Greedy versus map-based optimized adaptive algorithms for random-telegraph-noise mitigation by spectator qubits

Behnam Tonekaboni [1,*] Areeya Chantasri [2,3,†] Hongting Song [3,4,‡] Yanan Liu [3,§] and Howard M. Wiseman [3,‖]

[1]*Centre for Quantum Dynamics, Griffith University, Yuggera Country, Brisbane, Queensland 4111, Australia*
[2]*Optical and Quantum Physics Laboratory, Department of Physics, Faculty of Science, Mahidol University, Bangkok 10400, Thailand*
[3]*Centre for Quantum Computation and Communication Technology (Australian Research Council), Centre for Quantum Dynamics, Griffith University, Yuggera Country, Brisbane, Queensland 4111, Australia*
[4]*Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology, Beijing 100094, China*

In a scenario where data-storage qubits are kept in isolation as far as possible, with minimal measurements and controls, noise mitigation can still be done using additional noise probes, with corrections applied only when needed. Motivated by the case of solid-state qubits, we consider dephasing noise arising from a two-state fluctuator, described by random telegraph process, and a noise probe which is also a qubit, a so-called spectator qubit (SQ). We construct the theoretical model assuming projective measurements on the SQ, and derive the performance of different measurement and control strategies in the regime where the noise mitigation works well. We start with the Greedy algorithm; that is, the strategy that always maximizes the data qubit coherence in the immediate future. We show numerically that this algorithm works very well, and find that its adaptive strategy can be well approximated by a simpler algorithm with just a few parameters. Based on this, and an analytical construction using Bayesian maps, we design a one-parameter (Θ) family of algorithms. In the asymptotic regime of high noise sensitivity of the SQ, we show analytically that this Θ family of algorithms reduces the data qubit decoherence rate by a divisor scaling as the square of this sensitivity. Setting Θ equal to its optimal value, $\Theta^\star$, yields the Map-based Optimized Adaptive Algorithm for Asymptotic Regime (MOAAAR). We show, analytically and numerically, that MOAAAR outperforms the Greedy algorithm, especially in the regime of high noise sensitivity of SQ.

I. INTRODUCTION

Qubit decoherence from environmental noises is one of the major obstacles to building useful large-scale quantum computers [1–4]. In the past years, a wide range of noise-mitigation techniques have been proposed to prolong the coherence time of computational qubits. Some of the techniques, such as the dynamical decoupling (DD), aim at removing average effects of noise continuously in time, via applying carefully designed controls onto qubits [5–12]. Other techniques, such as those that fit in the category of quantum error correction (QEC), are designed to monitor effects of noise and correct errors via information-redundancy encoding [13–15].

However, both DD and QEC approaches require direct controls or measurements on the qubits of interest, which could in turn create more sources of noise. Therefore, in the particular case where it is possible to keep the computing (*data*) qubits well isolated from their environment, new techniques, that can both correct noise effects while minimizing direct contact to the data qubits, are required. Such techniques would demand an additional entity, e.g., a sensor or a probe, in proximity of the data qubits, that can sense the problematic noise and any correction to the data qubits can be applied only when needed.

In this paper, we are interested in the recently proposed idea of spectator qubits (SQs) [16–18], which are qubit-type probes that are assumed to be much more sensitive to target noise than data qubits and can be easily measured. This could be applicable in solid-state quantum computing, where different types of qubits with very different properties can co-exist. For example, the data qubits could be the spin qubits with low sensitivity to noise [19–21] and SQs could be quantum dots nearby with an ability of high-efficiency readout [22–24]. Assuming a well-isolated spin-donor qubit in silicon substrate, there will still be lingering effect from low-frequency noise, such as charge noise, which can affect qubits at different locations simultaneously [25,26] and is considered difficult to remove [27]. Therefore, we are interested in constructing a theoretical model using the data-spectator setup to investigate a plausible optimal regime for the noise mitigation. It turns out that the solution is not trivial, even with a simple regime with single data and spectator qubits.

We have presented our most important findings on this problem in a companion letter (CL) [28]. Specifically, we showed that by performing projective measurements on the

―――――
*behnam.tfn@gmail.com
†Corresponding author: areeya.chn@mahidol.ac.th
‡shtfc@163.com
§yanan.liu@griffith.edu.au
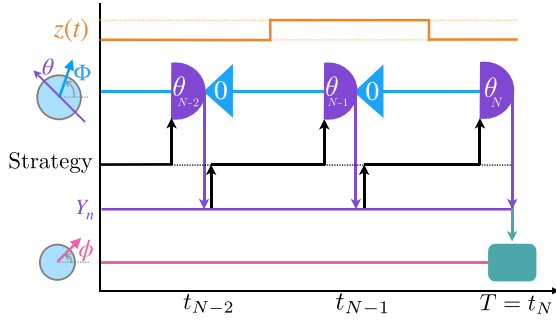‖h.wiseman@griffith.edu.au

FIG. 1. Schema for the adaptive measurement strategy on SQ. The orange line shows the dynamics of RTP, which jumps between two values indicated with pale dashed orange lines. The blue line represents the SQ, which is reset (triangle) after each measurement (purple semicircle). Stored measurement results are represented by purple arrows toward the measurement record line Y_n (purple). These measurement records are used via the adaptive strategy (black arrows) to decide the next measurement time t_n and next measurement angle θ_n. Last, at final time T, the data qubit (pink) is corrected using information Y_N (downward teal arrow).

SQ at certain times and in certain bases, that are chosen adaptively in an optimized manner, the decoherence of the data qubit can be suppressed by an amount which is quadratic in the SQ sensitivity. This result holds under ideal conditions, and the optimal suppression is quite nontrivial, as it is better than that can be achieved by simply maximizing the data qubit coherence in each, arbitrarily small, time step (Greedy algorithm). In this paper, we give the detailed derivations of the results in the CL [28], and explore in much more depth the comparison between the Greedy algorithm and the optimized algorithm. But to explain more precisely the contents of this paper, it is necessary first to give some more theoretical background.

We model the decoherence of the data qubit as being caused by low-frequency phase noise arising from the random motion of a trapped charge caused by impurities, as is common in the experimental solid-state qubit setting [27,29–32]. Specifically, we model the noise $z(t)$ as a two-state fluctuator [33–36], with values ± 1 as shown in Fig. 1. This can be mathematically described by the random telegraph process (RTP), with two flip rates: $\gamma_\uparrow$ and $\gamma_\downarrow$ [37,38]. To monitor the noise, a SQ is placed nearby, experiencing the same phase noise, but with the higher sensitivity. This is described by a total Hamiltonian of both qubits,

$$\hat{H}_{\text{tot}} = \frac{\kappa}{2}\hat{\sigma}_z^{\text{d}} z(t) + \frac{K}{2}\hat{\sigma}_z^{\text{s}} z(t). \tag{1}$$

Here κ and K are the data qubit's and SQ's noise sensitivities and $\hat{\sigma}_z^{\text{d,s}}$ are their respective z-Pauli matrices. Also, we are working in a rotating frame for both qubits (their frequencies need not be the same or even similar) and with units where $\hbar = 1$.

We assume that, in order to obtain information about the noise $z(t)$, the SQ can be measured after arbitrary waiting times $\{\tau_1, \tau_2, \ldots\}$, with arbitrary angles $\{\theta_1, \theta_2, \ldots\}$, and reprepared in an arbitrary state $|\theta\rangle^{\text{s}}$. Here θ always indicates an angle on the equator of the Bloch sphere, defining either

a Pauli observable, $\hat{\sigma}_\theta^{\text{s}} \equiv \cos\theta\, \hat{\sigma}_x^{\text{s}} + \sin\theta\, \hat{\sigma}_y^{\text{s}}$, or a state $|\theta\rangle^{\text{s}}$ which is the $+1$ eigenstate of $\hat{\sigma}_\theta^{\text{s}}$. The resulting string of measurement readouts can then be used in estimating the data qubit phase correction at some final time T, which is when the data qubit is required. The goal is to maximize the resultant coherence of the data qubit, averaging over all possible realizations of the noise.

Given this model and goal, the question is how one should choose the τ_n's, the θ_n's, and the final phase correction. In this work, we apply the *separation principle* from control theory, and the structure of the reward function (which defines the goal), to greatly simplify the problem. We show that the information necessary for all future controls can be stored in a complex 2-vector and propagated from one measurement to the next by a *Bayesian map*, a complex 2×2 matrix. We then use the maps to numerically find the locally optimal—also known as *Greedy*—algorithm.

Locally optimal algorithms, in general, are not always globally optimal, and we find that to be the case for the problem here. We are interested in the regime where the SQ approach is most useful, to wit,

$$T^{-1}, \kappa \ll \gamma_{\uparrow,\downarrow} \ll K. \tag{2}$$

We analyze how the Greedy algorithm works in this regime, and this suggests a simple, single parameter (Θ) family of models where $\tau_n = \Theta/K$ and $\theta_n = \pm\Theta$. Here the sign of θ_n is determined *adaptively*, and reflects the algorithm's best estimate of $z(t_n)$, the current state of RTP. In the asymptotic regime (2), we can solve analytically for the performance of this family of algorithms. Choosing the optimal parameter value $\Theta^\star$ yields a measurement and control strategy that is plausibly optimal, and which we call MOAAAR (Map-based Optimized Adaptive Algorithm in the Asymptotic Regime). This can reduce the decoherence rate of the data qubit, relative to the no control rate, by the multiplicative factor

$$H^\star (\bar{\gamma}/K)^2, \quad \text{where } H^\star \approx 1.254, \tag{3}$$

where $\bar{\gamma} := (\gamma_\uparrow + \gamma_\downarrow)/2$. That is, as mentioned above, the data qubit lifetime enhancement offered by the MOAAAR algorithm scales like the square of the sensitivity of the SQ. While the same quadratic scaling is observed for the Greedy algorithm, we find that the prefactor $H^\star$ for the decoherence rate with MOAAAR is smaller than that from Greedy.

The structure of this paper is as follows. First, in Sec. II, we briefly discuss the physics and mathematical description of the RTP. We then, in Sec. III, calculate the dephasing of the data qubit in the absence of any control, and formulate it as a matrix product, which will be useful later. We end this section by looking at how, in general, additional information can be used to improve the coherence in the data qubit. In Sec. IV we construct the setup and measurement on the SQ and discuss the notion of adaptive measurement strategy in the context of the SQ. Then, in Sec. V, we use Bayesian statistics and formulate the expected coherence using the gathered information via the SQ. We write the Bayesian map as a matrix product which updates a *coherence vector* after each measurement on the SQ. We also introduce a phase space formalism which includes all the sufficient information. We then optimize our measurement strategy using the locally optimization algorithm (Greedy) in Sec. VI and the global

optimization approach (MOAAAR) in Sec. VII. In Sec. VIII we compare these two optimization strategies by constructing a closed-form formula for each that applies even outside the asymptotic regime. We conclude our work with a discussion of future research directions in Sec. IX.

II. CHARGE NOISE AND RANDOM TELEGRAPH PROCESS (RTP)

Charge noise is one of the causes of decoherence in a wide range of physical systems including superconducting qubits [39], quantum dots [40], Nitrogen vacancy centers [41], trapped ions [42], and semiconductors [27,31,33,35,43,44]. The noise is often described by stochastic motion of electrons or holes trapped at impurities in the devices, which can be mathematically modeled as a random telegraph process (RTP), whose value jumps between two values [37,38]. Consider the RTP noise parameter $z_t := z(t)$, which can be either $+1$ or -1. It makes a jump from $z = -1$ to $z = +1$ with rate $\gamma_\uparrow$, and $z = +1 \rightarrow -1$ with rate $\gamma_\downarrow$. Define the probability vector

$$\underline{P}_t := \begin{pmatrix} \wp(z_t = +1) \\ \wp(z_t = -1) \end{pmatrix}, \quad (4)$$

where $\wp(z_t = \pm 1)$ is the probability of $z_t = \pm 1$ at time t. Then, with the jump rates defined above, one may write a (classical) master equation for vector $\underline{P}_t$ as

$$\dot{\underline{P}}_t = \mathbf{J} \underline{P}_t, \quad (5)$$

where $\mathbf{J}$ is the jump rate matrix,

$$\mathbf{J} = \begin{pmatrix} -\gamma_\downarrow & +\gamma_\uparrow \\ +\gamma_\downarrow & -\gamma_\uparrow \end{pmatrix}. \quad (6)$$

If the RTP evolves for a duration of τ, between times t' and $t'' = t' + \tau$, the master equation (5) has a solution of the form $\underline{P}_{t''} = e^{\mathbf{J}\tau} \underline{P}_{t'}$. One may rearrange the matrix exponential and write the solution as

$$\underline{P}_{t''} = \left(\mathbf{I} + \frac{1 - e^{-2\bar{\gamma}\tau}}{2\bar{\gamma}} \mathbf{J} \right) \underline{P}_{t'}. \quad (7)$$

where $\bar{\gamma} = (\gamma_\uparrow + \gamma_\downarrow)/2$ as before, and $\mathbf{I}$ is the identity matrix. It is worth mentioning that when $\tau \rightarrow \infty$, the probability vector $\underline{P}_{t''}$ converges to its steady state given by

$$\underline{P}_{ss} = \frac{1}{2\bar{\gamma}} \begin{pmatrix} \gamma_\uparrow \\ \gamma_\downarrow \end{pmatrix}. \quad (8)$$

This will be used throughout the paper as the state of RTP when no information about the noise is available, such as at the initial time.

We model the effect of the charge noise on the data and spectator qubits by the Hamiltonian in Eq. (1). From this, it is apparent that, under free evolution in some time interval (t', t''), the qubits' accumulated phase (angle of rotation around the $\hat{\sigma}_z$ axis) will be proportional to the integral of the noise in that interval,

$$x := \int_{t'}^{t''} z(s) \, ds. \quad (9)$$

To determine the dephasing caused by this random accumulated phase (more on this in the next section) it is useful to calculate the average, $\int \wp(x) e^{ikx} \, dx$, of the exponential factor, e^{ikx}, for arbitrary k. Actually, it is even more useful to calculate the integral $\int e^{ikx} \wp(x, z_{t''}|z_{t'}) \, dx$ for arbitrary k, $z_{t''}$, and $z_{t'}$. This is, of course, just the Fourier transform of $\wp(x, z_{t''}|z_{t'})$, and can be calculated using Eqs. (7) and (9) as

$$\int e^{ikx} \wp(x, z_{t''}|z_{t'}) \, dx = \exp(-\bar{\gamma}\tau) \begin{cases} \cosh\left(\frac{\lambda}{2}\tau\right) - s\frac{\eta}{\lambda} \sinh\left(\frac{\lambda}{2}\tau\right), & \text{for } z_{t'} = z_{t''} = s \\ \frac{2\gamma_\uparrow}{\lambda} \sinh\left(\frac{\lambda}{2}\tau\right), & \text{for } z_{t''} = -z_{t'} = +1 \\ \frac{2\gamma_\downarrow}{\lambda} \sinh\left(\frac{\lambda}{2}\tau\right), & \text{for } z_{t''} = -z_{t'} = -1 \end{cases} \quad (10)$$

where $s \in \{-1, +1\}$ and

$$\lambda(k) = \sqrt{(\gamma_\downarrow + \gamma_\uparrow)^2 - 4ik(\gamma_\downarrow - \gamma_\uparrow) - 4k^2}, \quad (11a)$$

$$\eta(k) = (\gamma_\downarrow - \gamma_\uparrow) - 2ik \quad (11b)$$

are functions of the Fourier variable k (see Appendix A and Refs. [35,45] for the details of its derivation). In the following sections, we will use these results to derive the statistics of data and spectator phases, conditioned on measurement results and/or qubit controls.

III. DATA QUBIT DEPHASING

In this section, we define the data qubit's coherence and analyze its dephasing due to the RTP noise, both for the cases with no control and with the control phase correction. For the no-control case, we show that the qubit's coherence can be

written as a matrix equation involving a complex two-by-two matrix, $\mathbf{H}$. The matrix representation for the no-control coherence will be extended to the version including measurement strategies and readouts of the spectator qubit in Sec. V.

A. Formulation of data qubit coherence

From the total Hamiltonian in Eq. (1), if we include the control (phase correction) applied at the final time, $t = T$, we can write the Hamiltonian of the data qubit as

$$\hat{H}_d(t) = \frac{\kappa}{2} \hat{\sigma}_z^d z(t) + \hat{H}_{ctrl}\delta(t - T). \quad (12)$$

The first term of the Hamiltonian describes a stochastic phase rotation around the z axis of the data qubit's Bloch sphere. To evaluate the data qubit's decoherence due to this term, we consider a qubit's density matrix in the z basis affected by a

random noise,

$$\rho^{\mathrm{d}}(\Xi) = \rho_{++}|+1\rangle_z^{\mathrm{d}}\langle+1| + \rho_{+-}e^{-i\phi(\Xi)}|+1\rangle_z^{\mathrm{d}}\langle-1|$$
$$+ \rho_{-+}e^{+i\phi(\Xi)}|-1\rangle_z^{\mathrm{d}}\langle+1| + \rho_{--}|-1\rangle_z^{\mathrm{d}}\langle-1|, \quad (13)$$

where $|+1\rangle_z^{\mathrm{d}}$ and $|-1\rangle_z^{\mathrm{d}}$ are the two $\hat{\sigma}_z^{\mathrm{d}}$ eigenstates, $\hat{\sigma}_z^{\mathrm{d}}|\pm1\rangle_z^{\mathrm{d}} = \pm|\pm1\rangle_z^{\mathrm{d}}$. The terms $\rho_{++}, \rho_{--}, \rho_{+-}, \rho_{-+}$ are density matrix elements of the data qubit's state at the initial time ($t = 0$), before affected by the noise. The Hamiltonian (12) thus introduces an additional phase, denoted by $\phi = \phi(\Xi)$, which is a function of a set of variables, Ξ, which can include both the noise that affects the data qubit's phase, and the control applied to it.

Since the value of Ξ varies from run to run, we are interested in the average final state of the qubit, $\langle\rho^{\mathrm{d}}(\Xi)\rangle_\Xi$, where the subscript Ξ indicates that this is the random variable with respect to which the average is taken. The data qubit's final purity is completely determined by the magnitude of the off-diagonal elements of Eq. (13), which gives $|\rho_{+-}\langle e^{i\phi(\Xi)}\rangle_\Xi|$. Since the element ρ_{+-} is not affected by the noise, we can thus define the quantity to be maximized, which we call the *coherence*, as

$$\mathcal{C} := |\langle e^{i\phi(\Xi)}\rangle_\Xi|. \quad (14)$$

The expected average of random phases typically leads to the value of coherence being less than one (its maximum value) and can be as low as zero for the case of uniformly distributed phases in $[0, 2\pi)$. The reduction of coherence from random phases is called qubit dephasing. It is also convenient later to define a *complex coherence*,

$$\mathcal{A} := \langle e^{i\phi(\Xi)}\rangle_\Xi, \quad (15)$$

without the absolute value, where $\mathcal{C} = |\mathcal{A}|$. The information of the argument of $\mathcal{A}$ will be useful later, in estimating the phase information of the data qubit.

Let us define capital-letter variables for a total accumulated noise, from time $t = 0$ to any time t,

$$X := \int_0^t z(s)\,ds, \quad (16)$$

and for a total information,

$$Y := \text{all acquired information about } X, \quad (17)$$

which is from any measurements up to the time t. In the following subsections, we simplify the representation of a data qubit's state in Eq. (13) to showing only its phase. Let us define a state of the data's qubit phase as

$$|\phi\rangle^{\mathrm{d}} := \frac{1}{\sqrt{2}}\left(|+1\rangle_z^{\mathrm{d}} + e^{i\phi}|-1\rangle_z^{\mathrm{d}}\right). \quad (18)$$

Then the initial phase of the data qubit is $|\phi = 0\rangle^{\mathrm{d}}$ (a zero noisy phase), and we consider two cases: the no-control case and the case with the noise correction. For the former, $\hat{H}_{\mathrm{ctrl}} = 0$, the total data noisy phase will be a function of noise, $\phi(\Xi) = \phi(X)$. For the latter, we introduce a phase-correction control, $\hat{H}_{\mathrm{ctrl}} = -ic(Y)\hat{\sigma}_z^{\mathrm{d}}/2$, which describes an instantaneous rotation of the data qubit's state around the z axis with an angle $-c(Y)$ that depends on the information, Y. In this case, the data qubit's phase will be a function of both the noise and the control, $\phi(\Xi) = \phi(X, Y)$. We will show how

to compute the data qubit's coherence for both cases in the following subsections.

B. Decay of no-control (nc) coherence

For the "no-control" (nc) case, $\hat{H}_{\mathrm{ctrl}} = 0$, the data qubit only evolves stochastically due to the RTP noise. Given the Hamiltonian, Eq. (12), the data qubit's state (represented by its noisy phase) at any time t, becomes

$$|\phi(\Xi)\rangle^{\mathrm{d}} = \exp\left(-i\frac{\kappa}{2}\hat{\sigma}_z^{\mathrm{d}}\int_0^t ds\, z(s)\right)|\phi = 0\rangle^{\mathrm{d}} = |\kappa X\rangle^{\mathrm{d}}, \quad (19)$$

which means that the data noisy phase, $\phi(\Xi) = \phi(X) = \kappa X$, is simply proportional to the total accumulated noise. Following Eq. (14), replacing Ξ with X, we can calculate the no-control coherence as

$$\mathcal{C}^{\mathrm{nc}} := |\langle e^{i\phi(X)}\rangle_X| = \left|\int e^{i\kappa X}\wp(X)\,dX\right|, \quad (20)$$

and its no-control complex coherence

$$\mathcal{A}^{\mathrm{nc}} := \int e^{i\kappa X}\wp(X)\,dX, \quad (21)$$

where $\wp(X)$ is the probability density function of the accumulated noise.

Since the variable X is an accumulated RTP noise, it is convenient to write the probability function $\wp(X)$ as marginalized over the initial $z_0 = z(t = 0)$ and the end point $z_t = z(t)$ of RTP,

$$\wp(X) = \sum_{z_t}\sum_{z_0}\wp(X, z_t|z_0)\wp(z_0). \quad (22)$$

Substituting the above equation into Eq. (20), we find that the no-control coherence is

$$\mathcal{C}^{\mathrm{nc}} = \left|\sum_{z_t}\sum_{z_0}\left[\int e^{i\kappa X}\wp(X, z_t|z_0)\,dX\right]\wp(z_0)\right|, \quad (23)$$

where the term inside the square brackets is the Fourier form as we defined in Eq. (10). Now, we see that Eq. (23) can be written as a map-based equation by defining a vector

$$\underline{A}_0 := \underline{P}_{t=0} = \begin{pmatrix}\wp(z_0 = +1)\\\wp(z_0 = -1)\end{pmatrix}, \quad (24)$$

using the definition in Eq. (4) and defining matrix elements,

$$H_{z_0}^{z_t}(t, \kappa) := \int e^{i\kappa X}\wp(X, z_t|z_0)\,dX. \quad (25)$$

Since $z_0, z_t \in \{-1, +1\}$, this defines a two-by-two matrix,

$$\mathbf{H}(t, \kappa) := \begin{pmatrix}H_{+1}^{+1}(t, \kappa) & H_{-1}^{+1}(t, \kappa)\\H_{+1}^{-1}(t, \kappa) & H_{-1}^{-1}(t, \kappa)\end{pmatrix}. \quad (26)$$

Thus, the coherence in Eq. (23) is then written as

$$\mathcal{C}^{\mathrm{nc}} = |\underline{l}^\top\mathbf{H}(t, \kappa)\underline{A}_0|, \quad (27)$$

where $\underline{l}^\top = (1, 1)$. We can calculate the coherence Eq. (27) for the no-control case analytically (see details in Appendix A). As we will see later, reformulating the coherence

as matrix equations simplifies analytical and numerical calculations for optimal controls, particularly when measurements on the spectator qubit are involved.

