle \psi_f|$.

Here, we make use of the full measurement record to compute the matrices $\rho_t$ and $E_t$, and analyze how they, through application of Eq.(2), yield more confident predictions for measurements on the system. For imperfect
Figure 1: Time evolution in a monitored system. (a) Simplified experimental setup consisting of a transmon circuit coupled to a waveguide cavity. (b) We prepare the qubit in an initial state $(\text{Tr}(\rho_i, \sigma_z) \approx +1)$ and propagate $\rho$ forward in time, which makes accurate predictions about a final projective measurement (in the $\sigma_z$ basis) labeled $M$. The dashed line is the prediction based on a single quantum trajectory, and the solid line is the result from projective measurements on an ensemble of experiments that have similar values of $\rho_i$.

measurement efficiency, non-pure states, and measurements that do not commute with the system evolution, the predictions of Eq. (2) vary dramatically from those based on $\rho$ alone [3] [20].

Our experiment, illustrated in figure 1a, is composed of a superconducting transmon circuit dispersive coupled to a waveguide cavity [22] [23]. The two lowest energy levels of the transmon form a qubit with transition frequency $\omega_q/2\pi = 4.0033$ GHz. The dispersive coupling between the transmon qubit and the cavity is given by an interaction Hamiltonian, $H_{\text{int}} = -\hbar \chi a^\dagger a \sigma_z$, where $\hbar$ is the reduced Plank’s constant, $a^\dagger (a)$ is the creation (annihilation) operator for the cavity mode at frequency $\omega_c/2\pi = 6.9914$ GHz, $\chi/2\pi = -0.425$ MHz is the dispersive coupling rate, and $\sigma_z$ is the qubit Pauli operator that acts on the qubit in the energy basis. A microwave tone that probes the cavity with an average intracavity photon number $\bar{n} = \langle a^\dagger a \rangle$ thus acquires a qubit-state-dependent phase shift. Since $2|\chi| \ll \kappa$, where $\kappa/2\pi = 9.88$ MHz is the cavity linewidth, qubit state information is encoded in one quadrature of the microwave signal. We amplify this quadrature of the signal with a near-quantum-limited Josephson parametric amplifier [24]. After further amplification, the measurement signal is demodulated and digitized. This setup allows variable strength measurements of the qubit state characterized by a measurement timescale $\tau$; by binning the measurement signal in time steps $\delta t \ll \tau$ we execute weak measurements of the qubit state [20] [23] while by integrating the measurement signal for a time $T \gg \tau$ we effectively accumulate weak measurements in a projective measurement [25] of the qubit in the $\sigma_z$ basis.

Our experimental sequences begin with a projective measurement of the qubit in the $\sigma_z$ basis followed by a variable rotation of the qubit state to prepare the qubit in an arbitrarily specified initial (nearly) pure state. Following this preparation, the qubit is subject to continuous rotations given by $H_R = \hbar \Omega_R \sigma_y/2$, where $\Omega_R/2\pi = 0.7$ MHz is the Rabi frequency, and continuous probing given by the measurement operator $\sqrt{\chi} \sigma_z$, where $k = 4\chi^2\bar{n}/\kappa = 1/4\delta \Omega$ parametrizes the measurement strength $(k/2\pi = 95$ kHz) and $\eta = 0.35$ is the quantum measurement efficiency [27]. During probing, we digitize the measurement signal $V_t$ in time steps $\delta t = 20$ ns.

The density matrix associated with a given measurement signal $V_t$ is obtained by solving the stochastic master equation [1]:

$$
\frac{d\rho}{dt} = -\frac{i}{\hbar}[H_R, \rho] + k(\sigma_z \rho \sigma_z - \rho) + 2\eta k(\sigma_z \rho + \rho \sigma_z - 2\text{Tr}(\sigma_z \rho)\rho)V_t. \tag{3}
$$

Here, the first two terms are the standard master equation in Lindblad form, and the last stochastic term updates the state based on the measurement result and leads to quantum trajectory solutions that are different for every repetition of the experiment.