To simplify the discussion later, let us also define a complex two-dimensional coherence vector for the no-control case,

$$\underline{A}^{\mathrm{nc}} = \mathbf{H}(t, \kappa)\underline{A}_0 = \begin{pmatrix} \wp\,(z_t = +1)\mathcal{A}_{|z_t=+1} \\ \wp\,(z_t = -1)\mathcal{A}_{|z_t=-1} \end{pmatrix}. \qquad (28)$$

Here we have introduced a complex *conditional* coherence,

$$\mathcal{A}_{|z_t} := \int e^{i\kappa X}\wp\,(X|z_t)\,dX = \langle e^{i\kappa X}\rangle_{X|z_t}, \qquad (29)$$

of the data qubit. This quantity is conditioned on (hypothetically) finding out the value of the RTP z_t at time t; later we will condition on other data, that may be hypothetical or actual or a combination of both. We can get back the no-control complex coherence in Eq. (21) by

$$\mathcal{A}^{\mathrm{nc}} = \underline{I}^{\top}\underline{A}^{\mathrm{nc}} = \underline{I}^{\top}\mathbf{H}(t, \kappa)\underline{A}_0 = \sum_{z_t}\wp\,(z_t)\mathcal{A}_{|z_t}, \qquad (30)$$

where the sum is over $z_t \in \{-1, 1\}$, showing the connection among matrix equations and the complex coherence. The absolute value of this complex coherence is the no-control coherence as shown in Eq. (27).

C. Decay of coherence in asymptotic regime

Here, we show that, in the long-time limit, $t \to \infty$, and the asymptotic regime, where κ is much smaller than the RTP jump rate, $\kappa \ll \bar{\gamma}$, the no-control coherence $\mathcal{C}^{\mathrm{nc}}$ decays exponentially. We start with the no-control complex coherence, $\mathcal{A}^{\mathrm{nc}}$ in Eq. (21), and calculate the ratio of the change over an infinitesimal time and its value. That is, the ratio is defined as

$$\Lambda_t := \frac{d\mathcal{A}^{\mathrm{nc}}/dt}{\mathcal{A}^{\mathrm{nc}}}, \qquad (31)$$

which in general can be a time-dependent function denoted by the subscript t. We use the RTP steady state as the initial vector $\underline{A}_0 = \underline{P}_{t=0} = \underline{P}_{\mathrm{ss}}$ in (8), and expand the right-hand side of Eq. (31) to second order in κ and find

$$\Lambda_t = i(\gamma_\uparrow - \gamma_\downarrow)\frac{\kappa}{2\bar{\gamma}} + (e^{-2t\bar{\gamma}} - 1)\frac{\kappa^2\breve{\gamma}}{2\bar{\gamma}^2} + O(\kappa^3), \qquad (32)$$

where $\breve{\gamma} \equiv 2\gamma_\uparrow\gamma_\downarrow/(\gamma_\uparrow + \gamma_\downarrow)$ is the harmonic mean of $\gamma_\uparrow$ and $\gamma_\downarrow$. Taking the long-time limit, i.e., $t \to \infty$, of above the expansion, we find that the factor Λ_t approaches a constant complex value,

$$\Lambda = \lim_{t\to\infty}\Lambda_t = -\Gamma^{\mathrm{nc}} + i\Omega^{\mathrm{nc}}, \qquad (33)$$

where, in the last line, we have defined the average no-control frequency (phase drift rate) and decay rate (phase diffusion rate) as

$$\Gamma^{\mathrm{nc}} := \frac{\kappa^2\breve{\gamma}}{2\bar{\gamma}^2} \quad \text{and} \quad \Omega^{\mathrm{nc}} := \frac{\kappa(\gamma_\uparrow - \gamma_\downarrow)}{2\bar{\gamma}}. \qquad (34)$$

As a result, we substitute Λ_t in Eq. (31) with the constant Λ and solve the differential equation for the no-control complex coherence to get

$$\mathcal{A}^{\mathrm{nc}}(t) = \mathcal{A}^{\mathrm{nc}}(0)\exp\left(-\Gamma^{\mathrm{nc}}t + i\Omega^{\mathrm{nc}}t\right), \qquad (35)$$

which has the initial condition $\mathcal{A}^{\mathrm{nc}}(0) = \underline{I}^{\top}\underline{P}_{\mathrm{ss}} = 1$. We note that the complex coherence has its imaginary contribution only when the RTP has asymmetric jump rates. That is, for the symmetric case, $\gamma_\uparrow = \gamma_\downarrow$, we have $\Omega^{\mathrm{nc}} = 0$. Finally, since $\mathcal{C}^{\mathrm{nc}} = |\mathcal{A}^{\mathrm{nc}}|$, we obtain the no-control coherence in the asymptotic regime,

$$\mathcal{C}^{\mathrm{nc}}(t) = \exp(-\Gamma^{\mathrm{nc}}t). \qquad (36)$$

In the CL [28] (Fig. 1), we show the comparison between Eq. (36) and the exact coherence in (27). They differ significantly at short times, this is expected since Eq. (33) is valid only in the limit $t \gg 1/\bar{\gamma}$. Moreover, even by $t = 1/\bar{\gamma}$ one can see that their decay is almost identical.

D. Noise correction and optimal choice of control

The data qubit's coherence can be improved, from the no-control case, by utilizing additional information gathered by a SQ. Before discussing how to obtain this information (in the following section), we will show how additional information, in general, helps increase the coherence of the data qubit.

Let us assume that we have access to an additional information Y about the unknown accumulated noise X. We can use the information to compute an appropriate phase correction, $-c(Y)$, and apply it to the data qubit as described in Sec. III A, so that $\phi(\Xi) = \phi(X, Y) = \kappa X - c(Y)$. Given the general form of coherence in Eq. (14), and a control function $c(\bullet)$, we write a phase-corrected coherence as

$$\mathcal{C}_{c(\bullet)} := |\langle e^{i[\kappa X - c(Y)]}\rangle_{X,Y}|, \qquad (37)$$

which is similar to Eq. (20), but now the expected average is over all possible values of X and Y.

If X were known, we could trivially find that $c(Y) = \kappa X$ is a perfect correction that completely removes the phase error from the noise and the data qubit's coherence is maximized at one. However, the noise is actually unknown and we only have its indirect information from Y. Therefore, the main task is to find the control function, $c(\bullet)$, that maximizes the coherence, $\mathcal{C}_{c(\bullet)}$ in Eq. (37). To accomplish this, we first write the expected average in Eq. (37) as an explicit summation over values for Y (assumed discrete) and an integral over the continuous value of X with appropriate probability weights as

$$\begin{aligned} \mathcal{C}_{c(\bullet)} &= \left|\sum_Y \int e^{i\kappa X}e^{-ic(Y)}\wp\,(Y)\wp\,(X|Y)\,dX\right| \\ &= \left|\sum_Y e^{-ic(Y)}\wp\,(Y)\langle e^{i\kappa X}\rangle_{X|Y}\right| \\ &\leqslant \sum_Y \wp\,(Y)|\langle e^{i\kappa X}\rangle_{X|Y}|, \end{aligned} \qquad (38)$$

where, in the second line, we have rearranged terms into a conditional expected average, $\langle e^{i\kappa X}\rangle_{X|Y}$. In the last line, the inequality is, in general, a strict inequality. However, it becomes an equality when

$$c(\bullet) = \arg\langle e^{i\kappa X}\rangle_{X|\bullet}, \qquad (39)$$

which thus maximizes the phase-corrected coherence $\mathcal{C}_{c(\bullet)}$. Given the formula of the coherence in Eq. (37), we can regard Eq. (39) as defining the *optimal* estimator of the real phase, $\phi = \kappa X$, for the purpose of the final phase correction. Indeed, it is just the mean under *circular statistics* [46]. We indicate the maximized *control coherence* defined as

$$\mathcal{C}^{\mathrm{c}} := \mathcal{C}_{c(\bullet)=\arg\mathcal{A}_{|\bullet}} = \sum_Y \wp(Y)|\mathcal{A}_{|Y}|. \qquad (40)$$

Here we are again considering the conditional complex coherence,

$$\mathcal{A}_{|Y} := \langle e^{i\kappa X}\rangle_{X|Y} = \int e^{i\kappa X}\wp(X|Y)\,dX, \qquad (41)$$

this time conditioned on the information Y.

IV. SPECTATOR QUBIT (SQ) FOR NOISE SENSING

In our model, the information Y used in the phase correction can be obtained by measuring the SQ which senses the same RTP noise affecting the data qubit. The Hamiltonian of the SQ,

$$\hat{H}_{\mathrm{s}}(t) = \frac{K}{2}\hat{\sigma}_z^{\mathrm{s}}z(t), \qquad (42)$$

is identical to that of the data qubit, but with K as the SQ's noise sensitivity in place of the much smaller κ for the data qubit. Thus the SQ's Bloch vector also rotates around the z axis. We assume that we have complete control over the SQ. To maximize its usefulness as a noise probe, we take it be reset to an equatorial state after each measurement. Thus its state can always be written as

$$|\Phi\rangle^{\mathrm{s}} = \frac{1}{\sqrt{2}}\left(|+1\rangle_z^{\mathrm{s}} + e^{i\Phi}|-1\rangle_z^{\mathrm{s}}\right), \qquad (43)$$

given the eigenstates $|\pm 1\rangle_z^{\mathrm{s}}$ of the $\hat{\sigma}_z^{\mathrm{s}}$ observable, $\hat{\sigma}_z^{\mathrm{s}}|\pm 1\rangle_z^{\mathrm{s}} = \pm|\pm 1\rangle_z^{\mathrm{s}}$. The definition in Eq. (43) should be compared to Eq. (18) for the data qubit.

A. SQ's measurement and likelihood function

To acquire information about the RTP noise, we probe the SQ at various times throughout the process. Let us denote the measurement times by

$$t_1, t_2, \ldots, t_n, \ldots, t_{N-1}, t_N = T, \qquad (44)$$

where T is the final time that the phase correction is applied on the data qubit. Given the SQ's Hamiltonian (42), the integrated noise over time intervals of duration

$$\tau_n := t_n - t_{n-1} \qquad (45)$$

can be independently probed. Without loss of generality, we initialize the SQ at the zero-phase state, $|\Phi = 0\rangle^{\mathrm{s}}$, at the initial time, $t_0 = 0$, and after the nth measurement, for $n \in \{1, N-1\}$.

Just before the nth measurement, at time t_n, the SQ's state is given by $|\Phi(t_n)\rangle^{\mathrm{s}}$, where

$$\Phi(t_n) = K\int_{t_{n-1}}^{t_n} dt\, z(t) = K(X_n - X_{n-1}) := Kx_n. \qquad (46)$$

Here we have defined the integrated noise up to time t_n,

$$X_n := \int_0^{t_n} z(s)\,ds, \qquad (47)$$

similar to Eq. (16), and defined $x_n := X_n - X_{n-1}$, similar to Eq. (9). We also define $z_n := z(t_n)$ for the RTP at the measurement time t_n.

We assume that the measurement of the SQ at time t_n is projective. Given the parameter of interest, $\Phi(t_n)$, the optimal observable to measure will be of the form

$$\hat{\theta}_n := \mathbb{1} - |\theta_n\rangle^{\mathrm{s}}\langle\theta_n|. \qquad (48)$$

Here θ_n is the chosen measurement angle and $|\theta_n\rangle^{\mathrm{s}} = (1/\sqrt{2})(|+1\rangle_z^{\mathrm{s}} + e^{i\theta_n}|-1\rangle_z^{\mathrm{s}})$, and $\mathbb{1} = |+1\rangle_z^{\mathrm{s}}\langle+1| + |-1\rangle_z^{\mathrm{s}}\langle-1|$ is the identity operator. The outcomes of the projective measurement,

$$y_n \in \{0, 1\}, \qquad (49)$$

are associated with projecting the SQ's state to the two eigenstates, $\{|\theta_n\rangle^{\mathrm{s}}, |\theta_n + \pi\rangle^{\mathrm{s}}\}$, respectively. We define $\hat{\theta}_n$, Eq. (48), in this way so that the outcomes have a significance as null and nonnull outcomes, respectively, for a natural choice of θ_n, as we will see in Sec. VI C.

Given the SQ's state, $|\Phi(t_n)\rangle^{\mathrm{s}} = |Kx_n\rangle^{\mathrm{s}}$, Born's rule gives the probability of outcome y_n as

$$\begin{aligned}\wp(y_n|\theta_n, x_n) &:= |\,{}^{\mathrm{s}}\langle\theta_n + \pi\, y_n|Kx_n\rangle^{\mathrm{s}}\,|^2 \\ &= y_n + (-1)^{y_n}\cos^2\left[\tfrac{1}{2}(\theta_n - Kx_n)\right].\end{aligned} \qquad (50)$$

In the context of Bayesian estimation (see next section), this is the *likelihood function*, expressed slightly differently as Eq. (9) of the CL [28].

B. Adaptive strategies and the separation principle

Say that one has just measured and reprepared the SQ at time t_{n-1}. There are two real parameters that define the next measurement on it: the measurement angle, θ_n, and the waiting time, τ_n. Let us combine them and define—for any n—a *measurement setting*

$$\mu_n := \{\theta_n, \tau_n\}, \qquad (51)$$

and define a *measurement strategy* as

$$\mathfrak{S} = \{\mu_n : n \in \{1, \ldots, N\}\}. \qquad (52)$$

The aim of our work is to search for measurement strategies, $\mathfrak{S}$, that maximizes the data qubit's coherence.

In particular, to be general, we must include adaptive strategies, where, at a local measurement time $t_{n-1} \in \{t_1, \ldots, t_{N-1}\}$, one can use the information

$$Y_{n-1} := \{y_1, y_2, \ldots, y_{n-1}\}, \qquad (53)$$

obtained in the past up to t_{n-1}, to find the optimal choice for the next measurement at t_n. That is, we can explicitly write a measurement setting, $\mu_n(Y_{n-1})$, as a function,

$$\mu_n : \{0, 1\}^{(n-1)} \mapsto U(1) \times \mathbb{R}_+, \qquad (54)$$

that maps measurement results, $Y_{n-1} \in \{0, 1\}^{(n-1)}$ (in a set of $(n-1)$-string of binary numbers), to a measurement angle,

$\theta_n \in U(1)$ (in a set of the 1D-circular group), and a waiting time, $\tau_n \in \mathbb{R}_+$ (in a set of positive real numbers).

From the above, it is apparent that the space of possible strategies $\{\mathfrak{S}\}$ is very large for N large. However, we can simplify the problem of trying to find an optimal strategy by applying the *separation principle* [47,48], as mentioned in the introduction. First, let us formally define the *expected reward function*:

$$\mathcal{C}^c(T) = \sum_{Y_N} \wp(Y_N)|\mathcal{A}_{|Y_N}|, \qquad (55)$$

which is the control coherence, Eq. (40), for the information, $Y = Y_N$, at the final time $T = t_N$. It is implicit that the complex conditional coherence $\mathcal{A}_{|Y_N}$, as in Eq. (41), is defined with $X = X_N$, i.e., at the same time ($T = t_N$) as Y_N. We also note that the summation, $\sum_{Y_N}$, is over $Y_N \in \{0, 1\}^{(N)}$ and that $\wp(Y_N) = \wp(y_1, y_2, \ldots, y_N)$ is a joint probability of the string of readouts. Now, the separation principle says that if the reward function is additive in time—which ours, Eq. (55), is trivially since it is evaluated at the final time T only—then the optimization problem separates into two parts.

The first part is optimal estimation of the system. This means finding, at any time t_n, the Bayesian probability for the relevant parameters conditioned on the results obtained so far. That is, one can construct a conditional distribution $\wp_n \equiv \wp(X_n, z_n|Y_n)$ for the RTP noise (z_n) and its accumulated one (X_n) at time t_n conditioned on the current information Y_n. The second part is determining the optimal control based on this information. That is, instead of considering $\mu_{n+1}(Y_n)$, we can consider $\mu_{n+1}(\wp_n)$. Moreover, the goal of the control in the future of t_n (e.g., at the final time T) is to maximize the expected reward function *given* Y_n, which is

$$\mathcal{C}^c_{|Y_n}(T) := \sum_{Y_N|Y_n} \wp(y_{n+1}, \ldots, y_N|Y_n)|\mathcal{A}_{|Y_N}|, \qquad (56)$$

where the summation is defined as over possible *future* measurement outcomes, $Y_N|Y_n := (y_{n+1}, \ldots, y_N)$.

Now, given that $\wp_n$ is a *function* of a real variable X_n and a binary variable z_n, it is not immediately obvious that replacing $\mu_{n+1}(Y_n)$ by $\mu_{n+1}(\wp_n)$ is progress. However, as we will see in the next section (after considerable calculation), for the case of our particular conditional expected reward function Eq. (56), only a couple of statistics derived from the distribution $\wp(X_n, z_n|Y_n)$ are relevant (see Sec. V D), which greatly simplifies the problem.

Before turning to the details of that calculation, we here also introduce a more general definition of the conditional control coherence:

$$\mathcal{C}^c_{|D_n}(t_m) := \sum_{Y_m} \wp(Y_m|D_n)|\langle e^{i\kappa X_m}\rangle_{X_m|Y_m}|, \qquad (57)$$

which will be of use in later sections. Here the generalizations beyond (57) are twofold. First, we allow conditioning on an arbitrary set of data, D_n, which is available at measurement time t_n, not necessarily equal to the full record Y_n up to that time. Second, the time at which we imagine implementing the control on the data qubit is any measurement time t_m,

with $m \geqslant n$, rather than necessarily the final time $t_N = T$ as in Eq. (56).

V. BAYESIAN MAPS FOR PHASE ESTIMATION

In this section, we formulate a Bayesian map, which is central to our results, especially the MOAAAR algorithm of Sec. VII. We start with the expected reward function, Eq. (55), and show that it can be written as a matrix equation, similar to that of the no-control case, with the map **H** in Eq. (27). In contrast to the no-control case, where there is only one map for the entire process, the coherence must be updated whenever new information from the SQ measurements at $t \in \{t_1, t_2, \ldots, t_N\}$ are obtained. We will see that the new map, denoted by $\mathbf{F}(\mu_n = \{\theta_n, \tau_n\}, y_n)$, is an explicit function of the measurement setting, μ_n, and the readout, y_n, and can be derived analytically using Bayes' theorem.

A. Bayes' theorem for the expected reward function

The expected reward function, defined in Eq. (55), can be explicitly written as

$$\begin{aligned} \mathcal{C}^c(T) &= \sum_{Y_N} \wp(Y_N)|\mathcal{A}_{|Y_N}|, \\ &= \sum_{Y_N} \wp(Y_N)\left|\int dX_N \wp(X_N|Y_N)e^{i\kappa X_N}\right|, \\ &= \sum_{Y_N} \left|\int dX_N \wp(Y_N, X_N)e^{i\kappa X_N}\right|, \end{aligned} \qquad (58)$$

where X_N is the final accumulated noise,

$$X_N = x_1 + x_2 + \cdots + x_N, \qquad (59)$$

and Y_N is a string of measurement results from the SQ,

$$Y_N = (y_1, y_2, \ldots, y_N), \qquad (60)$$

up until the final time.

The integral in Eq. (58) is not trivial since we do not know an explicit form of the joint probability function $\wp(Y_N, X_N)$. However, we have the likelihood functions $\wp(y_n|\theta_n, x_n)$ and $\wp(x_n, z_n|z_{n'})$, via Eqs. (50) and (10). respectively. Thus, we write Eq. (58) in terms of these likelihood functions, by first defining a noise vector, $\vec{x} := \{x_1, \ldots, x_N\}$, and its 1-norm $||\vec{x}||_1 = \sum_n x_n$. Using Eq. (59) allows us to write

$$\begin{aligned} \wp(Y_N, X_N) &= \int \wp(Y_N, \vec{x})\delta(X_N - ||\vec{x}||_1)\,d^N\vec{x}, \\ &= \int \wp(Y_N|\vec{x})\wp(\vec{x})\delta(X_N - ||\vec{x}||_1)\,d^N\vec{x}. \end{aligned} \qquad (61)$$

The integral measure is $d^N\vec{x} \equiv dx_1 \cdots dx_N$. Substituting the above joint distribution back into Eq. (58) and integrating over X_N, we obtain

$$\mathcal{C}^c(T) = \sum_{Y_N} \left|\int \wp(Y_N|\vec{x})\wp(\vec{x})e^{i\kappa||\vec{x}||_1}\,d^N\vec{x}\right|. \qquad (62)$$

For the joint distribution, $\wp(Y_N|\vec{x})$, since we reset the SQ after every measurement, each result y_n depends only on the

accumulated noise x_n during the waiting time τ_n. Thus, we can write

$$\wp(Y_N|\vec{x}) \equiv \wp(y_1, \ldots, y_N|x_1, \ldots, x_N),$$

$$= \prod_{n=1}^{N} \wp_{\mu_n}(y_n|x_n), \tag{63}$$

where we have added the measurement setting $\mu_n = \{\theta_n, \tau_n\}$ as a subscript of the likelihood function (50). Furthermore, we construct a joint probability of the accumulated noises, $x_1, \ldots, x_N$, with the help of RTP variables and Eq. (10), and write

$$\wp(\vec{x}) \equiv \wp(x_1, \ldots, x_N),$$

$$= \sum_{z_0, \ldots, z_N} \left(\prod_{n=1}^{N} \wp(x_n, z_n|z_{n-1}) \right) \wp(z_0), \tag{64}$$

as a marginalized probability function over the RTP noises, $z_0, \ldots, z_N$, at all times. Here the sum $\sum_{z_0, \ldots, z_N}$ is over all possible values of the RTP. We then substitute Eqs. (63) and (64) into the reward function in Eq. (62) and get

$$\mathcal{C}^c(T) = \sum_{Y_N} \wp(Y_N) \left| \mathcal{A}_{|Y_N} \right| = \sum_{Y_N} \left| \sum_{z_0, \ldots, z_N} \left(\prod_{n=1}^{N} \int dx_n \, \wp_{\mu_n}(y_n|x_n) \wp(x_n, z_n|z_{n-1}) e^{i\kappa x_n} \right) \wp(z_0) \right|,$$

$$= \sum_{Y_N} \left| \sum_{z_0, \ldots, z_N} \left(\prod_{n=1}^{N} F_{z_{n-1}}^{z_n}(\mu_n, y_n) \right) \wp(z_0) \right|, \tag{65}$$

where in the second line we have defined

$$F_{z_{n-1}}^{z_n}(\mu_n, y_n) := \int dx_n \, \wp_{\mu_n}(y_n|x_n) \wp(x_n, z_n|z_{n-1}) e^{i\kappa x_n}. \tag{66}$$

These are the linear maps that are the core elements of our Bayesian-map formalism. In the next subsection, we will derive an analytic expression for $F_{z_{n-1}}^{z_n}(\mu_n, y_n)$, and rewrite it explicitly as a matrix.