Let us first recall how the density matrix makes predictions about the outcome of measurements. In figure 1b, we consider the probabilities $P(\pm z)$ for the outcome of the projective measurement operators $\Omega_{\pm z} = (\sigma_z \pm 1)/2$. We prepare the initial state, $\text{Tr}(\rho_i, \sigma_z) \approx +1$, by heralding the ground state and applying a $\pi/2$ rotation about the y axis. We then propagate $\rho_i$ forward from this initial state, and at each point in time we display the calculated $P(\pm z) = \text{Tr}[\Omega_{\pm z} \rho_i \Omega_{\pm z}^\dagger]$. By performing projective measurements of $\Omega_{\pm z}$ at time $t$ on an ensemble of experiments that have similar values of $\rho_i$ (within ±0.02) we obtain the corresponding experimental result $P(\pm z)$. We perform this analysis at different times and we observe close agreement between the single quantum trajectory prediction $P(\pm z)$ and the observed $P(\pm z)$. Note that the same procedure was used to tomographically reconstruct and verify the quantum trajectory associated with the mean value $\langle \sigma_z \rangle = 2P(\pm z) - 1$ in [3] [20].

We now turn to the application of measurement data to retrodict the outcome of an already performed measurement. Eq. (2) applies for any set of POVM measurement operators $\Pi_m$ at time $t$, and accumulates the information retrieved from the later probing in the matrix $E_t$ that is propagated backwards in time according to [2].

$$
\frac{dE}{dt} = \frac{1}{\hbar}[H_R, E] + k(\sigma_z E \sigma_z - E) + 2\eta k(\sigma_z E + E \sigma_z - \text{Tr}(\sigma_z E)\rho)V_t. \tag{4}
$$

We assume that no measurements take place beyond the time $T$, leading to the final condition $E_T = \tilde{I}$ [2] (Note that $\text{Tr}(E) = 2$). If no measurements take place at
We examine different initial states,

\[ \T(\rho E_\uparrow) \approx 1, \quad \T(\rho E_o) \approx +1, \quad \text{blue}, \]

and compare the retrodictions, \( P_p(z) \), based on 5 μs of probing, to the outcomes of measurements \( M \) that yielded similar values of \( E_0 \). In the lower panel, we display histograms of \( P_p(z) \) for different propagation times. As more of the record is included, the retrodictions become more confident, taking values that are more often near 0 or 1.

all before \( T \), for example because \( \eta = 0 \), Eq. (4) yields a solution for \( E(t) \) that remains proportional to the identity operator for all times, and Eq. (2) leads to the conventional expression that depends only on \( \rho_t \).

In figure 2 we test the retrodictions made by \( E \) and Eq. (2). We examine different initial states, \( \T(\rho E_\uparrow) \approx +1, \quad \T(\rho E_o) \approx +1, \quad \text{blue} \), which are prepared by heralding the ground state and applying \( \pi/2 \) rotations about the \( y \) and \(-x\) axes respectively. We propagate \( E \) backwards from \( E_T \) to \( E_0 \) to make a retrodiction about a projective measurement \( M \). Note that the initial states \( \rho_t \) make ambiguous predictions about the outcome of \( M \), \( P(z) = 1/2 \), yet information is available after \( M \) and by propagating \( E \) for longer times, the retrodiction for the outcome of \( M \) becomes more confident.

We verify that the retrodictions are correct by averaging the outcomes of many measurements \( M \) that corresponded to similar values of \( E_0 \) to obtain an experimentally derived probability, \( \T^z P(z) \). Figure 2a displays two sample trajectories for the retrodiction \( P_p(z) \) along with \( \T^z P(z) \). As more information is included, the retrodictions converge to fixed values. Figure 2b displays the results of \( 3 \times 10^5 \) experimental tests for the two different initial states \( \rho_t \). For both initial states and for a wide range of measurement outcomes we are able to tomographically verify the retrodictions. We also display histograms of the different values \( P_p(z) \) for different propagation times of \( E \). These show that as information is included in the propagation of \( E \) the retrodictions become more confident.

Having verified the predictions based on \( \rho \), and the retrodictions based on \( E \), we now aim to illustrate the application of \( \rho \) and \( E \) to create a \textit{smoothed} prediction (which uses both past and future information) for the outcome of a POVM measurement. The POVM measurement that we consider is simply a short segment of the measurement signal received between \( t \) and \( t + \Delta t \) and is given by the measurement operators \( [1, 28] \),

\[
\Omega_V = (2\pi a^2)^{-1/4} e^{-(V - \sigma_z)^2/4a^2} \]  