B. Analytical expression for *F* map

In the similar manner as in the construction of the matrix $\mathbf{H}$ in Eq. (26), we see that the sums and products in Eq. (65) can be rewritten as matrix multiplication. Specifically, defining the two-by-two matrix

$$\mathbf{F}(\mu_n, y_n) := \begin{pmatrix} F_{+1}^{+1}(\mu_n, y_n) & F_{-1}^{+1}(\mu_n, y_n) \\ F_{+1}^{-1}(\mu_n, y_n) & F_{-1}^{-1}(\mu_n, y_n) \end{pmatrix}, \tag{67}$$

we can write the expected reward function Eq. (65) as

$$\mathcal{C}^c(T) = \sum_{Y_N} |\underline{I}^\top \mathbf{F}(\mu_N, y_N) \cdots \mathbf{F}(\mu_2, y_2) \mathbf{F}(\mu_1, y_1) \underline{A}_0|$$

$$= \sum_{Y_N} |\underline{I}^\top \underline{A}_N|, \tag{68}$$

where $\underline{A}_0$ and $\underline{I}$ are defined in Eqs. (24) and (27). Also, in the second line of Eq. (68), we defined a new coherence vector

$$\underline{A}_N \equiv \mathbf{F}(\mu_N, y_N) \cdots \mathbf{F}(\mu_2, y_2) \mathbf{F}(\mu_1, y_1) \underline{A}_0. \tag{69}$$

Moreover, we can use convolutions and Fourier transform (see details in the Appendix B) to find that the matrices, $\mathbf{F}$ and $\mathbf{H}$, are related via

$$\mathbf{F}(\mu_n = \{\theta_n, \tau_n\}, y_n)$$

$$= \tfrac{1}{4}[2\,\mathbf{H}(\tau_n, \kappa) + (-1)^{y_n} e^{-i\theta_n} \mathbf{H}(\tau_n, \kappa + K)$$

$$+ (-1)^{y_n} e^{+i\theta_n} \mathbf{H}(\tau_n, \kappa - K)]. \tag{70}$$

One can easily confirm from Eq. (70) that

$$\sum_{y_n} \mathbf{F}(\mu_n, y_n) = \mathbf{H}(\tau_n, \kappa), \tag{71}$$

which can be understood that the no-control map $\mathbf{H}$ (no measurement information) is equivalent to ignoring the gained information from the conditional map $\mathbf{F}(\mu_n, y_n)$.

C. Interpretation of complex 2-vector $\underline{A}_n$

Following the definition of $\underline{A}_N$ in Eq. (69), we define a coherence vector $\underline{A}_n = (A_n^{z=+1}, A_n^{z=-1})^\top$ for an intermediate time t_n. This can be computed from the vector at the previous time step using the measurement result y_n via

$$\underline{A}_n = \mathbf{F}(\mu_n, y_n) \underline{A}_{n-1}. \tag{72}$$

We can interpret the two elements of $\underline{A}_n$, similar to Eq. (28), as

$$\underline{A}_n = \begin{pmatrix} \underline{A}_n^{+1} \\ \underline{A}_n^{-1} \end{pmatrix} = \begin{pmatrix} \wp(Y_n, z_n = +1) \mathcal{A}_{|Y_n, z_n = +1} \\ \wp(Y_n, z_n = -1) \mathcal{A}_{|Y_n, z_n = -1} \end{pmatrix}. \tag{73}$$

That is, the coherence vector $\underline{A}_n$ contains information about the measurement results $Y_n = (y_1, \ldots, y_n)$ and the complex conditional coherence defined as

$$\mathcal{A}_{|Y_n, z_n} := \langle e^{i\kappa X_n} \rangle_{X_n|Y_n, z_n}, \tag{74}$$

conditioned on both the results and the hypothetically knowable z_n, at time t_n. With this, we can see a relationship between two types of complex conditional coherences, in Eq. (41) and Eq. (74), via

$$\wp(Y_n) \mathcal{A}_{|Y_n} = \underline{I}^\top \underline{A}_n = \sum_{z_n} \wp(Y_n, z_n) \mathcal{A}_{|Y_n, z_n}, \tag{75}$$

where the summation is over the RTP state $z_n \in \{-1, +1\}$.

The vector $\underline{A}_n$ does not let one calculate exactly the probability of the measurement results, $\wp(Y_n)$. However, if one considers κ (the sensitivity of the data qubit to noise) to be a variable rather than a fixed number, then one can take the

limit $\kappa \to 0$ in Eq. (74). This gives $\lim_{\kappa\to 0} \mathcal{A}_{|Y_n,z_n} = 1$ and, from Eq. (73), we see that

$$\wp(Y_n) = \lim_{\kappa\to 0} \underline{I}^\top \underline{A}_n. \tag{76}$$

This motivates defining a probability vector,

$$\underline{\check{A}}_n = \lim_{\kappa\to 0} \underline{A}_n = \begin{pmatrix} \wp(Y_n, z_n = +1) \\ \wp(Y_n, z_n = -1) \end{pmatrix}, \tag{77}$$

which should not be confused with the unconditioned RTP probability vector $\underline{P}_t$.

Similar to the coherence vector $\underline{A}_n$ in Eq. (72), the probability vector can be updated after every SQ measurement via

$$\underline{\check{A}}_n = \check{\mathbf{F}}(\mu_n, y_n)\underline{\check{A}}_{n-1}, \tag{78}$$

where we have defined a new probability map with $\kappa \to 0$,

$$\check{\mathbf{F}}(\mu_n, y_n) := \lim_{\kappa\to 0} \mathbf{F}(\mu_n, y_n). \tag{79}$$

This $\check{\mathbf{F}}$ matrix contains only real numbers and thus can map between two real-number vectors containing only probability functions. Interestingly, the initial state of the probability vector is the same as $\underline{A}_0$ and the RTP initial state, $\underline{P}_0$,

$$\underline{\check{A}}_0 = \underline{A}_0 = \underline{P}_0 = \underline{P}_{\rm ss}. \tag{80}$$

Here we have also set them to be equal to the RTP steady state as per Eq. (8). We will see later that $\underline{\check{A}}_n$ and $\wp(Y_n)$ are useful to both understanding dynamics of the SQ and simulating the data qubit's coherence numerically.

D. Two real sufficient statistics: (α, ζ)

As mentioned in Sec. IV B, the separation principle allows us to separate the problem at any time t_n into the estimation part, using the information Y_n obtained up to that time, and the control part, which uses that estimate to maximize the expected future cost function, conditioned on the present Y_n. Here we show that, for our cost function, the estimation problem reduces to calculating the complex 2-vector $\underline{A}_n$. Furthermore, except at the final time T, just two real numbers derived from $\underline{A}_n$ are sufficient for the control problem.

We start with the expected reward function conditioned on Y_n given in Eq. (56), which we rewrite here for convenience:

$$\mathcal{C}^{\rm c}_{|Y_n}(T) := \sum_{Y_N|Y_n} \wp(y_{n+1}, \ldots, y_N|Y_n)\big|\mathcal{A}_{|Y_N}\big|. \tag{81}$$

We then use Bayes' rule to write $\wp(y_{n+1}, \ldots, y_N|Y_n) = \wp(Y_N)/\wp(Y_n)$ and use the relationship $\wp(Y_N)|\mathcal{A}_{|Y_N}| = |\underline{I}^\top \underline{A}_N|$ from Eq. (75) to obtain

$$\mathcal{C}^{\rm c}_{|Y_n}(T) = [\wp(Y_n)]^{-1} \sum_{Y_N|Y_n} |\underline{I}^\top \underline{A}_N|, \tag{82a}$$

$$= [\wp(Y_n)]^{-1} \sum_{Y_N|Y_n} |\underline{I}^\top \mathbf{F}(\mu_N, y_N) \cdots$$

$$\mathbf{F}(\mu_{n+2}, y_{n+2})\mathbf{F}(\mu_{n+1}, y_{n+1})\underline{A}_n|, \tag{82b}$$

where in the second line we used Eq (69), only expanding out the $\mathbf{F}$ maps for the future outcomes $y_{n+1}, \ldots, y_N$.

Now, to maximize this reward function, the only things we have the ability to control, in the future of t_n, in Eq. (82b)

is the SQ measurement strategy, $\mu_{n+1}, \ldots, \mu_N$. What current information is relevant to this strategy? First, since Y_n is already known, $\wp(Y_n)$ is a fixed scalar multiple, and so is not relevant. That leaves only the vector $\underline{A}_n$. In other words, if we know $\underline{A}_n$, we know *everything* relevant for choosing the SQ measurement strategy to maximize the final objective $\mathcal{C}^{\rm c}_{|Y_n}(T)$. From Eq. (40), it also contains all the relevant present information for calculating, at the final time, the optimal control to apply to arrive at Eq. (82b). Here we are concerned with the question "what parts of $\underline{A}_n$ are relevant for the SQ measurement strategy?" The complex-2 vector $\underline{A}_n$ defines four real parameters. It turns out that only two of them are relevant as we now show.

Since the two (complex) elements of $\underline{A}_n$ are associated with two plausible values of the unknown RTP, $z_n \in \{-1, +1\}$ as in Eq. (73), one might guess that these elements should tell us something relevant about the data qubit's phase and coherence if one could, hypothetically, find out which state z_n the RTP is actually in. One such quantity is the scaled ratio of the two elements of $\underline{A}_n$,

$$\alpha_n := \frac{K}{\kappa} \arg \frac{A_n^{z=+1}}{A_n^{z=-1}}, \tag{83a}$$

$$= \frac{K}{\kappa}\big(\arg \mathcal{A}_{|Y_n,+1} - \arg \mathcal{A}_{|Y_n,-1}\big), \tag{83b}$$

using the definition in Eq. (73). This quantity α_n is the first of the two sufficient statistics and can be interpreted as the *maximum difference of an optimal control phase* from finding out if the RTP state was $z_n = +1$ or $z_n = -1$, scaled so as to be of order unity.

To understand Eq. (83a), we recall that the optimal control (phase correction) is given by $c(Y) = \arg \mathcal{A}_{|Y}$ from Eq. (39). If, at a current time t_n, a phase correction, $c(Y_n)$, were to be applied to the data qubit, by further assuming the hypothetical z_n, one could consider refining the control to

$$c'(Y_n, z_n) = \arg \mathcal{A}_{|Y_n,z_n} = \arg \langle e^{i\kappa X_n}\rangle_{X_n|Y_n,z_n}, \tag{84}$$

where z_n is included in the condition of the expected average. The bracketed term in Eq. (83b) is therefore the difference of the control phases for the two hypothetical z_n. To understand the scaling factor that makes this phase difference of order unity, $O(1)$, it is necessary to consider the dynamics of the RTP and the measurements, as we do in the following paragraph.

The value of z_n is relevant to the data qubit phase only for the recent period of evolution, of order $1/\check{\gamma}$. For any time earlier than that, the RTP state is pretty much uncorrelated with its current state. Thus the maximum difference that finding out z_n could make on the data qubit's phase estimate is the difference in the cumulative phase, κX_n, over that time scale, for $z_n = +1$ versus $z_n = -1$. That is, of order $\kappa/\check{\gamma}$, which is small. In actuality, when we are gathering information about z through a good adaptive protocol (see later sections) the difference is even smaller, because from the preceding measurement result we have a very good idea of the value of z. Hence if z is unknown that is substantially because it may have changed its value since the last measurement. In other words, the preceding time scale for which finding out the value of z is relevant is of order τ, not of order $1/\check{\gamma}$. As we will see later,

the optimal value for τ is of order $1/K$. Thus, the difference that finding out z_n could make for the phase estimate is of order κ/K. Hence, to scale the phase difference to the order unity, we divide it by the factor κ/K and get α_n as in Eq. (83a).

The second of the two sufficient statistics is related to the modulus of the two elements of $\underline{A}_n$, defined as

$$\zeta_n := \frac{\left|A_n^{z=+1}\right| - \left|A_n^{z=-1}\right|}{\left|A_n^{z=+1}\right| + \left|A_n^{z=-1}\right|}. \tag{85}$$

This can be interpreted as an *approximated mean of z_n conditioned on Y_n*. To see this, we consider the conditional complex coherence, $\mathcal{A}_{|Y_n, \pm 1}$. As we have already seen above, their arguments are very close, differing by only $O(\kappa/K)$ as per Eq. (83b). But this implies that their moduli must also be very close. In fact, the relative difference in their moduli is quadratically smaller than the difference in their arguments. That is because the relative change in the data qubit coherence $|\mathcal{A}|$ over some short time interval is quadratic in the uncertainty in the data qubit's phase. (One can see this by Taylor expanding the decoherence to the second order of the phase.) But the *difference* in the uncertainty in the data qubit's phase, for the different values of z_n, will not be larger than the corresponding difference in their mean, because the difference between the two RTP values is the maximum uncertainty the RTP value can have.

Therefore, in the relevant time for the refined best-estimate phase shift, $\tau = O(1/K)$, the difference in the relative decoherence resulting from refining one's knowledge to $z = +1$ versus $z = -1$ would be at most of order $O((\kappa/K)^2)$. That is, for a good measurement protocol, we can be confident that $|\mathcal{A}_{|Y_n, +1}|/|\mathcal{A}_{|Y_n, -1}| = 1 + O((\kappa/K)^2)$. With relative errors of the same magnitude, we can thus replace $|\mathcal{A}_{|Y_n, z_n}|$ by $|\mathcal{A}_{|Y_n}|$ in Eq. (73), to get

$$\left|A_n^{z_n}\right| \approx \left|\mathcal{A}_{|Y_n}\right| \wp(Y_n, z_n). \tag{86}$$

Hence, ζ_n in Eq. (85) can be approximated as

$$\zeta_n \approx \frac{\sum_{z_n} z_n \wp(Y_n, z_n)}{\sum_{z_n} \wp(Y_n, z_n)}, \tag{87}$$

which is exactly the mean of z_n conditioned on the current measurement record, Y_n.

There are two more (third and fourth) parameters from the vector $\underline{A}_n$ that are not relevant to finding the optimal SQ control. The third parameter is the argument of the sum of elements,

$$\varphi_n := \arg \mathcal{A}_{|Y_n} = \arg \left(A_n^{z=+1} + A_n^{z=-1}\right). \tag{88}$$

This would be relevant as the final control on the *data* qubit, $c(Y_n)$, if $t_n = T$ were the final time. However, because t_n is not the end of the protocol, the quantity simply expresses how far the mean phase has drifted so far. Since the random phase driven by the RTP is a true $U(1)$ process, having an estimate of the phase at any particular time is not relevant to future decoherence. One can see this from Eq. (82b) that a global phase change to the vector $\underline{A}_n$, i.e., changing the value of φ_n, does not affect the coherence on the left of the equation.

Fourth, we can define the final real parameter from the sum of moduli of the elements as

$$r_n := \left|A_n^{z=+1}\right| + \left|A_n^{z=-1}\right| \approx \wp(Y_n)\left|\mathcal{A}_{|Y_n}\right|, \tag{89}$$

TABLE I. Four (real) statistics parameters along with their physical explanations. The vector $\underline{A}_n$ is described by these parameters.

Parameter	Description
r_n, Eq. (89)	$\approx Y_n$ probability $\times$ conditional coherence.
φ_n, Eq. (88)	= optimal control phase, not knowing z_n.
ζ_n, Eq. (85)	$\approx$ mean of z_n conditioned on Y_n.
α_n, Eq. (83a)	= maximum difference that z_n could make to optimal control phase, scaled.

where we have used Eq. (86) with the same degree of approximation as discussed around that equation. As already argued, $\wp(Y_n)$ is irrelevant to optimizing the future control, and so is any loss of coherence already suffered such that $|\mathcal{A}_{|Y_n}| < 1$, by a similar argument used for φ_n above.

We summarize the meaning of the four real parameters that make up the vector $\underline{A}_n$ in Table I. Another way to summarize them is to note that in the regime of interest it is possible to approximate $\underline{A}_n$ by a simple expression using the above four parameters,

$$A_n^{z=\pm 1} \approx r_n \exp(i\varphi_n) \times \frac{1 \pm \zeta_n}{2} \exp\left(\pm i \frac{1 \mp \zeta_n}{2} \frac{\alpha_n \kappa}{K}\right). \tag{90}$$

Here the approximation sign allows for a *relative* error in r_n of $O((\kappa/K)^2)$ and an *absolute* error in φ_n of $O((\kappa/K)^3)$. This expression is for interest only; we do not use it in the remainder of the paper.

Returning to the two sufficient statistics, $\alpha_n \in \mathbb{R}$ in Eq. (83a) and $\zeta_n \in [-1, 1]$ in Eq. (85), these can thus be regarded as the only two moments of $\wp(X_n, z_n|Y_n)$ (the solution to the estimation part of the problem) that are relevant to the SQ measurement part of the problem. Thus, as mooted in Sec. IV B, we can vastly reduce in size the set of control protocols $\mathfrak{S}$ Eq. (52) that we consider by simplifying the functional form of μ_{n+1} from that in Eq. (54) to

$$\mu_{n+1} : \mathbb{R} \times [-1, 1] \mapsto U(1) \times \mathbb{R}_+. \tag{91}$$

That is, μ_{n+1} is now a function of (α_n, ζ_n) rather than of Y_n directly. Finding the exactly optimal control strategy $\{\mu_n : n \in \{1, \ldots, N\}\}$ is still a difficult task, especially as, in general, N will not, in fact, be fixed for a fixed T. However, if we seek a close-to-optimal, rather than exactly optimal, control strategy $\mathfrak{S}$, then this difficulty can be avoided. Because we are interested in the long-time limit, we will have $N \gg 1$, and for the vast majority of the control choices we can ignore the *edge effects* (where n is small or n is close to N), and choose the μ_{n+1} to all be the *same* function of their two arguments. That is, we have only to optimize a single function,

$$\mu : \mathbb{R} \times [-1, 1] \mapsto U(1) \times \mathbb{R}_+. \tag{92}$$

In the future sections, we will make further assumptions to restrict the function μ, and we will see how the stochastic behavior of α_n and ζ_n is greatly useful in developing the intuition to do this.

VI. LOCAL OPTIMIZATION ALGORITHM (GREEDY)

The first strategy we explore is a "greedy" one [49]. That is, it based on the natural idea of maximizing the expected reward

function $\mathcal{C}^c$ locally in time. This is easier than maximizing the actual reward function, which is $\mathcal{C}^c$ at the final time T. Greedy algorithms are not optimal for typical hard problems, but even if they are not, they may be relatively close to optimal. That appears to be the situation for our problem, as we will find. We call this strategy the Greedy algorithm, or simply "Greedy" for short.

We conceptualize the Greedy algorithm as follows. Say that we are at some time t such that $t_n < t \leqslant t_{n+1}$. By this we mean that exactly n measurements on the SQ have already happened and their measurement results, Y_n, are known, and it is still possible to decide to make the $(n+1)$th measurement at time t. Imagine now that the final time is the immediate future,

$$\mathfrak{T} = t + dt. \tag{93}$$

Then the Greedy algorithm aims to maximize the conditional expected reward function $\mathcal{C}^c_{|Y_n}(\mathfrak{T})$ as in Eq. (82a) where we have replaced T with $\mathfrak{T}$.

We emphasize that any—even random—measurement on the SQ provides some information that can be used to increase the coherence of the data qubit. Thus to maximize $\mathcal{C}^c_{|Y_n}(\mathfrak{T})$ at least one measurement on the SQ at time $\mathfrak{T} = t + dt$ should be performed. Greedy maximizes the expected reward function at the immediate future time, by deciding whether or not to measure at time t and optimizing the measurement angles at both times. In the following subsections, we present the details of Greedy in Sec. VI A and then demonstrate its numerical implementation in Sec. VI B. Studying the patterns in the numerically found Greedy algorithm in Sec. VI C then motivate us to propose and analyze various "experimentally" simpler greedy algorithms in Secs. VI D and VI F, where in Sec. VI E we analyze the phase space dynamics of the simpler greedy algorithms.

A. The Greedy algorithm to maximize $\mathcal{C}^c_{|Y_n}(\mathfrak{T})$

Let the current time be denoted $t = t_n + \tau$, with $\tau > 0$. The goal is to determine the next measurement setting, μ_{n+1}, based on the current information, Y_n. To do that, the Greedy algorithm compares the reward $\mathcal{C}^c_{|Y_n}(\mathfrak{T})$ from two prospective scenarios:

(i) The SQ is measured only at time $\mathfrak{T} = (t_n + \tau) + dt$,

(ii) The SQ is measured now, at $t = t_n + \tau$, and again at $\mathfrak{T} = (t_n + \tau) + dt$.

We then calculate the coherence vectors associated with the two scenarios at time $\mathfrak{T}$ using Eq. (72):

$$\underline{A}^{(\mathrm{i})} := \mathbf{F}(\mu_{(\mathrm{i})}, y_{n+1})\underline{A}_n, \tag{94a}$$

$$\underline{A}^{(\mathrm{ii})} := \mathbf{F}(\mu'_{(\mathrm{ii})}, y_{n+2})\mathbf{F}(\mu_{(\mathrm{ii})}, y_{n+1})\underline{A}_n, \tag{94b}$$

where y_{n+1} and y_{n+2} are measurement results at time t and $\mathfrak{T} = t + dt$ respectively. Scenario (i) consists of only one measurement with a setting $\mu_{(\mathrm{i})} := \{\theta_{(\mathrm{i})}, \tau + dt\}$ whereas scenario (ii) consists of two consecutive measurements with settings $\mu_{(\mathrm{ii})} := \{\theta_{(\mathrm{ii})}, \tau\}$ and $\mu'_{(\mathrm{ii})} := \{\theta'_{(\mathrm{ii})}, dt\}$ for the first and second measurements, respectively. Note that these angles should be thought of not as numbers but as *functions* of $\underline{A}_n$ and, for $\theta'_{(\mathrm{ii})}$, also of y_{n+1}.

Despite the fact that scenario (ii) has one more measurement than scenario (i), the coherence vectors are both calculated at the hypothetical final time $\mathfrak{T} = t + dt$. Thus, the reward functions—as it is defined in Eq. (82a)—for the two cases at time $\mathfrak{T}$ are

$$\mathcal{C}^c_{(\mathrm{i})}(\theta_{(\mathrm{i})}) := \mathcal{C}^c_{|Y_n}(\mathfrak{T})\big|_{(\mathrm{i})} = [\wp\,(Y_n)]^{-1} \sum_{y_{n+1}} |\underline{I}^\top \underline{A}^{(\mathrm{i})}|, \tag{95a}$$

$$\mathcal{C}^c_{(\mathrm{ii})}(\theta_{(\mathrm{ii})}, \theta'_{(\mathrm{ii})}) := \mathcal{C}^c_{|Y_n}(\mathfrak{T})\big|_{(\mathrm{ii})} = [\wp\,(Y_n)]^{-1} \sum_{y_{n+2}, y_{n+1}} |\underline{I}^\top \underline{A}^{(\mathrm{ii})}|, \tag{95b}$$

where on the left-hand sides of the above equations, we defined $\mathcal{C}^c_{(\mathrm{i})}$ and $\mathcal{C}^c_{(\mathrm{ii})}$ as functions of the prospective angles. We then find the optimal angles by maximizing both $\mathcal{C}^c_{(\mathrm{i})}$ and $\mathcal{C}^c_{(\mathrm{ii})}$ over all possible angles. Formally, the maximized reward functions are given by

$$\mathcal{C}^{\mathrm{op}}_{(\mathrm{i})} := \max_{\theta_{(\mathrm{i})}} \mathcal{C}^c_{(\mathrm{i})}(\theta_{(\mathrm{i})}), \tag{96a}$$

$$\mathcal{C}^{\mathrm{op}}_{(\mathrm{ii})} := \max_{\theta_{(\mathrm{ii})}, \theta'_{(\mathrm{ii})}} \mathcal{C}^c_{(\mathrm{ii})}(\theta_{(\mathrm{ii})}, \theta'_{(\mathrm{ii})}). \tag{96b}$$

We compare the optimum values in Eqs. (96) to decide whether or not to measure at time t. The decision-making procedure is as follows: if $\mathcal{C}^{\mathrm{op}}_{(\mathrm{i})} > \mathcal{C}^{\mathrm{op}}_{(\mathrm{ii})}$, it means that no measurement at time t is required to maximize the reward function at $\mathfrak{T}$. Thus, Greedy decides to wait and moves to the next infinitesimal time to repeat the procedure again. If, on the other hand, $\mathcal{C}^{\mathrm{op}}_{(\mathrm{i})} \leqslant \mathcal{C}^{\mathrm{op}}_{(\mathrm{ii})}$, Greedy chooses to measure the SQ at t, so that τ is the actual waiting time τ_{n+1} from t_n to $t_{n+1} = t$ and the measurement angle $\theta_{(\mathrm{ii})}$ is the actual θ_{n+1}. Thus, we write

$$\mu_{n+1} = \{\theta_{(\mathrm{ii})}, \tau\}, \tag{97}$$

as the next measurement setting chosen by the Greedy. We note that, the prefactor $[\wp\,(Y_n)]^{-1}$ in Eq. (95) does not influence the Greedy decision-making procedure and knowing only $\underline{A}_n$ suffices. Indeed, following the argument of Sec. V D, knowing only the two real parameters α_n and ζ_n suffices.

Now that we have identified μ_{n+1}, we repeat the procedure, starting at time t_{n+1}, to find the measurement settings for the next step, μ_{n+2}, using $\underline{A}_{n+1}$. Repeating in this fashion till $t_N = T$ (where we always measure), equips us with a single realization of Y_N that corresponds to a specific sequence of settings, $\{\mu_1, \ldots, \mu_N\}$ and a trajectory of the reward function, $\mathcal{C}^c_{|Y_n}(t_n)$ for $n \in \{1, \ldots, N\}$. Then, to calculate the expected reward function, similar to what we defined in Eq. (58), we average over the information Y_n as

$$\mathcal{C}^c(t_n) = \left\langle \mathcal{C}^c_{|Y_n}(t_n) \right\rangle_{Y_n} = \sum_{Y_n} \wp\,(Y_n)\, \mathcal{C}^c_{|Y_n}(t_n). \tag{98}$$

The hope, then, is that by locally maximizing $\mathcal{C}^c$ in time, that the final $\mathcal{C}^c(T)$ will be much larger than the unconditioned coherence, $\mathcal{C}^{\mathrm{nc}}$, and perhaps even close to optimal.

In the following subsection, we will implement the "full-Greedy" algorithm as presented here and evaluate Eq. (98) numerically. However, instead of numerically generating all possible trajectories associated with all possible measurement results Y_N, we will perform a stochastic simulation where the

measurement output at each step, y_n, is generated randomly according to its probability. Then, by generating enough stochastic trajectories, we can evaluate the expected reward, $\mathcal{C}^c(t_n)$ in Eq. (98) as an average. But in later subsections we will see how it can be possible to simplify the algorithm so as to perform the exact sum.

B. Numerical implementation of Greedy

For our numerical calculations, we choose $\gamma_\uparrow = \gamma_\downarrow = \bar{\gamma} = 1$ for simplicity and the other dynamical parameters (κ and K) are in the units of $\bar{\gamma}$. Equating RTP up and down jump rates yields symmetric dynamics that are easier to analyze. Also, we assume that, at $t = 0$, the RTP is in the steady state, $\underline{A}_0 = \underline{P}_{ss}$, where $\underline{P}_{ss}$ is defined in Eq. (8). We choose $\kappa = 0.2$ and a range of K values from 2 to 100. At the higher end of ranges for K, condition (2) will be satisfied. For the (theoretically infinitesimal) time step dt, it is most critical that it is small compared to $1/K$. We choose $dt = 0.001/K$.

The Greedy algorithm requires optimization over the three measurement angles as in Eqs. (96). Fortunately, our locally optimal strategy has the same structure as that introduced in Ref. [50]. This means we can use analytical solutions described in [51] to reduce the range of each possible angle from a real interval $[-\pi/2, \pi/2]$ to only three discrete choices. This is described in Appendix C.

As we discussed above, after finding an optimal μ_{n+1}, we need to simulate the measurement outcome y_{n+1} in order to calculate $\underline{A}_{n+1}$. For our stochastic simulation we randomly generate y_{n+1} using a probability function $\wp(y_{n+1}|Y_n)$, where Y_n is the measurement record assumed known up to that time. This function is encoded in vector $\underline{\check{A}}_{n+1}$, following Eq. (76), where we write

$$\wp(y_{n+1}|Y_n) = \frac{\wp(Y_{n+1})}{\wp(Y_n)} = \frac{\underline{I}^\top \underline{\check{A}}_{n+1}}{\underline{I}^\top \underline{\check{A}}_n}. \quad (99)$$

To be able to use the above equation, we update the vector $\underline{\check{A}}_n$ after each measurement using Eq. (78), with the initial probability vector $\underline{\check{A}}_0 = \underline{P}_{ss}$, same as $\underline{A}_0$ as in Eq. (80).

We show the conditional coherence $\mathcal{C}^c_{|Y_n}(t_n)$ in Fig. 2 from numerical implementation of the Greedy algorithm using aforementioned parameters and $K = 20$. The light green lines are 100 trajectories of $\mathcal{C}^c_{|Y_n}(t_n)$ where three of them are exaggerated (bold-dashed lines) as examples. To calculate the expected reward function $\mathcal{C}^c(t_n)$, we average over these 100 trajectories. It should be noted that the waiting times, τ_n, in different trajectories do not match, so we could not simply calculate the average decoherence at each t_n. Instead, for each trajectory we interpolated a line between consecutive $(t_n, \mathcal{C}^c_{|Y_n}(t_n))$ points, giving rise to the broken lines shown in Fig. 2. Then we took the average of the broken lines, which is shown in dark green. Comparing the average with the no-control coherence $\mathcal{C}^{nc}(t)$, we see that the Greedy algorithm can massively improve the data qubit's coherence.

C. Greedy choices of measurement settings

We would like to understand how the greedy algorithm chooses the measurement settings $\mu_{n+1} = \{\theta_{n+1}, \tau_{n+1}\}$, based on the measurement result y_n and relevant statistics α_n and
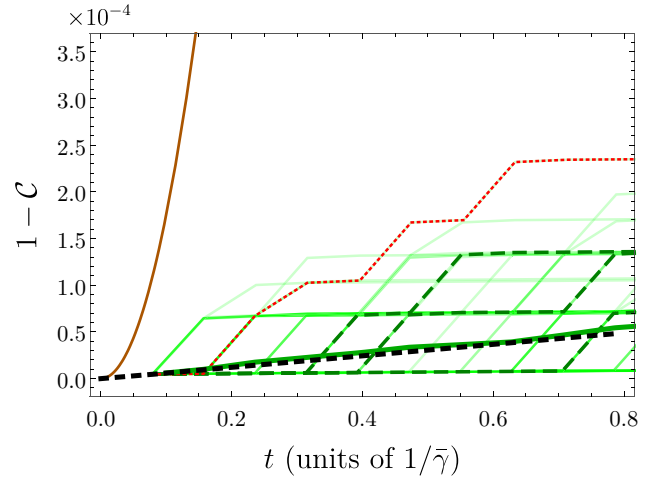

FIG. 2. Decoherence in data qubit for $\kappa = 0.2$ and $K = 20$. The thick brown curve represents no-control decoherence ($1 - \mathcal{C}^{nc}$) in the asymptotic regime as in Eq. (36), whereas the green lines represent 100 trajectories from numerical simulations of the Greedy algorithm. We plotted the trajectories in light green so that the saturation in the color shows the relative probability of a particular decoherence value at a particular time. Note that Greedy selects a different waiting time at each step, but the difference between them in this case is of order $10^{-3}/\bar{\gamma}$, which is too small to see here, giving the appearance that all branches occur at the same time. Three typical trajectories are bolded with thick dashed lines, and one atypical trajectory, which happened only once, is shown in dotted red. The thick dark green line is the average of the 100 trajectories, while the black dashed line represents the exact mean of Greedy$_4$, introduced in Sec. VI D.

ζ_n. To do that, we look at individual trajectories of the Greedy chosen settings. A typical one is presented in Fig. 3. For ζ_n and θ_{n+1}, their values can change sign, so we plotted the absolute values of them in order to see the pattern better, while the sign of them are depicted by filled and hollow markers for positive and negative respectively.

We observe in Fig. 3 that, after an initial transient, the statistical parameters, α_n and $|\zeta_n|$, jump between only two values each, and these jumps are correlated with the measurement result y_n. We also observe that whenever a nonnull measurement ($y_n = 1$) occurs, the sign of ζ_n changes, that is depicted by the filled markers switched to hollow markers. Moreover, from Fig. 3(d), we see that the Greedy choices for the waiting time τ_{n+1} and the absolute value of measurement angle $|\theta_{n+1}|$ are, apart from transients, again confined to two values and their jumps are correlated with jumps in $\{\alpha_n, \zeta_n\}$. Finally, the sign of θ_{n+1} is the same as the sign of ζ_n. These observations lead us to proposing a new approach to implement a simpler version of Greedy algorithm that we explain next.

D. "Greedy$_4$" with four parameters

We can capture the Greedy choices of measurement parameters as just explored by setting

$$\theta_{n+1} := s_n \Theta_G(\alpha_n, |\zeta_n|), \quad (100a)$$

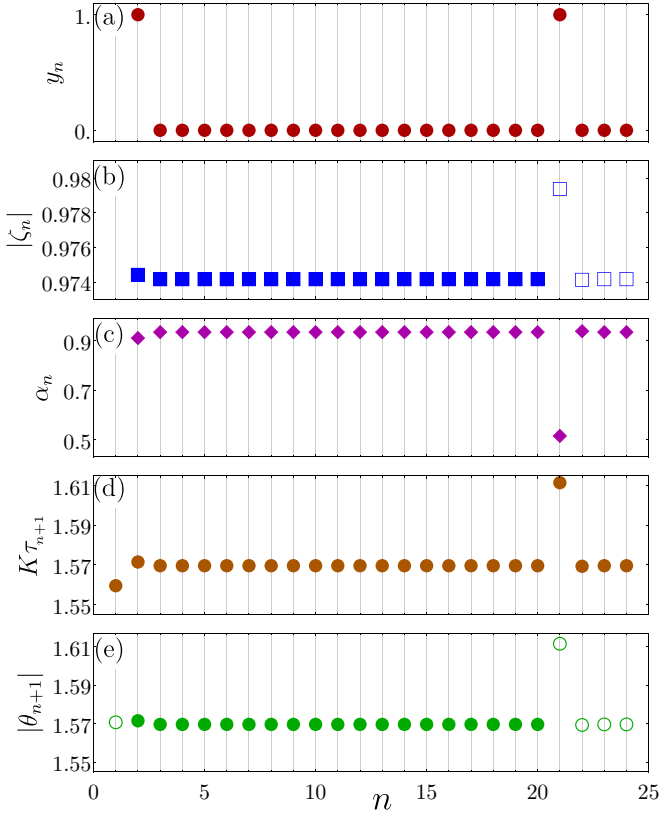$$\tau_{n+1} := \frac{\Delta_G(\alpha_n, |\zeta_n|)}{K}. \quad (100b)$$

FIG. 3. Greedy chosen values of next measurement parameters, $\{\theta_{n+1}, \tau_{n+1}\}$, together with the information gained via nth measurement, ζ_n, α_n, and y_n, in a single trajectory of the Greedy algorithm for $\kappa = 0.2$ and $K = 20$. For θ_{n+1} and ζ_n, their absolute values are shown in the vertical axes, while their signs are depicted via filled and hollow markers for positive and negative signs, respectively.

Here the function Θ_{G} sets the absolute value of the Greedy choice of measurement angle, while

$$s_n := \mathrm{sgn}(\zeta_n), \tag{101}$$

determines its sign, and the function Δ_{G} sets the waiting time τ_{n+1}. This Δ_{G} is dimensionless and of order unity, since, as we have discussed, the waiting time scales with $1/K$.

Although, in general, Δ_{G} and Θ_{G} may be functions of both α_n and ζ_n, we see from Fig. 3 that knowing α_n is sufficient. In fact, since (after some transient steps) α_n jumps between two values, denoted by $\{\alpha^>, \alpha^<\}$, it is sufficient to know the binary variable $a_n \in \{>, <\}$, defined as

$$a_n = \begin{cases} > & \text{if } \alpha_n > \bar{\alpha}, \\ < & \text{if } \alpha_n < \bar{\alpha}, \end{cases} \tag{102}$$

where $\bar{\alpha} := (\alpha^> + \alpha^<)/2$. Note that, most of the time, $a_n = >$. Thus, we simplify Θ_{G} and Δ_{G} as binary functions, each with two values as

$$\Theta_{\mathrm{G}}(a_n) = \begin{cases} \Theta_{\mathrm{G}}^> & \text{if } a_n = >, \\ \Theta_{\mathrm{G}}^< & \text{if } a_n = <, \end{cases} \tag{103a}$$

and

$$\Delta_{\mathrm{G}}(a_n) = \begin{cases} \Delta_{\mathrm{G}}^> & \text{if } a_n = >, \\ \Delta_{\mathrm{G}}^< & \text{if } a_n = < . \end{cases} \tag{103b}$$
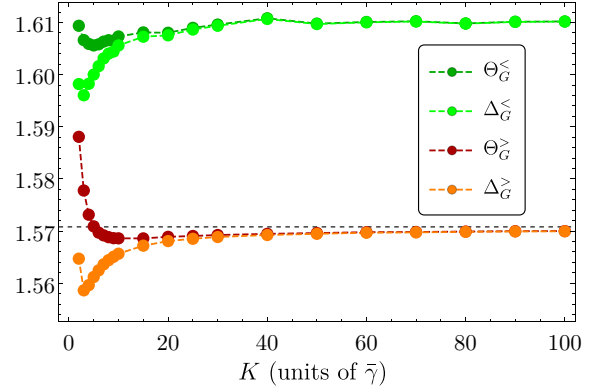
FIG. 4. Calculated values of the Greedy$_4$ parameters, from top to bottom, $\Theta_{\mathrm{G}}^<$, $\Delta_{\mathrm{G}}^<$, $\Theta_{\mathrm{G}}^>$, and $\Delta_{\mathrm{G}}^>$, for a range of K. See text for details of calculation. The horizontal dashed line shows the value of $\pi/2$; see Eq. (112).

It should be noted that the above equations could have been defined in terms of $|\zeta_n|$ rather than α_n. However, because the difference between $\alpha^>$ and $\alpha^<$ is more pronounced, it is more natural to use this as the variable on which Θ_{G} and Δ_{G} depend.

We calculate numerical values of $\{\Delta_{\mathrm{G}}^>, \Delta_{\mathrm{G}}^<, \Theta_{\mathrm{G}}^>, \Theta_{\mathrm{G}}^<\}$ as follows. Each Greedy trajectory provides us with a sequence of Greedy chosen measurement angles $\{\theta_n\}$, a sequence of waiting times $\{\tau_n\}$, and an associated sequence of $\{\alpha_n\}$. First, we discard the transient data (the first few time steps) and then separate the chosen angles into two groups by checking if the associated α_n is larger or smaller than the threshold $\bar{\alpha}$. That is, we have two sets of angles defined as

$$\vartheta^> = \{|\theta_n| : a_n = >\}, \tag{104a}$$

$$\vartheta^< = \{|\theta_n| : a_n = <\}. \tag{104b}$$

Similarly for the waiting times, scaled by K, we define two sets as

$$\Upsilon^> = \{\tau_n/K : a_n = >\}, \tag{105a}$$

$$\Upsilon^< = \{\tau_n/K : a_n = <\}. \tag{105b}$$

The above procedure is repeated for all 100 trajectories. Then we take all members of $\vartheta^>$s and average them to be the numerical value of $\Theta_{\mathrm{G}}^>$. The numerical values of $\Theta_{\mathrm{G}}^<$, $\Delta_{\mathrm{G}}^>$, and $\Delta_{\mathrm{G}}^<$ are calculated similarly, following the conditions of a_n in Eqs. (103). Figure 4 shows the values of $\Delta_{\mathrm{G}}^>$, $\Delta_{\mathrm{G}}^<$, $\Theta_{\mathrm{G}}^>$, $\Theta_{\mathrm{G}}^<$ for a range of the sensitivity K from 2 to 100. All four of these parameters are within 2.5% of $\pi/2$ across the whole range.

Now that we know the binary values of Θ_{G} and Δ_{G}, we no longer need the full optimization part of the Greedy algorithm. As a result, we obtain a new adaptive algorithm in which an experimenter can simply use (103) with the knowledge of α_n to decide measurement settings for the next step. We call this new adaptive algorithm "Greedy$_4$" (Greedy with subindex 4), since it is defined by four parameters, namely $\Delta_{\mathrm{G}}^>$, $\Delta_{\mathrm{G}}^<$, $\Theta_{\mathrm{G}}^>$, and $\Theta_{\mathrm{G}}^<$. We emphasize again that in order to determine these four parameters, the full Greedy algorithm, including the optimization part, must be run. Once the four parameters are found, the experimenter could use them for future experiments.
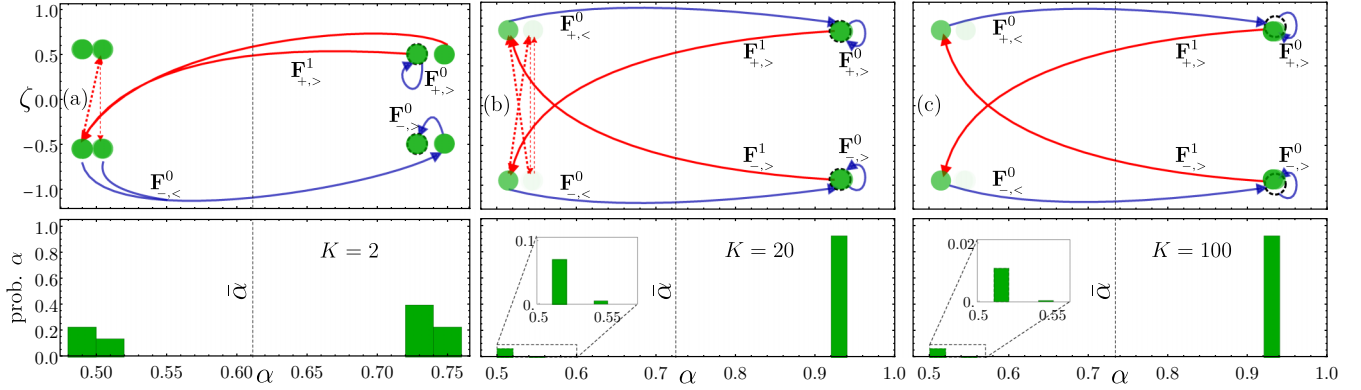
FIG. 5. Dynamics of the Greedy$_4$ algorithm in phase space (top row) and probability of α (bottom row). We used $\kappa = 0.2$ in all plots, while $K = 2, 20,$ and 100, from left to right. The value of $\bar{\alpha}$ is indicated by a dashed line in each plot. The blue arrows represent maps due to null results ($y = 0$), while the red arrows represent nonnull outcomes ($y = 1$). Less frequent transitions are shown by red dashed arrows. We have shown only half of the arrows in the instance of $K = 2$; the reader can visualize the other half by mirroring the image around the $\zeta = 0$ axis.

To calculate the expected reward $\mathcal{C}^c(t_n)$ as in Eq. (98) for Greedy$_4$, we numerically generate all possible Y_N and average the associated reward functions with the probability weight $\wp(Y_N)$ from Eq. (76). Each realization of Y_N, like full-Greedy, is associated with a different sequence of waiting times τ_n, which take two values. However, as shown in Fig. 4, the difference between $\Delta_G^>$ and $\Delta_G^<$ is not great. Moreover, as we have already seen in Fig. 3, most of the time the value $\Delta_G^>$ is chosen. The other value is chosen only when the system detects that a jump has occurred, which happens at approximately the same rate as the jumps, $\check{\gamma}$. Thus it is not a bad approximation to take the average as if every τ_n were the same, an effective waiting time is given by

$$\tau_{\text{eff}} = \frac{\Delta^>}{K}(1 - \check{\gamma}\,\tau_{\text{eff}}) + \frac{\Delta^<}{K}\check{\gamma}\,\tau_{\text{eff}}. \tag{106}$$

Here τ_{eff} appears on the right-hand side because $\check{\gamma}\tau_{\text{eff}}$ is the probability, in one waiting time, that the RTP jumps. Solving for τ_{eff} we obtain

$$\tau_{\text{eff}} = \frac{\Delta^>}{K - \check{\gamma}(\Delta^< - \Delta^>)}. \tag{107}$$

Using this τ_{eff}, we find the expected reward function for Greedy$_4$, which is shown with a black dashed line in Fig. 2. This line is an *exact* average in the sense that it is not stochastic. Moreover, this line agrees with the stochastic average using the full Greedy, within the stochastic error, and confirms that the expected reward function calculated from the full Greedy is in agreement with Greedy$_4$.

E. Dynamics of Greedy$_4$ in (α, ζ) phase space

In this subsection, we use the sufficient statistics (α, ζ) as a way to visualize the dynamics of the Greedy$_4$ algorithm in "phase space." The plots in Fig. 5 show phase space dynamics of 100 trajectories for $K = 2, 20,$ and 100. We use light green dots to represent all (α, ζ) states, and thus the green color saturation indicates the relative probability of being in that state. Blue arrows in the figure show the state transition due to the null result, $y = 0$, while the red arrows show that of the nonnull result, i.e., $y = 1$. The maps used for the tran-

sitions are shown as labels. Since Greedy$_4$ chooses between the binary values of the measurement angle and the waiting time, we can simplify the notation for the **F** map introduced in Sec. V B as

$$\mathbf{F}_{s,a}^y := \mathbf{F}\left(s\Theta_G(a), \Delta_G(a)/K, y\right), \tag{108}$$

where $\Theta_G(a)$ and $\Delta_G(a)$ are the functions defined in Eqs. (103). Note that, as a consequence of taking $\gamma_\uparrow = \gamma_\downarrow$, the phase diagrams for (α, ζ) are symmetric in reflection around the $\zeta = 0$ line. Furthermore, the bottom panel of each subfigure depicts the normalized histograms for different values of α, which can be interpreted as the probability of the system being on that specific α.