where, \( 1/4a^2 = k\eta\Delta t \). The operators \( \Omega_V \) satisfy \( \int \Omega_V^\dagger \Omega_V \, dV = \mathbb{I} \) as expected for POVMs, and if we assume that \( \rho_t \) can be treated as a constant during \( \Delta t \), the probability of the measurement yielding a value \( V \) is \( P(V) = \T(\Omega_V \rho_t \Omega_V^\dagger) \), which is the sum of two Gaussian distributions with variance \( a^2 \) centered at \(+1\) and \(-1\) and weighted by the populations \( \rho_{00} \) and \( \rho_{11} \) of the two qubit states. The \( \sigma_z \) term in \( \Omega_V \) causes the back action on the qubit degree of freedom, \( \rho \rightarrow \Omega_V \rho \Omega_V^\dagger \), due to the readout of the measurement result \( V \). If the effects of damping and the Rabi drive can be ignored during \( \Delta t \), the operators \( [5] \) also describe a stronger measurement, yielding ultimately the limit where the two Gaussian distributions are disjoint, and the readout causes projective back action of the qubit on one of its \( \sigma_z \) eigenstates, with probabilities \( \rho_{00} \) and \( \rho_{11} \).

Since the system is also subject to probing and evolution after \( t \), we now examine what smoothed predictions can be made for the outcome of the measurement \( \Omega_V \) based on both \textit{earlier} and \textit{later} probing. We must hence evaluate the conditioned density matrix \( \rho_t \) and the matrix \( E_t \) and Eq. (2) yields the outcome probability distribution expressed in terms of their matrix elements,

\[
P_p(V) \propto \rho_{00} E_{00} e^{-(V_1 - 1)^2/2a^2} + \rho_{11} E_{11} e^{-(V_1 - 1)^2/2a^2}
\]

\[
+ (\rho_{10} E_{01} + \rho_{01} E_{10}) e^{-(V_2 + 1)^2/2a^2}.
\]

We observe that the information obtained after the measurement of interest plays a formally equally important role as the conditional quantum state represented by \( \rho \).

The predicted mean value is \( \langle V \rangle_p = \int P_p(V) V \, dV \), and can be evaluated,

\[
\langle V \rangle_p = \frac{\rho_{00} E_{00} - \rho_{11} E_{11}}{\rho_{00} E_{00} + \rho_{11} E_{11} + \exp(-\frac{\Delta}{a^2})(\rho_{10} E_{01} + \rho_{01} E_{10})}.
\]  

Here we note that if the measurement is strong, \( a \) is small, and the coherence contribution is cancelled in the denominator, yet if the measurement is weak, a single measurement is dominated by noise and reveals only little information (and causes infinitesimal back action). This is
the situation that leads to so-called weak values. If the measurement signal is proportional to an observable $\hat{A}$, and the system is initialized in $|\psi_i\rangle$ and post-selected in state $|\psi_f\rangle$, the mean signal is given by [29],

$$\langle \hat{A}_w \rangle = \text{Re} \left[ \frac{\langle \psi_f | \hat{A} | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle} \right],$$

which may differ dramatically from the usual expectation value $\langle \psi_i | \hat{A} | \psi_i \rangle$. Our Eq. (7) has, indeed, been derived by Wiseman [30] to clarify how weak values are related to continuous quantum trajectories and correlations in field measurements.

In figure 3 we display results of our experiments that test the predictions of Eq. (6). For many iterations of the experiment we choose a measurement time interval, $\Delta t = 180$ ns that is short enough that the effect of the continuous Rabi drive is nearly negligible in the time interval $(t, t+\Delta t)$. Based on 800 ns of probing before $t$, we calculate $P(V)$, and based on 800 ns of probing before and after the measurement interval, we calculate $P_p(V)$ for the result of the measurement. In Fig. 3 we show that both the conventional and the past quantum state formalism yield agreement between the predicted mean value and the measured values. The measured results are noisy, and we plot the data with the predicted average value along the horizontal axes, and the measured values along the vertical axes.

While $\langle V \rangle = \langle \sigma_z \rangle$, and thus never exceeds 1, a fraction of the experiments lead to prediction and observation of values $|\langle V \rangle_p| > 1$. Such anomalous weak values in connection with Eq. (7) have been typically identified with the intentional post selection of final states with a very small overlap with the initial state. Surprisingly, continuous probing leads to similar effects [3]. In figure 4 we examine the states that lead to different weak value predictions. We represent pairs of $\rho$ and $E$ as connected points on the Bloch sphere. Indeed, predictions outside the spectral range of the operator are accompanied by large weak values of $\rho$ and $E$. In agreement with the pure state case, large weak values of $\sigma_z$ do not occur when $\rho_t$ or $E_{t+\Delta t}$ are close to the $\sigma_z$ eigenstates, but rather when they are close to opposite $\sigma_x$ eigenstates.