In these phase diagrams, we see that the statistical states (α, ζ) have a tendency to be in only a few states, instead of spreading over the entire phase space. This was already noted in the discussion of Fig. 3 for $K = 20$, but here we visualize the correlated changes in α, ζ more easily, as well as the transition from K of order 1 to $K \gg 1$. In all cases, the two most occupied states turn out to correspond to the eigenstates with the largest absolute eigenvalues of $\mathbf{F}_{+,>}^0$ and $\mathbf{F}_{-,>}^0$ (Recall that these are 2×2 complex linear maps acting on $\underline{A}$, and it turns out that they each have two eigenstates). We denote these stable eigenstates (the ones with largest absolute eigenvalues) by $\underline{E}_{+,>}^0$ and $\underline{E}_{-,>}^0$, respectively. The (α, ζ) values for these eigenstates are depicted as the dashed circles in the phase space plot.

To understand the state dynamics, let us first consider a state being in $\underline{E}_{+,>}^0$, the stable eigenstate of $\mathbf{F}_{+,>}^0$. In this case, any null-result ($y = 0$) measurement will map the state to itself and no effect on the system's phase space configuration. On the other hand, a nonnull measurement ($y = 1$) will map the state to a point on the left side, with $\alpha < \bar{\alpha}$ while switching the sign of ζ from positive to negative. Any subsequent nonnull result, which is less common, will cause the sign of ζ to switch again, while α remains in the same region $\alpha < \bar{\alpha}$. However, since the probability of two consecutive nonnull results is low, the subsequent measurement most likely yields a null result, mapping the system state to a point on the right side with $\alpha > \bar{\alpha}$, while keeping the sign of ζ unchanged. The

following null results, as they are more likely, will move the state towards the other stable eigenstate, $\underline{E}^0_{-,>}$.

The above description holds for any value of $K > \bar{\gamma}$. But there are differences as K changes. First consider the case for a relatively low sensitivity, $K = 2$, i.e., of the same order as $\bar{\gamma} = 1$. In this case we observe in Fig. 5(a) that the dynamics in the phase space can be grouped into eight locations. We understand the dynamics as follows. For small K, we find that the likelihood of consecutive nonnull measurement results is higher compared to that of the larger K. So, if the system configuration has moved to the left side of the phase space, as a consequence of a measurement result, any further nonnull measurements yield maps that changes the sign of ζ but keeps the α on the left side. These maps as we showed them with red dashed arrows, reveals four main locations on the left side of the phase space. Furthermore, we observe that a map due to a null result, ($y = 0$), brings the system from the left to the right side of the phase space, but the resulting configuration is distanced from the eigenstates. Then, further null results bring the system to the eigenstate. As a result, there are four main locations on the right side. Note that we have not shown all the transitions in the $K = 2$ case to avoid messiness (see caption).

As K increases, the first thing that happens is that the rightmost points on the phase diagram all converge to the eigenstates $\underline{E}^0_{\pm,>}$. That is, for $\alpha > \bar{\alpha}$ we have only these two eigenstates to consider. This is seen already in Fig. 5(b). As K increases even further, reaching the asymptotic regime $K \gg \bar{\gamma}$, the chance of two consecutive nonnull results is negligible. This leads to there also being only two points on the left side, as seen in Fig. 5(c). Thus, the dynamics are simplified and can be grasped by concentrating on just four points: $\underline{E}^0_{\pm,>}$, and $\mathbf{F}^1_{\pm,>} \underline{E}^0_{\pm,>}$. Moreover, the system spends the majority of its time in one of the eigenstates and occasionally, with a nonnull ($y = 1$) measurement result, goes to the $a = <$ state while altering $s = \text{sgn}(\zeta)$. The next measurement almost always yields $y = 0$, causing a jump back to an eigenstate while keeping s the same.

F. Greedy$_2$ in the asymptotic regime

In the asymptotic regime $K \gg \bar{\gamma}$, we show that the Greedy$_4$ algorithm in Eqs.(103) can be simplified even further. By inspecting Fig. 4 carefully, we observe that in this regime the four parameters converge to only two, because we have the identities $\Theta^>_G = \Delta^>_G$ and $\Theta^<_G = \Delta^<_G$. From the original definitions in Eq. (100), it means that the Greedy-chosen measurement angles and waiting times are related via a simple relation:

$$\theta_{n+1} = s_n K \tau_{n+1}. \tag{109}$$

As a result, we have a simpler algorithm with just two parameters: $\Theta^>_G$ and $\Theta^<_G$. We refer to this algorithm as "Greedy$_2$," which is valid only for large K or the asymptotic regime. The algorithm can be summarized as

$$\theta_{n+1} = s_n \Theta_G(a_n), \tag{110a}$$

$$\tau_{n+1} = \Theta_G(a_n)/K, \tag{110b}$$

where $\Theta_G(a_n)$ is given by Eq. (103a).

This choice has an intuitive explanation, as follows. The term $K\tau_{n+1}$ is simply a SQ's phase evolved (after the SQ's phase reset to zero) during the time interval (t_n, t_{n+1}), where $\tau_{n+1} = t_{n+1} - t_n$, given that the RTP stays unchanged with a value of s_n during that time. That is, under the approximation that $z(t)$ stays equal to our best estimate of its value, s_n at the start of the interval, the evolved SQ's phase is given by

$$\Phi(t_{n+1}) = K \int_{t_n}^{t_{n+1}} z(t)dt \approx s_n K \tau_{n+1}. \tag{111}$$

This approximation can be justified in the asymptotic regime, as our confidence in the value of z_n is high ($|\zeta_n| \approx 1$) and the probability of a jump in the interval is low, $O(\check{\gamma}/K)$. Therefore, Eq. (109) simply says that the measurement angle should be chosen to be exactly the most likely SQ phase at the time of measurement, $\theta_{n+1} = \Phi(t_{n+1})$. In other words, the measurement probes whether the SQ is in the most likely state. A null result ($y_{n+1} = 0$) is the answer "yes," while the other result ($y_{n+1} = 1$) is the answer "no." That is, in this aspect, Greedy$_2$ performs an optimal choice for minimum-error binary state discrimination [52,53], where one state (which is pure) arises if and only if $z(t) = s_n$ for the entire interval, and the other state (which is mixed) arises if and only if $z(t) = -s_n$ for some of the interval.

Note that the above simple intuition does not explain the value chosen by Greedy$_2$ for its two parameters, $\Theta^>_G$ and $\Theta^<_G$. However, inspecting Fig. 4 again, we see that one of these values does have a simple interpretation in the asymptotic limit:

$$\lim_{K/\bar{\gamma} \to \infty} \Theta^>_G = \pi/2. \tag{112}$$

Note that this is the more important of the two parameters, as in the asymptotic regime, the measurement result is almost always a null result, the system state is almost always in the region $\alpha_n > \bar{\alpha}$, and the measurement choice is almost always $\Theta_G(a_n) = \Theta^>_G$. The choice of $\pi/2$ for $s_n\theta_{n+1}$ and $K\tau_{n+1}$ implies that the most likely pure state [from $z(t) = s_n$ for the entire interval] and the secondmost likely pure state [from $z(t) = -s_n$ for the entire interval] are orthogonal at the measurement time, and can be perfectly distinguished by the measurement.

VII. MAP-BASED OPTIMIZED ADAPTIVE ALGORITHM FOR ASYMPTOTIC REGIME (MOAAAR)

While using the Greedy algorithm for the SQ offers a great reduction in the decoherence of the data qubit, it is only locally optimal. There is no reason to expect it to be globally optimal. In this section we introduce and analytically study a new algorithm that is plausibly optimal in the asymptotic regime, $K \gg \gamma_{\uparrow,\downarrow} \gg \kappa$. However, to design the algorithm we use the knowledge and intuitions gained from the numerical investigation of Greedy in this regime.

First, recall from Sec. VI F that in asymptotic regime, Greedy chooses measurement settings that are perfectly matched according to Eqs. (110). Second, recall from Sec. VI E that out of the two values of $\Theta_G(a_n)$, the choice is almost always $\Theta_G(>)$, which we called $\Theta^>_G$. Considering just the single parameter, $\Theta^>_G$, one could use it to define

what one might call a Greedylike$_1$ algorithm, in which $\Theta_{\rm G}(a)$ in Greedy$_2$ is replaced by the constant $\Theta_{\rm G}^\gtrless$. This would reproduce the performance of the Greedy algorithm in the asymptotic regime, although it would not actually approximate the Greedy algorithm itself, which always uses two distinctly different angles, even in the asymptotic regime, as per Greedy$_2$.

Like the Greedy algorithm, this Greedylike$_1$ algorithm would still be adaptive, in general, because of the s_n in Eq. (110a). However, we already know the value of $\Theta_{\rm G}^\gtrless$ in the asymptotic regime, Eq. (112). While this choice is intuitive, it in fact corresponds to a *nonadaptive* measurement, as $\theta = s\pi/2$ is measuring in the same basis for $s = \pm 1$. It would be surprising if the globally optimal performance in the asymptotic regime could be attained by a nonadaptive algorithm, and indeed we find that this is not the case. We do this by cleaving to the well-motivated relations between the measurement angles and the waiting times in Eqs. (110), and to the understanding that the asymptotic performance can be expected to be reproduced with Θ fixed, but without cleaving to the particular choice of Θ made by the Greedy algorithm. That is, as expected, we can do better by choosing Θ in a globally optimal manner rather than a locally optimal (greedy) manner.

To be precise, we here propose an adaptive (in general) algorithm for measurement angles and waiting times,

$$\theta_{n+1} = s_n\Theta, \tag{113a}$$

$$\tau_{n+1} = \Theta/K, \tag{113b}$$

as in Eqs. (110), but with the function $\Theta_{\rm G}$ replaced by a fixed parameter Θ. This simplifies the calculation of the expected reward function in the asymptotic regime, as we need consider only the following set of four single-parameter maps:

$$\mathbf{F}_s^y(\Theta) := \mathbf{F}(s\Theta, \Theta/K, y), \tag{114}$$

where $s \in \{+1, -1\}$ and $y \in \{0, 1\}$. Similar to the preceding section, we denote the stable eigenstates of these maps by $\underline{E}_s^y$.

In this section, we first show how the problem of maximizing the expected reward function is, in the regime of interest (2), equivalent to minimizing an expected decoherence rate. We then apply the adaptive strategy, Eqs. (113), and derive an analytical expression for the decoherence rate as a function of Θ in the asymptotic regime. This can be used to search for a *globally* optimal parameter value, $\Theta^\star$, which defines our Map-based Optimized Adaptive Algorithm for Asymptotic Regime (MOAAAR). We call it Map-based because of the central role of the Bayesian maps, as mentioned in the preceding paragraph. We emphasize that it is adaptive because we will find $\Theta^\star \neq \pi/2$, which means that, via s_n, previous measurement results do alter the basis in which the SQ is measured. (Recall that this is in contrast with the $\Theta_{\rm G}^\gtrless = \pi/2$ choice of the Greedylike$_1$ algorithm in this regime.)

A. Expected reward and decoherence rate

Recall the original expected reward function Eq. (55) defined for the final time T, when the phase-correction control is to be applied. This is a sum over exponentially many terms, which makes it infeasible to do an analytic maximiza-

tion. However, from the numerical results in the previous section (Greedy), we know that the system's state (α, ζ) effectively jumps between only a finite number of states, and explores all of them over time. As we will see, this is also true for the globally optimized MOAAAR strategy. This means that the stochastic evolution of our state of knowledge is ergodic [37], so the average decay of the coherence over a typical trajectory over many steps $(T \gg K^{-1})$ is the same as an ensemble average over one step. There will be deviations from this due to initial $[t = O(1/\check\gamma)]$ and final $[T - t = O(1/\check\gamma)]$ transients, but these can be ignored for long times as we are interested in. Thus we can then safely use an expected reward function at an intermediate time t_n, $\mathcal{C}^{\rm c}(t_n)$ as a proxy of the expected reward at the actual final time, $\mathcal{C}^{\rm c}(T)$.

We also saw in the preceding section that in the regime $K \gg \bar\gamma$, the Greedy algorithm makes the system spend almost all of the time close to the two stable eigenstates $\underline{E}_\pm^{y=0}$. We find this to be even more true with MOAAAR; we will quantify deviations from this assumption in Sec. VII C. Thus we can approximate the coherence (the reward function) at time t_n as

$$\mathcal{C}^{\rm c}(t_n) = \sum_{Y_n} \wp(Y_n)\big|\mathcal{A}_{|Y_n}\big| = \sum_{Y_n} |\underline{I}^\top \underline{A}_n|$$

$$= \sum_{\substack{Y_n| \\ s_n=+1}} |\beta_+(Y_n)\underline{I}^\top \underline{E}_+^0| + \sum_{\substack{Y_n| \\ s_n=-1}} |\beta_-(Y_n)\underline{I}^\top \underline{E}_-^0|. \tag{115}$$

The first line follows from the definitions of the reward function in Eq. (98) or Eq. (55). In the second line, we have separated the summation, $\sum_{Y_n} |\underline{I}^\top \underline{A}_n|$, into two contributions associated with $s_n = +1$ and $s_n = -1$, and used the fact that the state $\underline{A}_n$ is almost always proportional to one of the two eigenstates ($\underline{E}_+^0$ and $\underline{E}_-^0$). Which of the two states pertains at time t_n is determined solely by s_n (which is a function of Y_n), while $\beta_+(Y_n)$ and $\beta_-(Y_n)$ are the appropriate proportionality scalars.

Let us now consider the expected reward at the next SQ's measurement at $t_{n+1} = t_n + \tau$, which can be computed by applying one additional map, $\mathbf{F}_\pm^y(\Theta)$, to Eq. (115) and summing over its measurement results, $y_{n+1} \in \{0, 1\}$, for which we use the dummy variable y. This gives

$$\mathcal{C}^{\rm c}(t_{n+1}) = \sum_y \sum_{\substack{Y_n| \\ s_n=+1}} |\beta_+(Y_n)\underline{I}^\top \mathbf{F}_+^y(\Theta)\underline{E}_+^0|$$

$$+ \sum_y \sum_{\substack{Y_n| \\ s_n=-1}} |\beta_-(Y_n)\underline{I}^\top \mathbf{F}_-^y(\Theta)\underline{E}_-^0|. \tag{116}$$

Note that the signs s_n determine which $\mathbf{F}$ map is applied, and matches the subscript of the eigenstates.

Next, we rearrange the formula Eq. (116) such that it is written in terms of the coherence at time t_n in Eq. (115), with one (justified) approximation. We can interpret the two terms in Eq. (115) as two weighted conditional coherence contributions, namely, $P_{\rm ss}(s_n = +1)\mathcal{C}^{\rm c}_{|s_n=+1}(t_n)$ and $P_{\rm ss}(s_n = -1)\mathcal{C}^{\rm c}_{|s_n=-1}(t_n)$, respectively. [Recall that we defined a general conditioned coherence in Eq. (57).] Since s_n is our best guess for the RTP z_n, and since we are not in a transient regime, we can take these probabilities to be the same as the steady-state

probabilities for the RTP. Also, in this long-time steady-state regime, we can safely say that the value that s_n happens to have at this particular time has almost no bearing on the coherence at this time, which is determined by a long history of stochastic events, only the most recent of which is relevant to s_n. That is, we can make the approximation $\mathcal{C}^c_{|s_n=+1}(t_n) \approx \mathcal{C}^c_{|s_n=-1}(t_n) \approx \mathcal{C}^c(t_n)$. Rewriting the terms involving $\beta_{\pm}(Y_n)$ in Eq. (115) in terms of conditional coherences, we can simplify the expected reward at time t_{n+1} to

$$
\begin{aligned}
\mathcal{C}^c(t_{n+1}) = \sum_y \Bigg\{ & \frac{P_{ss}(s_n=+1)|\underline{I}^\top \mathbf{F}^y_+(\Theta)\underline{E}^0_+|}{|\underline{I}^\top \underline{E}^0_+|} \mathcal{C}^c_{|s_n=+1}(t_n) \\
& + \frac{P_{ss}(s_n=-1)|\underline{I}^\top \mathbf{F}^y_-(\Theta)\underline{E}^0_-|}{|\underline{I}^\top \underline{E}^0_-|} \mathcal{C}^c_{|s_n=-1}(t_n) \Bigg\}
\end{aligned}
$$

$$
\approx [1 - \bar{\Gamma}(\Theta)\tau]\,\mathcal{C}^c(t_n). \tag{117}
$$

This describes an exponential decay with an expected decoherence rate defined as

$$
\bar{\Gamma}(\Theta) := \sum_{s=\pm} P_{ss}(s) \frac{\left|\underline{I}^\top \underline{E}^0_s\right| - \sum_y \left|\underline{I}^\top \mathbf{F}^y_s(\Theta)\underline{E}^0_s\right|}{\tau \left|\underline{I}^\top \underline{E}^0_s\right|}. \tag{118}
$$

Here we have dropped the n subscript entirely as it is no longer necessary.

The $\bar{\Gamma}(\Theta)\tau$ in Eq. (117) can also be thought of as a steady-state average of the relative change of the conditional coherence, as the system's state evolves from $\underline{A}_n$ (conditioned on Y_n) at time t_n to the next time $t_{n+1} = t_n + \tau$ with an unknown y_{n+1}. This relative change was called $\delta(\underline{A}_n)$ in CL [28] (where the vector $\underline{A}_n$ is its argument, not the scalar quantity which is changing). In terms of this, we have a conditional rate,

$$
\begin{aligned}
\Gamma(\Theta, \underline{A}_n) &= \frac{\delta(\underline{A}_n)}{\tau}, \\
&= \frac{1}{\tau} \frac{\left|\mathcal{A}_{|Y_n}\right| - \sum_{y_{n+1}} \wp(y_{n+1}|Y_n)\left|\mathcal{A}_{|Y_{n+1}}\right|}{\left|\mathcal{A}_{|Y_n}\right|}, \tag{119}
\end{aligned}
$$

and the average rate in Eq. (118) is obtained by averaging over Y_n. This average exponential-decay rate well approximates the coherence decay rate for a typical trajectory, as argued above. Therefore we can use $\bar{\Gamma}(\Theta)$ as a proxy for the global reward function (55). That is, we reformulate the problem of maximizing the expected final coherence as the problem of minimizing the average decoherence rate. This can be done essentially analytically, as we see in the following subsection.

B. Analytical expression for decoherence

We wish to find the minimum of the function $\bar{\Gamma}(\Theta)$ in Eq. (118). Recall that Θ is the single parameter that defines our adaptive algorithm as per Eqs. (113). This in term defines the four maps in Eq. (114) (from two choices of $s \in \{+1, -1\}$ and two choices of $y \in \{0, 1\}$), as a function of Θ, via the original definition in Eq. (66). We are interested in the stable eigenstates $\underline{E}^0_\pm$ of the map $\mathbf{F}^0_\pm(\Theta)$. With the help of *Mathematica*, we analytically solve for these eigenvectors and their expansions in the asymptotic regime, taking $\gamma_\uparrow, \gamma_\downarrow \ll K$. Ignoring an irrelevant complex scalar multiple, these eigenvectors can be represented by the two sufficient statistics introduced in Sec. V D. These are the parameters α^0_s

and ζ^0_s found by substituting $\underline{A}_n = \underline{E}^0_s$ into $\alpha_n(\underline{A}_n)$ (83a) and $\zeta_n(\underline{A}_n)$ (85), respectively. Keeping terms up to $O(1/K)$, we get

$$
\alpha^0_s = \frac{1}{12(2\Theta + \sin 2\Theta)^2}\left(N_\Theta + s\frac{\gamma_\downarrow - \gamma_\uparrow}{K}M_\Theta\right), \tag{120a}
$$

$$
\zeta^0_s = s \times \left\{1 - \frac{\gamma_s}{K}[\csc\Theta(\cos\Theta + \Theta\csc\Theta)]\right\}, \tag{120b}
$$

where we have used γ_s with the meaning $\gamma_+ = \gamma_\downarrow$, $\gamma_- = \gamma_\uparrow$. The terms N_Θ and M_Θ are lengthy functions of Θ alone, and are shown in Appendix. D.

By substituting α^0_s and ζ^0_s in Eqs. (120) for the eigenstates $\underline{E}^0_s$ in Eq. (118) and expanding terms to the lowest order of (κ/K), we obtain, to lowest order,

$$
\bar{\Gamma}(\Theta) \approx H_\Theta \breve{\gamma}\kappa^2/(2K^2), \tag{121}
$$

where

$$
H_\Theta = 3\Theta^2\csc^4\Theta - [2\Theta(\Theta - \cot\Theta) + 1]\csc^2\Theta + \tfrac{1}{3}\Theta^2 - 1. \tag{122}
$$

This is our most important result, also shown in the CL [28], where we plot H_Θ. We find that this function has an absolute minimum of $H^\star \approx 1.254$ at $\Theta = \Theta^\star \approx 1.50055$. This measurement angle $\Theta^\star$ defines our MOAAAR, and is plausibly the lowest possible decoherence rate using projective measurements on a SQ within the regime of Eq. (2). We show in Fig. 6 the phase space plots for $K = 2, 20, 100$ using the MOAAAR algorithm. This confirms that, even more so than for the Greedy algorithm in Fig. 5, as $K/\bar{\gamma}$ grows, the system spends more and more time in the two stable eigenstates.

C. Approximation errors and approaching rates

With the analytical expressions for MOAAAR, we can analyze any errors that could occur from the approximations we have made. First, let us revisit the approximation, $z_n \approx s_n = \text{sgn}(\zeta_n)$, used prior to Eq. (117) to justify using the steady-state probability of z_n in place of those for s_n. From the analytical expression in Eqs. (120), we can see that $\zeta^0_s = s[1 - O(\gamma_s/K)]$, for any value of Θ. Using the result (correct to even higher order) for ζ_n in Eq. (87), we can thus say that

$$
\begin{aligned}
\zeta_n &\approx \frac{\wp(Y_n, z_n=+1) - \wp(Y_n, z_n=-1)}{\wp(Y_n)} \\
&= \wp(z_n=+1|Y_n) - \wp(z_n=-1|Y_n). \tag{123}
\end{aligned}
$$

Combining the above, we have that, in the stable eigenstates,

$$
\wp(z_n \neq s_n) = O(\gamma_s/K). \tag{124}
$$

Therefore, not only do the probabilities of s_n and z_n match on average, but the variables themselves are almost always identical when the system is in the stable eigenstates. However, this raises the question of the probability that the system is in one of these eigenstates, to which we now turn.