In conclusion, we have demonstrated the use of the quantum trajectory formalism to infer the quantum state of a superconducting qubit conditioned on the outcome of continuous measurement. We have also demonstrated a quantum hindsight effect, where probing of a quantum system modifies and improves the predictions about measurements already performed in the past. These advances may be used to improve the state preparation and readout fidelity for quantum systems and increase their potential for use as probes [12][17] of time-dependent interactions and parameter estimation.

*murch@physics.wustl.edu
I. MEASUREMENT CALIBRATION

We calibrate the measurement strength by recording measurement signals for the qubit prepared in the $|0\rangle$ and $|1\rangle$ states. Measurement results are Gaussian distributed and we scale the signal such that the distributions are centered at $+1$ and $-1$ for the $|0\rangle$ and $|1\rangle$ states respectively. The variance, $a^2 = 1/4k\eta\Delta t$ is related to the measurement strength $k$, the measurement quantum efficiency $\eta$, and the integration time $\Delta t$.

\[ P(V_t|0)) = \frac{1}{\sqrt{2\pi a}} \exp\left(-\frac{(V_t + 1)^2}{2a^2}\right), \quad P(V_t|1)) = \frac{1}{\sqrt{2\pi a}} \exp\left(-\frac{(V_t - 1)^2}{2a^2}\right). \tag{8} \]

The measurement quantum efficiency is limited by losses in the microwave components and added noise from the amplifiers. The total quantum efficiency, $\eta_{tot} = \eta_{col}\eta_{amp}\eta_{env}$ also includes environmental decoherence of the qubit. Environmental inefficiency is small for the parameters of our experiment $\eta_{env} = (1 + \kappa/8\chi^2\bar{n}T^*_2)^{-1} = 0.95$, yet we include it in our calculation of $\rho$ and $E$ which is described below.

The measurement quantum efficiency $\eta = \eta_{col}\eta_{amp}$ was measured by fitting the distributions $P(V_t|0))$ and $P(V_t|1))$ to determine $a^2 = \kappa/(16\chi^2\bar{n}\Delta t)$, using $\bar{n} = 1.3$ and $\Delta t = 700$ ns. The dispersive coupling rate $\chi/2\pi = -0.43$ MHz was determined by using a Ramsey measurement to measure both the ac Stark shift $(2\chi\bar{n})$ and the measurement induced dephasing rate $(\Gamma_m = 8\chi^2\bar{n}/\kappa)$ for intracavity photon numbers ranging between $\bar{n} = 0$ and $\bar{n} = 1.3$. 