Because the phase space variables, α and ζ, are continuous, we cannot strictly talk about the probability of the system being in a stable eigenstate. However, we can calculate the probability that it is arbitrarily close to such an eigenstate. For a given s, if null results ($y = 0$) keep occurring, the system's state moves towards the eigenstate $\underline{E}^0_s$ exponentially fast. This
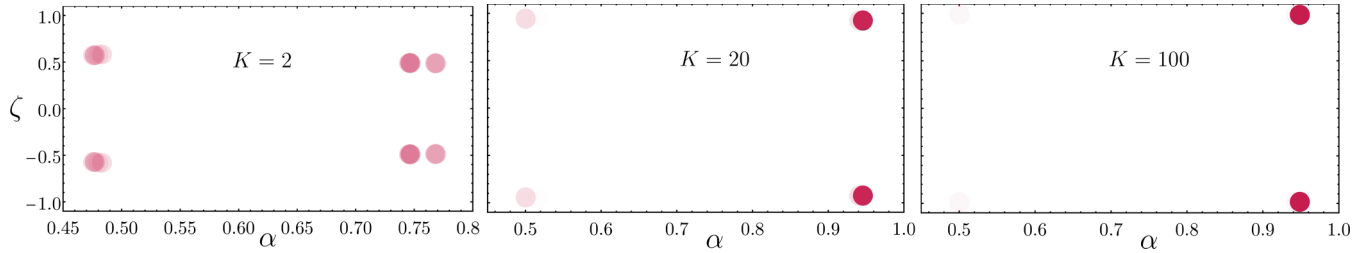
FIG. 6. The (α, ζ) phase-space plots for MOAAAR using optimal measurement angle $\Theta = \Theta^\star$ for different values of K. Similar to the Greedy case, we chose $\bar{\gamma} = 1$ and $\kappa = 0.2$. It is clear that when K is large ($K = 100$), the system's states are almost always in only two stable states (corresponding to the stable eigenstates of the two maps $\mathbf{F}^0_\pm(\Theta^\star)$.

is shown in Fig. 7 (blue dots), which takes an arbitrary value $\Theta = 1.0$ as an example, as well as the optimal value $\Theta^\star$. (We do this because we want the results in this subsection, used to justify the arguments of the preceding subsection, to hold for any Θ, not just the optimal value $\Theta^\star$ found via those arguments.) The sign of ζ stays fixed, while the value of α moves towards α^0_s exponentially fast for both $s = \pm 1$. Moreover, we see that the transient values of $|\zeta|$ are even closer to 1 than the values in the stable eigenstates. Thus, it is very safe to assume Eq. (124) for the whole evolution.

The approximation that the system state is almost always in a stable eigenstate $\underline{E}^0_\pm$ was also used in the first step, Eq. (115), towards finding the decoherence rate in Eq. (118). Thinking about the alternate expression for the decoherence rate,
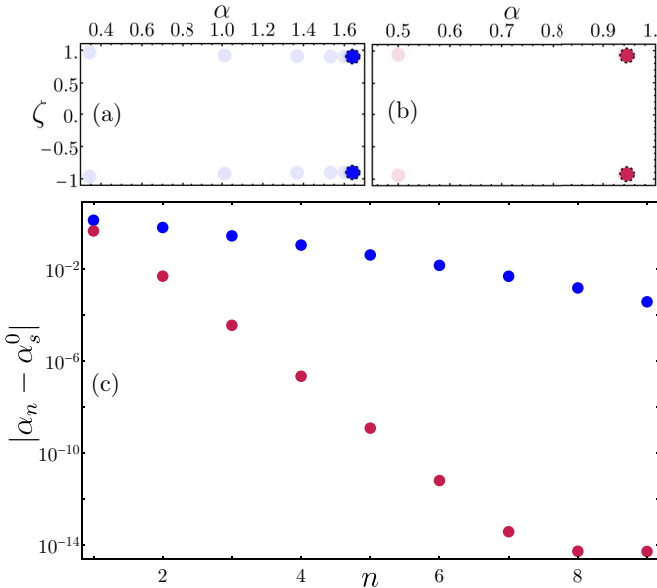

FIG. 7. Plots showing the exponential approach to the two stable eigenstates on the right side of the phase space. (a), (b) Phase-space (α, ζ) plots under the adaptive algorithm using $\Theta = 1.0$ (blue dots) and $\Theta = \Theta^\star$ (MOAAAR, maroon dots), respectively. To generate these points, we use $\mathbf{F}^1_s E^0_s$ for both $s \in \{+1, -1\}$ as a starting point and let the state evolve for the number of time steps $n = 9$. In each case, the eigenstates are shown as dashed circles. (c) A log-scale plot of distances from the eienstate's values, $|\alpha_n - \alpha^0_s|$, using the same data in (a) and (b). The last maroon data point (at $n = 9$) has reached the numerical machine precision in *Mathematica*. Here $K = 20$, and $\kappa = 0.2$ as usual.

Eq. (119), we have to ask how much difference arises, when averaging over all records, as a result of the system state's deviation from the stable eigenstates. After a nonnull result, α jumps a finite distance from α^0_s. After n time steps, the distance (in α) between the system's state and the fixed point is an exponentially decreasing function, $|\alpha_n - \alpha^0_s| \sim e^{-kn}$, where k is some dimensionless approaching constant, the slope of the curves in Fig. 7(c). We can thus approximate the number of steps after a jump that the system is still more than ϵ-distanced away from the eigenstates as $n_\epsilon = O[\log(1/\epsilon)/k]$. Now, the proportion of its evolution that the system spends this far from the eigenstates is n_ϵ times the probability per time step of a jump away from the eigenstates. But the latter is just equal to the probability of a nonnull result. This is the same order as Eq. (124), the probability that the best estimate of z_n is wrong, which is $O(\check{\gamma}/K)$. Since a jump takes a system a finite distance away in the α direction, we make the pessimistic assumption that if the system is more than ϵ-distanced away from the eigenstates then the relative error in the quantity of interest (the decoherence rate) could be of order unity. If, on the other hand, the system is less than ϵ-distanced away from the eigenstates (which is almost all the time), then if we are again pessimistic, we will say that the relative error in the quantity of interest is $O(\epsilon)$. Adding these two contributions together, we get the relative error to be $O[\log(1/\epsilon) \check{\gamma}/Kk] + O(\epsilon)$. The minimum value of this relative error is achieved when we choose ϵ to scale as $\check{\gamma}/Kk \ll 1$, which gives a relative error that is small in the asymptotic regime. Thus the errors in the approximations leading to Eq. (121) are insignificant in the regime (2).

D. Numerical comparison of performance of Greedy and MOAAAR and other algorithms

Having completed the analytical analyses of MOAAAR, we turn to numerics to compare it with the Greedy algorithm. We also compare it to some other, nonoptimized, algorithms. In Fig. 8 we plot the data qubit's decoherence as a function of time. Although the time is not long compared to $\bar{\gamma} = 1$, the decoherence for the optimized algorithms are very close to linear in time, which is a consequence of the very simple phase-space dynamics of those algorithms. All of the numerical results were calculated *exactly* using the coherence (or the expected reward) definition in Eq. (98), where all possible trajectories were generated using the $\mathbf{F}$ maps that correspond to the algorithm-chosen measurement angles and waiting times, for all possible realizations of Y_n. For the adaptive schemes
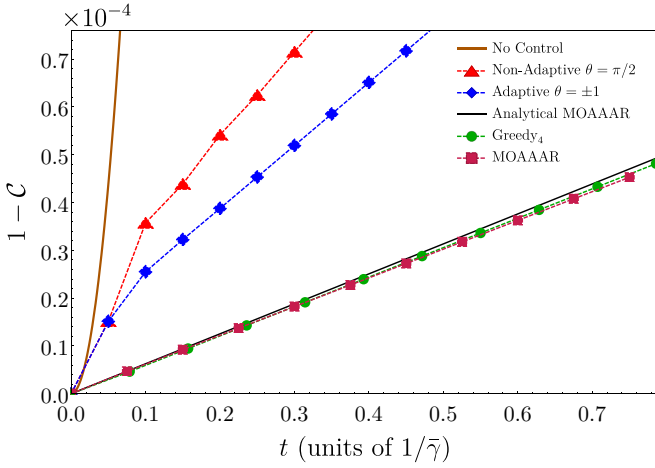
FIG. 8. Plots of data qubit's decoherence using different SQ's measurement and control strategies. Analytical results are shown with solid curves, while numerical results are shown with points connected via dashed lines. From top to bottom: the no-control case (solid brown), a nonadaptive algorithm with $\theta_{n+1} = \pi/2$ and $\tau_{n+1} = 1/K$ (red squares), an adaptive algorithm with $\theta_{n+1} = s_n$ and $\tau_{n+1} = 1/K$ (blue squares), the analytical MOAAAR from Eq. (121) (solid black), the Greedy$_4$ algorithm (green dots), and the MOAAAR with $\Theta^\star = 1.50055$ (maroon square). In all calculation, we used $\gamma_\uparrow = \gamma_\downarrow = 1$, $\kappa = 0.2$, and $K = 20$.

FIG. 9. The decoherence rates of the data qubit for the two competing strategies: the MOAAAR (maroon) and the Greedy$_4$ algorithm (green), for different values of K, setting $\gamma_\uparrow = \gamma_\downarrow = 1.0$ and $\kappa = 0.2$. Numerical results for the rates, shown as maroon disks and the green circles, are calculated from the slopes of curves in Fig. 8 (throwing away the first two points with transient effects). Here the solid maroon and dashed green curves are not simply lines connecting data points, but rather the closed-form approximation of the rates in Eq. (131) for MOAAAR and Eq. (130) for Greedy$_4$, respectively. The fainter maroon horizontal dashed line near the top shows the analytical result of the MOAAAR decoherence rate in the asymptotic regime Eq. (121) with $\Theta^\star$, while the fainter green solid line above it shows the same for Greedy, with $\Theta = \pi/2$ as in Eq. (112).

we took the first measurement angle to be $\pi/2$, since in the absence of any initial information about z_1, there is no reason to choose one sign for θ rather than the other. This is the automatic choice of the full Greedy algorithm, but not of Greedy$_4$. We also do this for later plots.

From worst to best, the curves in Fig. 8 are as follows:

(i) The no-control case, Eq. (36).

(ii) A nonadaptive algorithm with $\theta_{n+1} = \pi/2$ and $\tau_{n+1} = 1/K$. Even this completely unoptimized algorithm already does much better.

(iii) The adaptive algorithm of Eqs. (113) with the choice $\Theta = 1$, giving $\theta_{n+1} = s_n$ and $\tau_{n+1} = 1/K$, this being the same waiting times as in (ii). Here we see that making the measurement angle adaptive significantly improves the performance.

(iv) The Greedy$_4$ algorithm of Eqs. (103), which does much better again.

(v) MOAAAR, which like (iii) uses Eqs. (113), but this time with the optimized $\Theta^\star = 1.50055$. This is better than Greedy, but only by a small amount.

Finally in Fig. 8, we also plot the analytical result for MOAAAR from Eq. (121). This agrees quite well with the numerics, but not perfectly. Indeed, the predicted decoherence is worse than what is achieved from the exact simulations, and in fact is worse even than the decoherence achieved by Greedy. This may surprise the reader, since we claimed in the preceding section that all the approximations leading to Eq. (121) were justified in the asymptotic regime, Eq. (2). The issue is that $K/\bar{\gamma} = 20$ is not far enough into the asymptotic regime to see the accuracy of the asymptotic analytical result.

We address this in Fig. 9, which compares the decay rates of the two algorithms for various values of K. These rates are obtained as the slopes of straight lines fitted to data
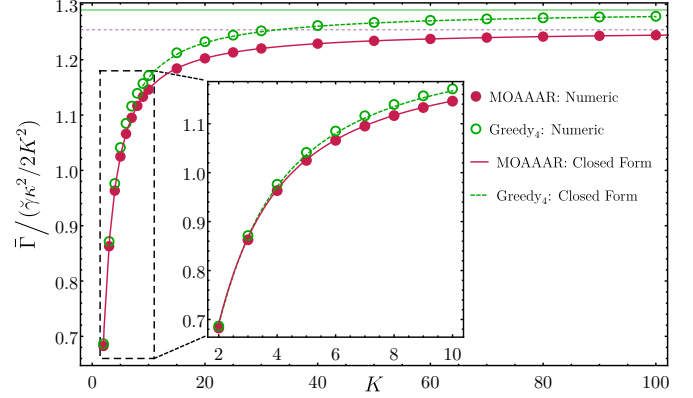
such as those shown in Fig. 8, throwing away the first two data points to avoid transient effects. We see that, using the rate in Eq. (121) and H_Θ in Eq. (122), the MOAAAR and Greedy decoherence rates, scaled by $2K^2/(\check{\gamma}\kappa^2)$, do appear to approach the analytic asymptotic values of $H^\star \approx 1.254$ and $H_{\pi/2} \approx 1.290$, respectively, for large K. This plot better shows that the decoherence rates for MOAAAR are always smaller than those for Greedy, even for relatively small K, where both algorithms give almost the same decoherence rates, which can be seen in the inset of Fig. 9.

In Fig. 9 we also plot the decoherence rates calculated with an analytical closed-form expressions, which is derived in Sec. VIII. This closed-form expression agrees very well with the numerical results, especially for MOAAAR, over almost two orders of magnitude of variation in K. Moreover, the closed form expression can be shown to approach the asymptotic values as $K \to \infty$, with a relative deviation scaling as $\bar{\gamma}/K$.

E. Further numerical evidence for optimality of MOAAAR

In the previous section, we showed that MOAAAR performs better than other discussed algorithms. Of course, there are infinitely many possible algorithms that one could compare to, even restricting to the form of strategy in Eq. (91), derived from the separation principle (Sec. IV B) applied to the sufficient statistics of Sec. V D. It is because we do not know any way to make such a comparison that we have claimed MOAAAR only to be *plausibly* optimal, in the asymptotic regime. The claim of plausibility is, naturally, subjective, but is based on our detailed study of the problem. Nevertheless, the reader may ask whether more evidence for

the optimality of MOAAAR can be given. To address this, in this subsection we consider a more general algorithm.

As the reader will recall, MOAAAR's best measurement angle $\Theta = \Theta^\star \approx 1.50055$ was found under the constraint in Eqs. (113), which assumed the relationship $\tau = \Theta/K$ between the magnitude of the measurement angle and the waiting time. This was chosen based on an analysis of Greedy$_2$ in the asymptotic regime, and has an intuitive interpretation in terms of optimal state discrimination; see Sec. VI F. In this subsection, we relax this constraint and numerically calculate the decoherence rate over a wide range of parameters: $\Theta \in [0, \pi]$ and $\tau \in [1/K, 6/K]$. We keep the natural adaptive strategy for the measurement angle, $\theta_{n+1} = s_n\Theta$, where $s_n = \text{sgn}(\zeta_n)$, as any other choice would break the symmetry between the RTP states.

We first investigate the phase-space dynamics and find that, outside the regime where $\tau \approx \Theta/K$, the very simple dynamics as shown in Fig. 6 or Fig. 7 are no longer evident. Instead, these strategies generally result in highly scattered state dynamics in the phase space. This means that we can no longer use eigenstates of the **F** maps to calculate the decoherence rate. We therefore need to use the brute-force calculation following Eq. (98) to compute the exact coherence for a given strategy from averaging over all possible realizations of measurement results Y_n. From the coherence data over time, similar to the plots in Fig. 8, we then compute a decoherence rate by fitting the data with a linear function and extracting its slope.

We also note that the further the measurement strategies are from the regime $\tau \approx \Theta/K$, the longer time it takes for the coherence process to reach its asymptotic behavior of exponential decay. (The "long" times we consider are still much shorter than the reciprocal of the decay rate, which is why using a linear fit for the long-time decoherence curve is permissible.) This is expected as, if more points in phase space are accessible, it may take longer to reach the ergodic regime. Thus, to avoid the transient effects, we perform the numerical calculations for more steps than previously (in this case, $T/\tau = 15$ steps) and use only the last five steps for the linear-model regression. Also, in order to be confident about the decoherence rate extracted from the slope, we calculate the coefficient of determination, known as the R^2 value [54], of the linear regression fit. Its value ranges from 0 to 1, with $R^2 = 1$ representing perfect fitting. We find that we can have extremely high confidence in our linear fits ($R^2 > 0.998$) except in a tiny region of the parameter space we explore, which is very far from the optimal regime; see Fig. 10(a).

We present in Fig. 10 our numerical results for the decoherence rate divided by the factor $\check{\gamma}\kappa^2/(2K^2)$, the same as in Fig. 9. In Fig. 10(a) we show this for the full range of parameters. There is one global minimum and another local minimum, which correspond to $\{\Theta, \tau\} \approx \{1.5, 1.5/K\}$ and $\{\Theta, \tau\} \approx \{1.5, (1.5 + \pi)/K\}$, respectively. (The pattern of local minima is presumably repeated for $\tau \approx (1.5 + n\pi)/K$ for all integers n.) In order to see clearly the optimal point, we zoom into the region $\Theta \in [1.4, 1.6]$ and $\tau \in [1.4/K, 1.6/K]$ in Fig. 10(b), where the decoherence rates are calculated with finer grid values of Θ and τ. The results show that the smallest values of the decoherence rate all occur around the line $\Theta = K\tau$ (shown as the white dashed line). We also find
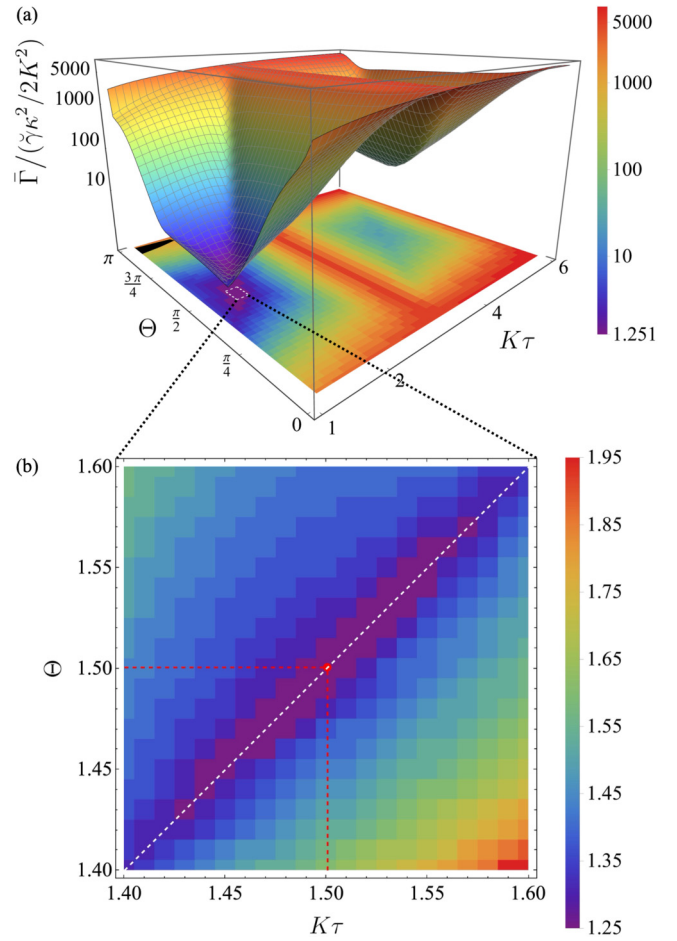

FIG. 10. Numerical decoherence rates for different SQ's adaptive measurement strategies defined by measurement angles $\theta_n = s_n\Theta$ and waiting times $\tau_n = \tau$, with $\gamma_\uparrow = \gamma_\downarrow = 1$, $\kappa = 0.2$, and $K = 100$. (a) The decoherence rate as a 3D colored surface plot with a colored contour plot (projection of the surface plot on the Θ–$K\tau$ plane) for $\Theta \in [0, \pi]$ and $\tau \in [1/K, 6/K]$. Different colors (on the color bar) indicate different values of the decoherence rate divided by the factor $\check{\gamma}\kappa^2/(2K^2)$. Note the log scale for this axis and color spread. The tiny black area of the plane at $\Theta \approx \pi$, $K\tau \approx 1$ shows where we do not have extreme confidence in the calculated rate. (b) A zoomed-in region, $\Theta \in [1.4, 1.6]$ and $\tau \in [1.4/K, 1.6/K]$, which includes a segment of the line $\tau = \Theta/K$ (white dashed line), which was the ansatz adopted to find MOAAAR, including the point $\{\Theta^\star, \tau^\star\} = \{1.50055, 1.50055/K\}$ (red dot) defining MOAAAR itself.

that $\Theta = 1.50055$ and $\tau = 1.50055/K$ (shown as the red dot) indeed gives the smallest value of $\bar{\Gamma}/(\check{\gamma}\kappa^2/2K^2) = 1.244$, which agrees perfectly with the optimal measurement strategy chosen by MOAAAR. These numerical results thus do offer more evidence that MOAAAR is a plausibly optimal adaptive algorithm for measuring the SQ.

VIII. CLOSED-FORM EXPRESSION FOR DECOHERENCE RATE IN NON-ASYMPTOTIC REGIME

In Sec. VII D we defined a simple adaptive strategy, MOAAAR, by finding the optimal value $\Theta^\star$ that minimizes the average decoherence rate, Eq. (118), computed from the

stable eigenstates of the two $\mathbf{F}$ maps in the asymptotic limit $K \gg \bar{\gamma}$. In this section, we show that we can generalize this to find a closed-form expression for the average decoherence rate even in the nonasymptotic regime. Moreover, it will apply for Greedy$_4$ and Greedy$_2$ as well as MOAAAR.

We do this by taking into account the small probabilities that the system's states can be outside of the eigenstates. In particular, guided by the numerical results in Fig. 5 and Fig. 6, we assume that as well as the two stable eigenstates (on the right side of the phase space) the system also spends time in another two less likely states on the left side of the phase space. Using all the four states and considering the probability transfers among them, we can then recalculate the average decoherence rate in closed form. For simplicity, we consider a symmetric-jump case, $\gamma_\uparrow = \gamma_\downarrow$ (though all of this can be simply adapted to the asymmetric case), where there is a symmetry in swapping between $s = +1$ and $s = -1$.

We first define the four states, using the maps defined in Eq. (108), which we reproduce here for convenience:

$$\mathbf{F}_{s,a}^{y} := \mathbf{F}\left(s\Theta(a), \Delta(a)/K, y\right). \tag{125}$$

Recall that $a = <$ refers to the left side of phase space and $a = >$ to the right side. The two right states (with $s = \pm 1$) are the stable (largest absolute eigenvalue) eigenvectors, $\underline{E}_s^0$, of $\mathbf{F}_{s,>}^0$. The left states are the states that these jump to when a nonnull result occurs, $\mathbf{F}_{s,>}^1 \underline{E}_s^0$. Because we are assuming the symmetry $\gamma_\uparrow = \gamma_\downarrow$, we need only one right and one left state, and we define two *normalized complex vectors* for these as

$$\underline{r} = \underline{E}_+^0 / \underline{I}^\top \underline{E}_+^0, \tag{126a}$$

$$\underline{l} = \mathbf{F}_{+,>}^1 \underline{r} / \underline{I}^\top \mathbf{F}_{+,>}^1 \underline{r}. \tag{126b}$$

Note that $\underline{l}$ is a state with $s = -1$ and $a = <$.