References

[1] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, Phys. Rev. 134, B1410 (1964).
[2] S. Watanabe, Rev. Mod. Phys. 27, 179 (1955).
[3] Y. Aharonov, S. Popescu, and J. Tollaksen, Physics Today 63, 27 (2010).
[4] Y. Aharonov, S. Popescu, and J. Tollaksen, Physics Today 64, 62 (2011).
[5] Y. Aharonov, S. Popescu, and J. Tollaksen, Physics Today 64, 9 (2011).
[6] L. Vaidman, Phys. Rev. A 87, 052104 (2013).
[7] S. Gammelmark, K. Mølmer, W. Alt, T. Kampschulte, and D. Meschede, Phys. Rev. A 89, 043839 (2014).
[8] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J.-L. O’Brien, A. G. White, and G. J. Pryde, Proc. Natl. Acad. Sci. U.S.A. 108, 1256 (2011).
[9] G. Waldherr, P. Neumann, S. F. Huelga, F. Jelezko, and J. Wrachtrup, Phys. Rev. Lett. 107, 090401 (2011).
[10] J. P. Groen, D. Ristè, L. Tornberg, J. Cramer, P. C. de Groot, T. Picot, G. Johansson, and L. DiCarlo, Phys. Rev. Lett. 111, 090506 (2013).
[11] P. Campagne-Ibarcq, L. Bretheau, E. Flurin, A. Auffèves, F. Mallet, and B. Huard, Phys. Rev. Lett. 112, 180402 (2014).
[12] M. Tsang, Phys. Rev. Lett. 102, 250403 (2009).
[13] M. Tsang, Phys. Rev. A 80, 033840 (2009).
[14] M. Tsang, H. M. Wiseman, and C. M. Caves, Phys. Rev. Lett. 106, 090401 (2011).
[15] M. A. Armen, A. E. Miller, and H. Mabuchi, Phys. Rev. Lett. 103, 173601 (2009).
[16] T. A. Wheatley, D. W. Berry, H. Yonezawa, D. Nakane, H. Arao, D. T. Pope, T. C. Ralph, H. M. Wiseman, A. Furusawa, and E. H. Huntington, Phys. Rev. Lett. 104, 093601 (2010).
[17] T. Rybarczyk, S. Gerlich, B. Peaudecerf, M. Penasa, B. Julsgaard, K. M. Imer, S. Gleyzes, M. Brune, J.-M. Raimond, S. Haroche, et al., arXiv p. 1409.0958 (2014).
[18] S. Gammelmark, B. Julsgaard, and K. Mølmer, Phys. Rev. Lett. 111, 160401 (2013).
[19] H. Wiseman and G. Milburn, Quantum Measurement and Control (Cambridge University Press, 2010).
[20] K. W. Murch, S. J. Weber, C. Macklin, and I. Siddiqi, Nature 502, 211 (2013).
[21] S. J. Weber, A. Chantasri, J. Dressel, A. N. Jordan, K. W. Murch, and I. Siddiqi, Nature 511, 570D573 (2014).
[22] J. Koch, T. M. Yu, J. Gambetta, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A 76, 042319 (2007).
[23] H. Paik, D. I. Schuster, L. S. Bishop, G. Kirchmair, G. Catelani, A. P. Sears, B. R. Johnson, M. J. Reagor, L. Frunzio, L. I. Glazman, et al., Phys. Rev. Lett. 107, 240501 (2011).
[24] M. Hatridge, R. Vijay, D. H. Slichter, J. Clarke, and I. Siddiqi, Phys. Rev. B 83, 134501 (2011).
[25] M. Hatridge, S. Shankar, M. Mirrahimi, F. Schackert, K. Geerlings, T. Brecht, K. M. Siwa, B. Abd, L. Frunzio, S. M. Girvin, et al., Science 339, 178 (2013).
[26] R. Vijay, D. H. Slichter, and I. Siddiqi, Phys. Rev. Lett. 106, 110502 (2011).
[27] Further details regarding the propagation of $\rho$ and $E$, the experimental setup, and data analysis are given in supplemental information.
[28] K. Jacobs and D. A. Steck, Contemp. Phys. 47, 279 (2006).
[29] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. 60, 1351 (1988).
[30] H. M. Wiseman, Phys. Rev. A 65, 032111 (2002).
II. CALCULATION OF THE DENSITY AND EFFECT MATRICES

The density matrix is calculated by propagating the stochastic master equation for our quantum system [1],

$$\frac{d\rho}{dt} = -i\frac{\hbar}{2} [\sigma_y, \rho] + \left( k + \frac{\gamma}{2} \right) (\sigma_z \rho \sigma_z - \rho) + 2\eta k (\sigma_z \rho + \rho \sigma_z - 2 \text{Tr}(\sigma_z \rho)) V_i. \tag{9}$$

This is the same as equation (4) in the main text with the exception that we include qubit dephasing which is characterized by the rate $\gamma = 1/T^*_2$. We propagate $\rho$ forward from the initial state $\rho_i = \frac{x_i \sigma_x}{2} + \frac{(I + z_i \sigma_z)}{2}$, where $x_i = 0.864$, $z_i = -0.058$ are determined from quantum state tomography at $t = 0$. These values differ from the ideal initial state $\rho_i = \frac{\sigma_x}{2} + \frac{I}{2}$ due to a time delay between the herald measurement and the start of data collection. The Rabi frequency is $\Omega/2\pi = 0.7$ MHz and we use time steps of $\delta t = 20$ ns.

The effect matrix $E(t)$ obeys a corresponding equation [2],

$$\frac{dE}{dt} = i\frac{\hbar}{2} [\sigma_y, E] + \left( k + \frac{\gamma}{2} \right) (\sigma_z E \sigma_z - E) + 2\eta k (\sigma_z E + E \sigma_z - \text{Tr}(\sigma_z E)) V_i dt. \tag{10}$$

Here we propagate $E$ backward in time from the final state $E_T = I$, and since $\text{Tr}(E) = 2$, the last terms in [9] and [10] are different by a factor of two. Since expressions involving $E$ are normalized, this factor is not essential, but we incorporate it to maintain a consistent value of the trace.

Our previous work [3] used a Bayesian argument to calculate the quantum trajectory for the density matrix $\rho$. In this work, we have focused on application of the stochastic master equation to connect with previous theoretical work. We find that the two techniques are equivalent when the continuous measurement is very weak, and the two techniques give very similar results for the parameters used in this experiment. Figure 5 displays a comparison of the two methods for a single trajectory.

III. EXPERIMENTAL SETUP

Figure 6 displays a detailed experimental schematic for the experiment.

The qubit was fabricated from double-angle-evaporated aluminum on silicon and is characterized by charging energy $E_C/h = 250$ MHz and Josephson energy $E_J/h = 9$ GHz. We obtain $E_C$ from the anharmonicity of the qubit levels and calculate $E_J$ from the relation $\hbar \omega_q = \sqrt{8 E_J E_C} - E_C$. The qubit frequency is $\omega_q/2\pi = 4.0033$ GHz, yet all experiments were performed with a constant intracavity photon number $\bar{n} = 1.3$ which resulted in an ac Stark shift of 1.05 MHz. All qubit rotations were performed at the ac Stark shifted frequency, 4.00225 GHz. We measured the qubit coherence properties, $T_1 = 30\mu s$, $T^*_2 = 16\mu s$ using standard techniques.

The lumped-element Josephson parametric amplifier (LJPA) is composed of a 1.5 pF capacitor shunted by a SQUID composed of two $I_0 = 1\mu A$ Josephson junctions. The LJPA is operated with negligible flux threading the SQUID loop and produces 20 dB of gain with an instantaneous bandwidth of 9 MHz.
IV. CONDITIONAL AVERAGING

The experimental verification of the predictions made by $\rho$ and $E$ in figures 1, 2, and 3 in the main text uses the concept of conditional averaging. In this section, we describe the procedures used in these figures in greater detail than space allows in the main text.

In figure 1, we present a trajectory for prediction of the outcome of a projective measurement of the qubit in the $\sigma_z$ basis, based on a single experimental sequence with a duration of $2 \mu s$ between the preparation and the final measurement $M$. This trajectory is obtained by propagating the stochastic master equation to obtain $\rho(t)$. We denote this as $\rho_{\text{target}}(t)$, the target trajectory. We confirm that the trajectory for $P(+z)$ is correct in the following manner. For each time point $t$ of the trajectory we conduct 3000 experiments with duration $t$ between the preparation and final measurement $M$. Of these 3000 experiments, we examine the subset of experiments that had values $\rho_t$ that were within $\pm 0.02$ of $\rho_{\text{target}}(t)$. Using this subset, we determine $P(+z)$ based on the corresponding subset of measurements.

Figure 6: Experimental schematic.
VI. SMOOTHED PREDICTIONS OF PROJECTIVE MEASUREMENTS

By increasing the strength of the measurement $\Omega_V$, we may ultimately perform a projective measurement of the qubit state in the $\sigma_z$ basis. In figure 8 we display histograms of $P(z)$ and $P_p(z)$, indicating the predicted and
Figure 8: The occurrence of different values of $P(+)z$ and $P_p(+)z$ obtained from many iterations of the experiment shown in grey and black respectively. We propagate $\rho$ for 800 ns from the initial state $\text{Tr}(\rho, \sigma_x) \approx 1$ and propagate $E$ for 800 ns from the final state $E = \hat{I}$. Panel (a) shows the case where $\Omega_R = 0$ and (b) shows the case where $\Omega_R/2\pi = 0.7 \text{ MHz}$. Both graphs show that the Eq.(2) in the main text makes more confident predictions for the outcomes of projective measurements by more often taking values closer to 0 and 1.

smoothed probability for finding the qubit in its ground state. We observe more occurrences of values of $P_p(+)z$ than of $P(+)z$ near 0 and 1 indicating that the past quantum state more often makes confident predictions about the outcome of a projective measurement.

When $\Omega_R = 0$ the stronger predictions given by the past quantum state are a consequence of the quantum non-demolition (QND) character of the measurement. The effects of the measurements commute and the past quantum state analysis merely accumulates measurements in the intervals $(0, t)$ and $(t + \Delta t, T)$. However, by setting $\Omega_R \neq 0$, we break the QND character of the measurement, and it is necessary to propagate $\rho$ and $E$ using their associated stochastic master equations.

[1] H. Wiseman and G. Milburn, *Quantum Measurement and Control* (Cambridge University Press, 2010).
[2] S. Gammelmark, B. Julsgaard, and K. Mølmer, Phys. Rev. Lett. **111**, 160401 (2013).
[3] S. J. Weber, A. Chantasri, J. Dressel, A. N. Jordan, K. W. Murch, and I. Siddiqi, Nature **511**, 570 (2014).