In order to calculate the average decoherence rate, we also need to know the probability that the system be in each of the four states. In the symmetric case, we can reduce the problem to only finding probabilities of the system being either on the right or left sides of the phase space, of which we denote the probabilities by $p^>$ and $p^< = 1 - p^>$, respectively. To compute the probabilities, we need the real matrices defined in Eq. (79). Using the same compact notation as in Eq. (125), we write these as $\check{\mathbf{F}}_{s,a}^y$. These matrices satisfy

$$\forall s, a, \quad \sum_y \underline{I}^\top \check{\mathbf{F}}_{s,a}^y = \underline{I}^\top. \tag{127}$$

That is, for all four cases $s \in \{+1, -1\}$ and $a \in \{>, <\}$, $\sum_y \check{\mathbf{F}}_{s,a}^y$ is a transition matrix, known in stochastic process theory as a column-stochastic matrix, with $\check{\mathbf{F}}_{s,a}^1$ and $\check{\mathbf{F}}_{s,a}^0$ known as substochastic matrices [55,56]. From these we define *normalized probability vectors*, $\check{\underline{r}}$ and $\check{\underline{l}}$ similar to Eqs. (126) as

$$\check{\underline{r}} = \check{\underline{E}}_+^0 / \underline{I}^\top \check{\underline{E}}_+^0, \tag{128a}$$

$$\check{\underline{l}} = \check{\mathbf{F}}_{+,>}^1 \check{\underline{r}} / \underline{I}^\top \check{\mathbf{F}}_{+,>}^1 \check{\underline{r}}, \tag{128b}$$

where $\check{\underline{E}}_+^0$ is the eigenvector of $\check{\mathbf{F}}_{+,>}^0$ with the largest absolute eigenvalue. Note that the divisors here ensure that $\check{\underline{r}}$ and $\check{\underline{l}}$ have positive elements, which are the conditioned Bayesian probabilities for $z = \pm 1$ when the system has normalized coherence vector $\underline{r}$ and $\underline{l}$, respectively. This is because the sequence of

results, Y_n, leading to $\underline{r}$ or $\underline{l}$ via the $\mathbf{F}$ maps will lead to $\check{\underline{r}}$ or $\check{\underline{l}}$ via the $\check{\mathbf{F}}$ maps.

Since a sum of elements (i.e., a 1-norm) of a probability vector, as defined in Eq. (77), tells us a probability of any stochastic measurement records leading up to that state, a norm of the form $\underline{I}^\top \check{\mathbf{F}}_{+,>}^y \check{\underline{r}}$ is the probability that a system in a right state $\underline{r}$ will yield the result y. If the result is $y = 0$ then it will remain in the same state. If it is $y = 1$ then it will jump to the left state $\underline{l}$. Similarly, the norm $\underline{I}^\top \check{\mathbf{F}}_{-,<}^y \check{\underline{l}}$ is the probability that a system in a left state $\underline{l}$ will yield the result y. (Note the change in s and a, reflecting the adaptive measurement.) Therefore, at any particular time in the steady-state regime, we can find a probability of the system being in the right states from the two contributions: (1) the right states in the past step [Eq. (128a), with the weight $p^>$] were mapped to themselves via the maps with null results and (2) the left states in the past step [Eq. (128b), with the weight $p^< = 1 - p^>$] were mapped to the right states via the maps with null results. This can be translated into an equation for $p^>$ as

$$p^> = p^> |\underline{I}^\top \check{\mathbf{F}}_{+,>}^0 \check{\underline{r}}| + (1 - p^>) |\underline{I}^\top \check{\mathbf{F}}_{-,<}^0 \check{\underline{l}}|, \tag{129}$$

which has a trivial exact solution.

Now that we have the probabilities $p^>$ and $p^<$, we can modify the formula of the average decoherence rate in Eq. (118) for the left and right states with the probability weights. The sum over $s \in \{+1, -1\}$ in Eq. (118) (which would be equally weighted in the symmetric case) can be replaced with an unequally weighted sum over the right and left states, while the sum over y stays unchanged. For the waiting time in Eq. (118), we replace it with a weighted average time $p^> \tau^> + p^< \tau^<$. (Recall that $\tau^a = \Delta^a/K$.) Thus, we have the 4-state approximation of the decoherence rate,

$$\bar{\Gamma}_4 = \frac{1 - \sum_y [|p^> \underline{I}^\top \mathbf{F}_{+,>}^y \underline{r}| + |p^< \underline{I}^\top \mathbf{F}_{-,<}^y \underline{l}|]}{p^> \tau^> + p^< \tau^<}. \tag{130}$$

Here the subscript 4 is used to indicate that this is a four-state approximation, though only two appear because of the symmetry assumed for simplicity in this section.

For Greedy$_4$, the measurement angles and the waiting times to be used in the $\mathbf{F}$ maps in Eq. (125) are as defined in Eqs. (103), namely $\Theta(a) = \Theta_G^a$, and $\Delta(a) = \Delta_G^a$. In the asymptotic regime, where the two parameters of Greedy$_2$ are sufficient, we have $\Theta(a) = \Delta(a) = \Theta_G^a$. For MOAAAR, we have just the single parameter: $\Theta(a) = \Delta(a) = \Theta^\star$. In this last case we can drop the a dependence on the $\mathbf{F}$ maps, and the average decoherence rate can be further reduced to

$$\bar{\Gamma}_4^\star = \frac{K}{\Theta^\star}\left(1 - \sum_y [|p^> \underline{I}^\top \mathbf{F}_+^y \underline{r}| + |p^< \underline{I}^\top \mathbf{F}_-^y \underline{l}|]\right). \tag{131}$$

As already discussed, Fig. 9 shows that Eq. (130) gives an excellent approximation, for both MOAAAR and Greedy, even far from the asymptotic regime. It is worth noting that we also used this closed-form approximation in Fig. 3 of CL [28]. In the regime of relatively small K, specifically for $K \lesssim 5\bar{\gamma}$, the decoherence due to Greedy and MOAAAR are practically identical according to the closed-form approximation Eq. (130), and extremely close according to slope fitting to the exact numerical coherence calculation.
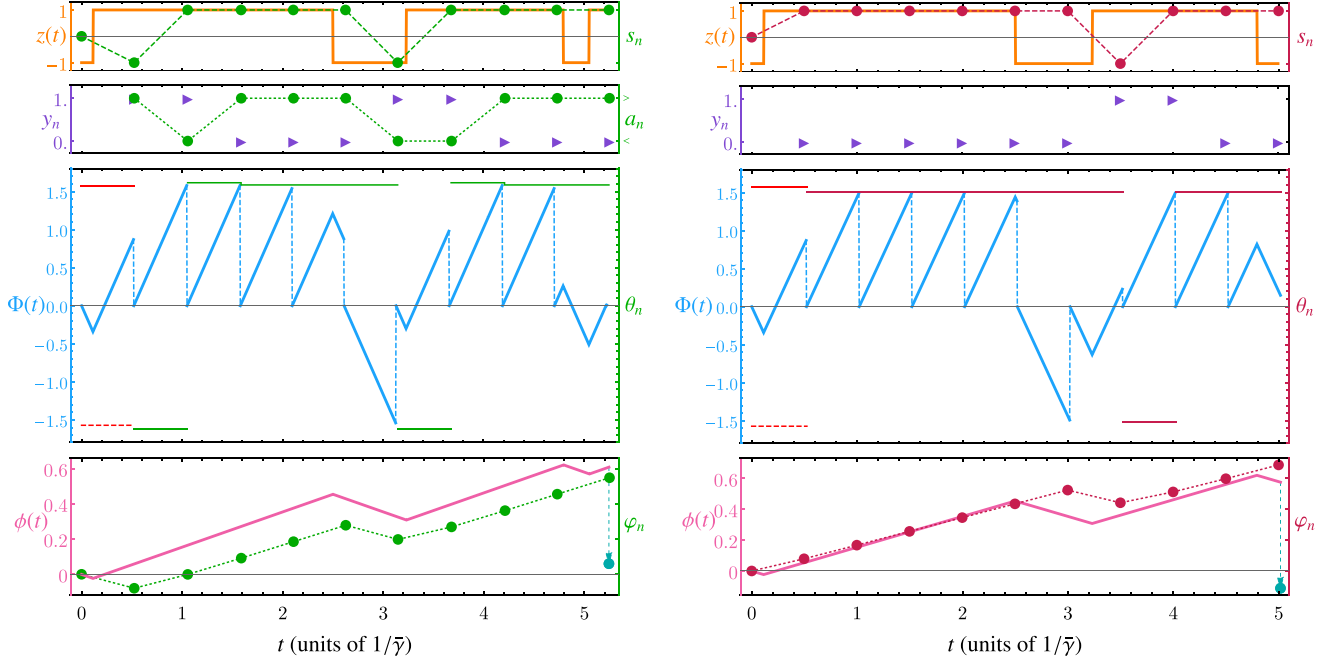
FIG. 11. Comparison of typical behavior for Greedy$_4$ (left panels) and the MOAAAR (right panels) for $K = 3$, $\gamma_\uparrow = \gamma_\downarrow = 1$, and $\kappa = 0.2$. The top row shows a typical realization of the RTP, $z(t)$ (orange solid line), which is the same in both panels, apart from the fact that the total time on the left is greater because, although 10 measurements are shown in both cases, Greedy chooses longer measurement waiting times. Also in the top row, $s_n = \mathrm{sgn}(\zeta)$ (circles at the measurement times) is the maximum likelihood estimate of $z(t)$ given Y_n (all measurement results up to and including y_n). The second row shows the results y_n (purple triangles). The left panel also shows a_n (circles), which is the second statistic derived from Y_n that the Greedy algorithm uses (together with s_n) to determine the waiting time τ_{n+1}, and angle θ_{n+1} for, the $(n + 1)$th measurement. MOAAAR uses only s_n for its adaptive choice of $\theta_{n+1} = s_n \Theta^\star$ while keeping $\tau_{n+1} = \Theta^\star / K$. The third row shows the phase of the SQ, $\Phi(t)$, which is reset to 0 after each measurement. The solid horizontal line segments from t_n to t_{n+1} shows the measurement angle θ_{n+1} chosen by the algorithms [except the first one (red) which is chosen to be $\pi/2$ (or, equivalently, $-\pi/2$, dashed)]. Having $\Phi(t_{n+1}) = \theta_{n+1}$ guarantees a null result, $y_{n+1} = 0$. The last row shows the phase of the data qubit, $\phi(t)$ (solid pink line) and the best estimate of it, φ_n (circles), conditioned on Y_n. At the final time the estimated phase is used to correct (teal arrow) the unwanted accumulated phase, taking $\phi(t_N)$ to $\phi(t_N) - \varphi_N$ (teal circle).

We close this section (and the results of this paper) by giving a comparison of the workings of Greedy and MOAAAR in the regime of relatively small K, in Fig. 11. In particular, we choose $K = 3\bar{\gamma}$, to accentuate the differences between the two algorithms. We plot a typical trajectory, both for the RTP, $z(t)$, which is the same in both cases, and for the measurement results Y_n, which are different in the two cases, because obviously the measurement strategy $\mu_n = (\theta_n, \tau_n)$ differ. We also plot the Bayesian parameters s_n and a_n which determine μ_n in the Greedy case. We omit a_n in the MOAAAR case because only s_n is used in its adaptive algorithm.

Note that in the left panel of the second row of Fig. 11, a_n is perfectly correlated with y_n [except for the very first measurement where $\theta_1 = \pm\pi/2$ and initial coherence vector $\underline{A}_0 = \underline{P}_{\mathrm{ss}}$ as in Eq. (80)]. Thus Greedy$_4$ does not actually need to calculate a_n in order to implement its algorithm. That is, a_n equals $>$ whenever there is a null result, $y_{n+1} = 0$, and equals $<$ whenever there is a nonnull result, $y_{n+1} = 1$. A null result is guaranteed if $\Phi(t_{n+1}) = \theta_{n+1}$. This condition can (and, in fact, usually does) occur for MOAAAR, but it never occurs for Greedy because Greedy always chooses $K\tau_{n+1} < |\theta_{n+1}|$. This is as shown in Fig. 4 and can be seen upon close inspection of Fig. 11: $\Phi(t)$ never quite reaches the green horizontal-line

segments on the Greedy side, while it often reaches the maroon line segments on the MOAAAR side.

It is also interesting to look at the data qubit phase $\phi(t) = \kappa \int_0^t z(t)$, and its best estimate, φ_n at time t_n. The absolute value of the slopes of the line segments for the former equal κ, and this is greater than the absolute values of the slopes of the line segments of the latter. That is, the Bayesian algorithms do not estimate $\phi(t)$ by simply integrating s_n, their most likely estimates of $z(t)$ at time t_n; in equation form, $\varphi_n \neq \sum_{j=1}^n s_j \kappa \tau_j$. The plot shows that the optimal Bayesian estimate and final control of the data qubit phase can be surprisingly effective, even when the algorithms are quite often wrong about the RTP, with $s_n \neq z(t_n)$. Though it is not precisely true that $\varphi_N = \kappa \tau \sum_{n=1}^N s_n$, this approximation is good enough for a rough derivation of the power-law scaling, $\check{\gamma} \kappa^2 / K^2$, for the decoherence rate arising for both MOAAAR and Greedy. The simple argument is as follows.

From Sec. III D, by approximating circular statistics by linear statistics, one sees that when the decoherence, $\mathcal{D} := 1 - \mathcal{C}^c$, is small, it equals the mean square error (MSE) in the final phase estimate, i.e.,

$$\mathcal{D} \approx \langle [\varphi_N - \phi(T)]^2 \rangle. \tag{132}$$

Now, from Fig. 11, there will be a substantial contribution to the error in the final estimate φ_N from the cases where the algorithm's guess, s_n, for z_n is wrong. Thus, since we are interested here only in the scaling of the MSE, we can approximate the MSE in Eq. (132) as

$$\mathcal{D} \sim (\kappa\tau)^2 \left\langle \left[\sum_{n=1}^{N} (s_n - z_n) \right]^2 \right\rangle, \tag{133}$$

where the contribution in the sum is nonzero only when the guess is wrong. The proportion of intervals where s_n and z_n differ is just the error probability of Eq. (124), which is of order γ/K. This scaling of the error probability, $O(\gamma/K)$, is also intuitive to understand, since the probability for a transition in the RTP between SQ measurements scales like this when the waiting times scale as $1/K$. Note that, for simplicity, we are not worrying about the distinction between $\gamma_\uparrow$ and $\gamma_\downarrow$, so $\check{\gamma} = \gamma_\downarrow = \gamma_\uparrow = \bar{\gamma} = \gamma$.

Thus the sum in Eq. (133) is just the sum of a large number ($\tilde{N} = O(N\gamma/K)$) of independent variables $e_{\tilde{n}}$ of random signs and magnitude 2:

$$\mathcal{D} \sim (\kappa\tau)^2 \left\langle \left[\sum_{\tilde{n}=1}^{\tilde{N}} e_{\tilde{n}} \right]^2 \right\rangle. \tag{134}$$

Since $\langle e_{\tilde{n}} \rangle = 0$ and the variance of a sum of independent random numbers is the sum of the variances, the mean square of that sum is $4\tilde{N}$. Now $N \sim T/\tau$, and $\tau \sim 1/K$ so we obtain finally

$$\mathcal{D} \sim T\gamma(\kappa/K)^2, \tag{135}$$

as expected for a decoherence rate $\Gamma \sim \gamma\kappa^2/K^2$.

IX. CONCLUSION

In this paper, we looked into noise mitigation of a data qubit using a spectator qubit (SQ). We considered dephasing in a data qubit caused by a random telegraph process (RTP) with transition rates $\gamma_\uparrow$ and $\gamma_\downarrow$. We assumed that the SQ is affected by the RTP in the same way as the data qubit, but with much greater sensitivity, $K \gg \kappa$. We then investigated various adaptive measurement and control strategies to reduce the data qubit dephasing. A locally optimal (Greedy) algorithm greatly reduced the no-control decoherence by a factor of order $(\bar{\gamma}/K)^2$ in the asymptotic regime $K \gg \bar{\gamma} \equiv (\gamma_\uparrow + \gamma_\downarrow)/2$. This enables us to develop a globally optimized algorithm that could reduce the decoherence rate even more, a Bayesian Map-based Adaptive Algorithm in the Asymptotic Regime (MOAAAR), with a reduction factor of $1.254(\bar{\gamma}/K)^2$.

Our results show that, in the right regime, the SQ, like techniques such as Dynamical Decoupling and Quantum Error Correction, can mitigate noise arbitrarily well. The fact that we have a data qubit decoherence rate that is, plausibly, the minimum possible achievable within our model, makes it a useful benchmark for comparison with other techniques for optimizing control, such as machine learning. Numerical and heuristic techniques may be necessary for more complex scenarios, as discussed below, so knowing how quickly, or closely, they can approach optimal performance is useful.

In our analysis, we assumed that the controller has full knowledge of the dynamical parameters κ, $\gamma_\uparrow$, $\gamma_\downarrow$, and K.

In practice, one would need to characterize these parameters prior to applying noise mitigation via the SQ. This characterization could be done via a suitable multiqubit *quantum noise spectroscopy* protocol such as those described in [57–59]. Interestingly, this can be done using SQ itself [60,61]. It would also be useful to study how the uncertainty in characterization would affect our algorithm's performance, and evaluate what the error thresholds in characterization are for our method to work well, and which parameters are most important. We expect that the least important parameters are $\gamma_\uparrow$ and $\gamma_\downarrow$, followed by K, and then κ. For practical applications we would also need to model imperfections including finite SQ's measurement readout times, imperfect projective measurement of the SQ, and additional decoherence of the SQ.

Beyond this first tranche of future work, several generalizations of our work are possible and may be useful. Typically, qubits are not affected by a single two-level RTP, but by several. This could be modelled by a multilevel RTP, and the generalization should be relatively straightforward. Beyond that, there is the question of general types of classical noise. Moreover, we assumed that SQ and data qubit feels the same noise, but this assumption could be relaxed; the SQ is still potentially useful as long as these noises are correlated. With regard to this topic, Youssry *et al.* [60] showed that machine learning can effectively find the correlation between the noises on the SQ and on the data qubit. Finally, we could consider continuous weak measurement of the SQ [62,63] rather than discrete-time projective measurements.

APPENDIX A: DERIVATION OF H MATRIX

Here we derive the matrix elements Eq. (26) which are described by Eq. (25), which we rewrite here:

$$H_{z_0}^{z_t}(t, \kappa) := \int e^{i\kappa X} \wp\,(X, z_t | z_0)\, dX. \tag{A1}$$

The above equation is just a Fourier transform of $\wp\,(X, z_t | z_0)$ evaluated at a conjugate variable κ. We note that the final result in (10) can be derived via many different techniques. In what follows, we derive the elements of matrix **H** using matrix diagonalization.

It is more convenient to work with discrete times and we can take the time-continuum limit at the end. We first divide the time duration of interest, $t \in [t_0, t]$, into M steps: $\{t_0, t_1, \ldots, t_m, \ldots, t\}$, where $t_m = t_{m-1} + \Delta t$ and Δt is an infinitesimal time. We note that t_m is not the same as the measurement time t_j in Eq. (44). We also assume that the

RTP value does not change during the infinitesimal time, i.e., $z(t \in (t_{m-1}, t_m]) = z(t_m)$. Later we will take the limit of $M \to \infty$ to recover the time-continuous limit.

The conditional probability in Eq. (A1) depends on z_0 and $z_t = z_M$, as well as the accumulated noise, X. Since we did break the total duration time into many infinitesimal steps, we can write the conditional probability, $\wp(X, z_t | z_0)$, in terms of the ones for the infinitesimal step, which we can solve analytically. We write the conditional probability as

$$\wp(X, z_t = z_M | z_0) = \sum_{z_1, \ldots, z_{M-1}} \delta\left(X - \sum_{m=1}^{M} z_m \Delta t\right) \wp(z_1, \ldots, z_M | z_0) = \sum_{z_1, \ldots, z_{M-1}} \delta\left(X - \sum_{m=1}^{M} z_m \Delta t\right) \wp(z_M | z_{M-1}) \cdots \wp(z_1 | z_0),$$
(A2)

where the first summation is over RTPs at multiple times and we have used $\delta(\cdots)$ as the Dirac δ function. We then substitute the above equation into Eq. (A1) and find

$$H_{z_0}^{z_t} = \int dX e^{i\kappa X} \sum_{z_1, \ldots, z_{M-1}} \delta\left(X - \sum_{m=1}^{M} z_m \Delta t\right) \wp(z_M | z_{M-1}) \cdots \wp(z_1 | z_0) = \sum_{z_1, \ldots, z_{M-1}} \exp\left(i\kappa \sum_{m=1}^{M} z_m \Delta t\right) \wp(z_M | z_{M-1}) \cdots \wp(z_1 | z_0),$$
(A3)

where in the second line we have used the integration property of Dirac δ function. We then distribute the exponential terms to accompany all the conditional probabilities and obtain

$$H_{z_0}^{z_t} = \sum_{z_1, \ldots, z_{M-1}} e^{i\kappa z_M \Delta t/2} [e^{i\kappa z_M \Delta t/2} \wp(z_M | z_{M-1}) e^{i\kappa z_{M-1} \Delta t/2}] \cdots [e^{i\kappa z_1 \Delta t/2} \wp(z_1 | z_0) e^{i\kappa z_0 \Delta t/2}] e^{-i\kappa z_0 \Delta t/2}$$

$$= e^{i\kappa z_M \Delta t/2} \left(\sum_{z_1, \ldots, z_{M-1}} M_{z_M, z_{M-1}} \cdots M_{z_1, z_0} \right) e^{-i\kappa z_0 \Delta t/2},$$
(A4)

where in the second line we have defined

$$M_{z, z'} := e^{i\kappa z \Delta t/2} \wp(z | z') e^{i\kappa z' \Delta t/2}.$$
(A5)

This $M_{z, z'}$ can be thought of as an element of a 2×2 matrix, where $z, z' \in \{-1, +1\}$. We can then use the results of Eqs. (6) and (7) and, and write

$$\wp(z | z') = \left(1 + \frac{1 - e^{-2\bar{\gamma} \Delta t}}{2\bar{\gamma}} \mathbf{J} \right)_{z, z'},$$
(A6)

which is simply the matrix term in Eq. (7). Therefore, the matrix $\mathbf{M}$ becomes

$$\mathbf{M} = \frac{1}{2\bar{\gamma}} \begin{pmatrix} (\gamma_\uparrow + \gamma_\downarrow e^{-2\bar{\gamma} \Delta t}) e^{i\kappa \Delta t} & \gamma_\uparrow - \gamma_\uparrow e^{-2\bar{\gamma} \Delta t} \\ \gamma_\downarrow - \gamma_\downarrow e^{-2\bar{\gamma} \Delta t} & (\gamma_\downarrow + \gamma_\uparrow e^{-2\bar{\gamma} \Delta t}) e^{-i\kappa \Delta t} \end{pmatrix}.$$
(A7)

Using the above matrix, we write the elements $H_{z_0}^{z_t}$ as

$$H_{z_0}^{z_t} = e^{i\kappa z_M \Delta t/2} (\mathbf{M}^M)_{z_M, z_0} e^{-i\kappa z_0 \Delta t/2}.$$
(A8)

Thus, the matrix $\mathbf{H}$ becomes

$$\mathbf{H} = \lim_{M \to \infty} \begin{pmatrix} e^{i\kappa \Delta t} & 0 \\ 0 & 1 \end{pmatrix} \mathbf{M}^M \begin{pmatrix} e^{-i\kappa \Delta t} & 0 \\ 0 & 1 \end{pmatrix},$$
(A9)

where we have included the limit of $M \to \infty$ to reach the continuous time limit (since $\Delta t = (t - t_0)/M$, in the limit of $M \to \infty$, we have $\Delta t \to 0$). We can calculate $\mathbf{M}^M$ by decomposing $\mathbf{M}$ into $\mathbf{M} = \mathbf{V} \mathbf{D} \mathbf{V}^{-1}$, where

$$\mathbf{D} = \begin{pmatrix} L^+ - S & 0 \\ 0 & L^+ + S \end{pmatrix}$$
(A10)

and

$$\mathbf{V} = \begin{pmatrix} \frac{2\bar{\gamma}}{\gamma_\downarrow} \frac{L^- - S}{1 - e^{-2\bar{\gamma} \Delta t}} & \frac{2\bar{\gamma}}{\gamma_\downarrow} \frac{L^- + S}{1 - e^{-2\bar{\gamma} \Delta t}} \\ 1 & 1 \end{pmatrix},$$
(A11)

with

$$L^{\pm} := \frac{1}{4\bar{\gamma}}[(\gamma_{\uparrow} + \gamma_{\downarrow}e^{-2\bar{\gamma}\Delta t})e^{i\kappa\Delta t} \pm (\gamma_{\downarrow} + \gamma_{\uparrow}e^{-2\bar{\gamma}\Delta t})e^{-i\kappa\Delta t}],\tag{A12}$$

$$S := \sqrt{(L^{-})^2 + \frac{\check{\gamma}}{4\bar{\gamma}}(1 - e^{-2\bar{\gamma}\Delta t})^2}.\tag{A13}$$

Since **D** is a diagonal matrix, we have $\mathbf{M}^M = \mathbf{V}\mathbf{D}^M\mathbf{V}^{-1}$, with

$$\mathbf{D}^M = \begin{pmatrix} (L^+ - S)^M & 0 \\ 0 & (L^+ + S)^M \end{pmatrix}.\tag{A14}$$

We then substitute these results in Eq. (A9), and after taking the limit, we find the final result as

$$\mathbf{H}(t,\kappa) = \exp[-\bar{\gamma}t]\begin{pmatrix} \cosh\left[\frac{\lambda(\kappa)}{2}t\right] - \frac{\eta(\kappa)}{\lambda(\kappa)}\sinh\left[\frac{\lambda(\kappa)}{2}t\right] & \frac{2\gamma_{\uparrow}}{\lambda(\kappa)}\sinh\left[\frac{\lambda(\kappa)}{2}t\right] \\ \frac{2\gamma_{\downarrow}}{\lambda(\kappa)}\sinh\left[\frac{\lambda(\kappa)}{2}t\right] & \cosh\left[\frac{\lambda(\kappa)}{2}t\right] + \frac{\eta(\kappa)}{\lambda(\kappa)}\sinh\left[\frac{\lambda(\kappa)}{2}t\right] \end{pmatrix}\tag{A15}$$

with

$$\lambda(\kappa) = \sqrt{(\gamma_{\downarrow} + \gamma_{\uparrow})^2 - 4i\kappa(\gamma_{\downarrow} - \gamma_{\uparrow}) - 4\kappa^2},\tag{A16a}$$

$$\eta(\kappa) = (\gamma_{\downarrow} - \gamma_{\uparrow}) - 2i\kappa.\tag{A16b}$$

This result is in agreement with the Fourier transform introduced in Eq. (10).

APPENDIX B: DERIVATION OF F

In this section, we derive the elements of the matrix **F**. We start with the definition in Eq. (66), which we rewrite here,

$$F_{z_{n-1}}^{z_n}(\mu_n, y_n) := \int dx_n\, \wp_{\mu_n}(y_n|x_n)\wp(x_n, z_n|z_{n-1})e^{i\kappa x_n}.\tag{B1}$$

We recognize that the above equation is a Fourier transform of $\wp_{\mu_n}(y_n|x_n)\wp(x_n, z_n|z_{n-1})$, where formally we write

$$F_{z_{n-1}}^{z_n} = \mathcal{F}_{x_n \to k}[\wp_{\mu_n}(y_n|x_n)\wp(x_n, z_n|z_{n-1})],\tag{B2}$$

where $\mathcal{F}_{x \to k}[f]$ is defined as

$$\mathcal{F}_{x \to k}[f] := \int dx\, e^{ikx}\, f.\tag{B3}$$

We then use the convolution theorem, where a Fourier transform of a product of two functions is equal to a product of their individual Fourier transforms:

$$\mathcal{F}_{x \to \kappa}[f \cdot g] = \int dk\, \mathcal{F}_{x \to k}[f] \cdot \mathcal{F}_{x \to (\kappa - k)}[g].\tag{B4}$$

Applying the convolution theorem to Eq. (B2), we obtain

$$F_{z_{n-1}}^{z_n} = \int dk\, \mathcal{F}_{x_n \to k}[\wp_{\mu_n}(y_n|x_n)]\mathcal{F}_{x_n \to (\kappa - k)}[\wp(x_n, z_n|z_{n-1})],\tag{B5}$$

where each term in the integrand can be computed as the following. For the first one, we use Eq. (50), which we rewrite here:

$$\wp_{\mu_n}(y_n|x_n) = y_n + (-1)^{y_n}\cos^2\left[\tfrac{1}{2}(\theta_n - Kx_n)\right],\tag{B6}$$

where its Fourier transform is

$$\mathcal{F}_{x_n \to k}[\wp_{\mu_n}(y_n|x_n)] = \tfrac{1}{4}[2\delta_{k,0} + (-1)^{y_n}e^{-i\theta_n}\delta_{k,-K} + (-1)^{y_n}e^{+i\theta_n}\delta_{k,+K}],\tag{B7}$$

where δ here is the Kronecker delta function. For the second Fourier transform, we have

$$\mathcal{F}_{x_n \to (\kappa - k)}[\wp(x_n, z_n|z_{n-1})] = \int dx_n\, \wp(x_n, z_n|z_{n-1})e^{i(\kappa - k)x_n}$$
$$= H_{z_{n-1}}^{z_n}(\tau_n, \kappa - k),\tag{B8}$$

which we have calculated in the previous section. Combining these two terms and performing the integration over k, we find

$$F_{z_{n-1}}^{z_n} = \tfrac{1}{4}\big[2H_{z_{n-1}}^{z_n}(\tau_n, \kappa) + (-1)^{y_n}e^{-i\theta_n}H_{z_{n-1}}^{z_n}(\tau_n, \kappa + K) + (-1)^{y_n}e^{+i\theta_n}H_{z_{n-1}}^{z_n}(\tau_n, \kappa - K)\big].\tag{B9}$$

Since, in the above equation, all the indices are the same, we can write a matrix form of the above equation as

$$\mathbf{F} = \tfrac{1}{4}\{2\mathbf{H}(\tau_n, \kappa) + (-1)^{y_n} \times [e^{-i\theta_n}\mathbf{H}(\tau_n, \kappa + K) + e^{+i\theta_n}\mathbf{H}(\tau_n, \kappa - K)]\}.\tag{B10}$$

APPENDIX C: GREEDY OPTIMAL ANGLES

Here we explain details of the optimization process of the Greedy algorithm. The task is to find an optimum value for the measurement angles $\theta_{(i)}, \theta_{(ii)}$ and $\theta'_{(ii)}$, which maximize the coherence $\mathcal{C}_{(i)}^c$ and $\mathcal{C}_{(ii)}^c$ as discussed around Eqs. (95) and (96). In general, this optimization task can be computationally heavy since it needs to search over all possible values of the three angles, each in the range of $[-\pi/2, \pi/2]$. However, we can significantly simplify the task by using analytical results shown in Berry and Wiseman [51], where our coherence can be rewritten in their particular form of a cost function which has only three possible optimal values. Following the method presented in Ref. [51], we find that the optimum of θ_n should
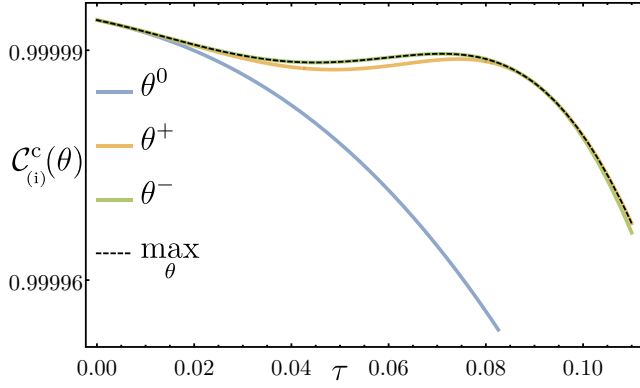
FIG. 12. An example of maximizing the coherence (reward function), $\mathcal{C}^c_{(i)}(\theta)$, over the three Berry-Wiseman angles in the Greedy algorithm. We can see that $\mathcal{C}^c_{(i)}(\theta^0)$ is always the smallest, while the other two are competing. The dashed line is the maximum of these three curves. The parameters for this plot are, $K = 20$, $\kappa = 0.2$ and $\gamma_\uparrow = \gamma_\downarrow = 1$.

be one of these three values:

$$\theta^0(\tau) = \arg(ba^* - c^*a), \tag{C1a}$$

$$\theta^\pm(\tau) = \arg\sqrt{\frac{c_2 \pm \sqrt{c_2^2 + |c_1|^2}}{c_1}}, \tag{C1b}$$

where

$$a = 2\underline{I}^\top \mathbf{H}(\tau, \kappa)\underline{A}_n, \tag{C2a}$$

$$b = \underline{I}^\top \mathbf{H}(\tau, \kappa + K)\underline{A}_n, \tag{C2b}$$

$$c = \underline{I}^\top \mathbf{H}(\tau, \kappa - K)\underline{A}_n, \tag{C2c}$$

and

$$c_1 = (a^*c)^2 - (ab^*)^2 + 4(|b|^2 - |c|^2)b^*c, \tag{C3a}$$

$$c_2 = -2i\,\mathrm{Im}(a^2b^*c^*). \tag{C3b}$$

We note that these angles are functions of the prospect waiting (to measure) time, τ. To implement the Greedy algorithm, we only need to check these three values at each time step, which is a much simpler task computationally. As an example, we show in Fig. 12 the coherence $\mathcal{C}^c_{(i)}$ for the three different Berry-Wiseman angles as functions of τ. The starting point is a value of $\mathcal{C}^c(t_n)$, right after the nth measurement, choosing $n = 2$. We show the optimal value, $\mathcal{C}^{op}_{(i)} = \max_{\theta_{(i)}} \mathcal{C}^c_{(i)}(\theta_{(i)})$, in the dashed curve as a function of τ.

In the Greedy algorithm in Eq. (95), we also have to optimize the coherence with two measurement angles, i.e., optimizing the function $\mathcal{C}^c_{(ii)}(\theta_{(ii)}, \theta'_{(ii)})$ over the total of nine pairs of $\theta_{(ii)}, \theta'_{(ii)}$(three options for each). We can simplify this task even further by realizing that the second measurement has a waiting time equal to dt after the first measurement. Thus we expect that during that infinitesimal time the absolute value of the accumulated phase by the spectator qubit is simply Kdt. Then we know the best angle should be $\theta'_{(ii)} = \mathrm{sign}(\zeta_n)Kdt$, which we also numerically checked that this angle does maximize the reward for the second measurement. So the optimization task only needs to search over the three possible angles of $\theta_{(ii)}$. We note here that we do not show plots for the two measurement case here, because it is very
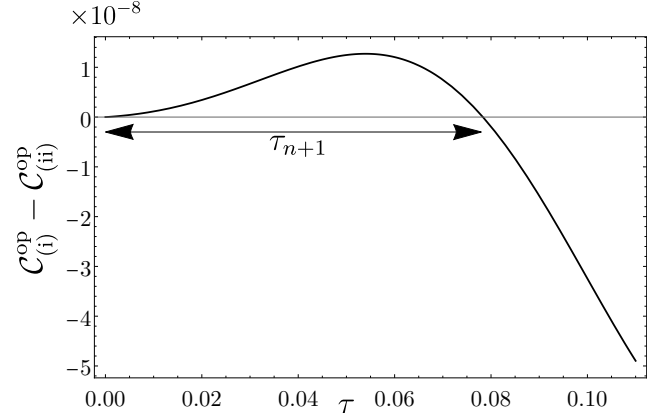

FIG. 13. An example of the process of choosing the next measurement (waiting) time τ_{n+1} in the Greedy algorithm. The plot shows the difference between $\mathcal{C}^{op}_{(i)}$ and $\mathcal{C}^{op}_{(ii)}$ as a function of τ. The waiting time is chosen when this function crossed the zero line. We used the same parameters as in Fig. 12.

similar to the one measurement case, with the difference of the order 10^{-8}.

The maximization of the coherence (reward function) at time t_n can then be used in choosing the next waiting time τ_{n+1} and angle θ_{n+1}. From the results above, we find the optimum angles for the one and two measurement scenarios as

$$\theta^{op}_{(i)}(\tau) := \arg\max_{\theta_{(i)}} \mathcal{C}^c_{(i)}(\theta_{(i)}), \tag{C4}$$

$$\theta^{op}_{(ii)}(\tau) := \arg\max_{\theta_{(ii)}} \mathcal{C}^c_{(ii)}(\theta_{(ii)}), \tag{C5}$$

which are still functions of arbitrary τ. To choose the next measurement setting, we first need to calculate to compare the reward between the two scenarios. That is, we compute the difference, $D(\tau) := \mathcal{C}^{op}_{(i)} - \mathcal{C}^{op}_{(ii)}$, which is a function of τ, shown in Fig. 13. So long as $D(\tau)$ is positive (i.e., $\mathcal{C}^{op}_{(i)} > \mathcal{C}^{op}_{(ii)}$) no measurement is required. On the other hand if $D(\tau) < 0$, we conclude that a measurement should have been done to increase the coherence. Thus, the next waiting time, τ_{n+1} is chosen when $D(\tau)$ crosses the zero line,

$$D(\tau_{n+1}) = 0. \tag{C6}$$

Now that we found τ_{n+1}, we can use it to identify the choice of the next measurement angle,

$$\theta_{n+1} = \theta^{op}_{(ii)}(\tau_{n+1}), \tag{C7}$$

using the maximized angles in Eq. (C4).

APPENDIX D: EXPRESSIONS FOR N_Θ AND M_Θ

Here we present the lengthy expressions of N_Θ and M_Θ that we used in Eq. (120),

$$N_\Theta = 24\big[\Theta - 2\Theta^3 + 3\Theta\cos 2\Theta + \tfrac{1}{4}\sin 4\Theta - \tfrac{1}{2}\sin 2\Theta + \Theta^2(8\cot\Theta + 4\Theta\csc^2\Theta - \sin 2\Theta)\big], \tag{D1}$$

$$\begin{aligned} M_\Theta = {}& -15 + 8\Theta^2(\Theta^2 - 15) \\ & - 12(2\Theta^2 - 1)\cos 2\Theta + 3\cos 4\Theta \\ & + 192\Theta^3\cot^3\Theta - 96\Theta^2(\Theta^2 - 1)\csc^2\Theta \\ & + 96\Theta^4\csc^4\Theta + 16\Theta^3\sin 2\Theta. \end{aligned} \tag{D2}$$

[1] K. Temme, S. Bravyi, and J. M. Gambetta, Phys. Rev. Lett. **119**, 180509 (2017).

[2] J. Preskill, Quantum **2**, 79 (2018).

[3] S. Endo, S. C. Benjamin, and Y. Li, Phys. Rev. X **8**, 031027 (2018).

[4] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Nature (London) **567**, 491 (2019).

[5] L. Viola, E. Knill, and S. Lloyd, Phys. Rev. Lett. **82**, 2417 (1999).

[6] L. Viola and E. Knill, Phys. Rev. Lett. **90**, 037901 (2003).

[7] M. Biercuk, A. Doherty, and H. Uys, J. Phys. B: At. Mol. Opt. Phys. **44**, 154002 (2011).

[8] H. K. Ng, D. A. Lidar, and J. Preskill, Phys. Rev. A **84**, 012305 (2011).

[9] A. M. Souza, G. A. Álvarez, and D. Suter, Phys. Rev. Lett. **106**, 240501 (2011).

[10] J. Medford, L. Cywiński, C. Barthel, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Phys. Rev. Lett. **108**, 086802 (2012).

[11] G. A. Paz-Silva and D. A. Lidar, Sci. Rep. **3**, 1530 (2013).

[12] J. Zhang, A. M. Souza, F. D. Brandao, and D. Suter, Phys. Rev. Lett. **112**, 050502 (2014).

[13] P. W. Shor, Phys. Rev. A **52**, R2493 (1995).

[14] A. M. Steane, Phys. Rev. Lett. **77**, 793 (1996).

[15] B. M. Terhal, Rev. Mod. Phys. **87**, 307 (2015).

[16] R. S. Gupta, C. L. Edmunds, A. R. Milne, C. Hempel, and M. J. Biercuk, npj Quantum Inf. **6**, 53 (2020).

[17] S. Majumder, L. Andreta de Castro, and K. R. Brown, npj Quantum Inf. **6**, 19 (2020).

[18] K. Singh, C. E. Bradley, S. Anand, V. Ramesh, R. White, and H. Bernien, arXiv:2208.11716 (2022).

[19] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. K. Vandersypen, Rev. Mod. Phys. **79**, 1217 (2007).

[20] J. J. Pla, K. Y. Tan, J. P. Dehollain, W. H. Lim, J. J. L. Morton, D. N. Jamieson, A. S. Dzurak, and A. Morello, Nature (London) **489**, 541 (2012).

[21] A. Morello, J. J. Pla, P. Bertet, and D. N. Jamieson, Adv. Quantum Technol. **3**, 2000005 (2020).

[22] A. Morello, J. J. Pla, F. A. Zwanenburg, K. W. Chan, K. Y. Tan, H. Huebl, M. Möttönen, C. D. Nugroho, C. Yang, J. A. van Donkelaar *et al.*, Nature (London) **467**, 687 (2010).

[23] D. Keith, M. G. House, M. B. Donnelly, T. F. Watson, B. Weber, and M. Y. Simmons, Phys. Rev. X **9**, 041003 (2019).

[24] J. Z. Blumoff, A. S. Pan, T. E. Keating, R. W. Andrews, D. W. Barnes, T. L. Brecht, E. T. Croke, L. E. Euliss, J. A. Fast, C. A. C. Jackson *et al.*, PRX Quantum **3**, 010352 (2022).

[25] D. Culcer and N. M. Zimmerman, Appl. Phys. Lett. **102**, 232108 (2013).

[26] A. Bermeister, D. Keith, and D. Culcer, Appl. Phys. Lett. **105**, 192102 (2014).

[27] D. Culcer, X. Hu, and S. Das Sarma, Appl. Phys. Lett. **95**, 073102 (2009).

[28] H. Song, A. Chantasri, B. Tonekaboni, and H. M. Wiseman, Phys. Rev. A **107**, L030601 (2023).

[29] A. B. Zorin, F.-J. Ahlers, J. Niemeyer, T. Weimann, H. Wolf, V. A. Krupenin, and S. V. Lotkhov, Phys. Rev. B **53**, 13682 (1996).

[30] E. Paladino, L. Faoro, G. Falci, and R. Fazio, Phys. Rev. Lett. **88**, 228304 (2002).

[31] T. Fujisawa and Y. Hirayama, Appl. Phys. Lett. **77**, 543 (2000).

[32] E. Paladino, Y. M. Galperin, G. Falci, and B. L. Altshuler, Rev. Mod. Phys. **86**, 361 (2014).

[33] T. Itakura and Y. Tokura, Phys. Rev. B **67**, 195320 (2003).

[34] J. Bergli, Y. M. Galperin, and B. L. Altshuler, Phys. Rev. B **74**, 024509 (2006).

[35] J. Bergli, Y. M. Galperin, and B. L. Altshuler, New J. Phys. **11**, 025002 (2009).

[36] J. T. Muhonen, J. P. Dehollain, A. Laucht, F. E. Hudson, R. Kalra, T. Sekiguchi, K. M. Itoh, D. N. Jamieson, J. C. McCallum, A. S. Dzurak, and A. Morello, Nat. Nanotechnol. **9**, 986 (2014).

[37] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 1985).

[38] K. Jacobs, *Stochastic Processes for Physicists: Understanding Noisy Systems* (Cambridge University Press, Cambridge, 2010).

[39] B. G. Christensen, C. D. Wilen, A. Opremcak, J. Nelson, F. Schlenker, C. H. Zimonick, L. Faoro, L. B. Ioffe, Y. J. Rosen, J. L. DuBois *et al.*, Phys. Rev. B **100**, 140503(R) (2019).

[40] B. M. Freeman, J. S. Schoenfield, and H. Jiang, Appl. Phys. Lett. **108**, 253108 (2016).

[41] M. Kim, H. J. Mamin, M. H. Sherwood, K. Ohno, D. D. Awschalom, and D. Rugar, Phys. Rev. Lett. **115**, 087602 (2015).

[42] M. Brownnutt, M. Kumph, P. Rabl, and R. Blatt, Rev. Mod. Phys. **87**, 1419 (2015).

[43] Y. M. Galperin, B. L. Altshuler, J. Bergli, and D. V. Shantsev, Phys. Rev. Lett. **96**, 097009 (2006).

[44] A. V. Kuhlmann, J. Houel, A. Ludwig, L. Greuter, D. Reuter, A. D. Wieck, M. Poggio, and R. J. Warburton, Nat. Phys. **9**, 570 (2013).

[45] S. Daniotti, C. Benedetti, and M. G. A. Paris, Eur. Phys. J. D **72**, 208 (2018).

[46] N. I. Fisher, *Statistical Analysis of Circular Data* (Cambridge University Press, Cambridge, 1995).

[47] D. P. Joseph and T. J. Tou, Trans. Am. Inst. Electr. Eng. II **80**, 193 (1961).

[48] K. J. Åström, *Introduction to Stochastic Control Theory* (Academic Press, New York, 1970).

[49] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, in *Introduction to Algorithms* (MIT Press, Cambridge, MA, 2009), Chap. 16, pp. 370–404.

[50] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).

[51] D. W. Berry, H. M. Wiseman, and J. K. Breslin, Phys. Rev. A **63**, 053804 (2001).

[52] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Mathematics in Science and Engineering Vol. 123 (Academic Press, New York, 1976).

[53] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2010).

[54] N. R. Draper and H. Smith, *Applied Regression Analysis*, 3rd ed., Wiley Series in Probability and Statistics (John Wiley & Sons, New York, 1998).

[55] A. Paz, *Introduction to Probabilistic Automata* (Academic Press, New York, 1971).

[56] J. Shallit and M.-W. Wang, J. Automat Lang. Comb. **6**, 537 (2001).

[57] G. A. Paz-Silva, L. M. Norris, and L. Viola, Phys. Rev. A **95**, 022121 (2017).

[58] U. von Lüpke, F. Beaudoin, L. M. Norris, Y. Sung, R. Winik, J. Y. Qiu, M. Kjaergaard, D. Kim, J. Yoder, S. Gustavsson *et al.*, PRX Quantum **1**, 010305 (2020).

[59] T. Chalermpusitarak, B. Tonekaboni, Y. Wang, L. M. Norris, L. Viola, and G. A. Paz-Silva, PRX Quantum **2**, 030315 (2021).

[60] A. Youssry, G. A. Paz-Silva, and C. Ferrie, arXiv.2103.13018 (2021).

[61] A. Youssry, G. A. Paz-Silva, and C. Ferrie, npj Quantum Inf. **6**, 95 (2020).

[62] A. N. Korotkov, Phys. Rev. B **60**, 5737 (1999).

[63] H.-S. Goan, G. J. Milburn, H. M. Wiseman, and H.-B. Sun, Phys. Rev. B **63**, 125326 (2001).