Maximum Information Gain in Weak or Continuous Measurements of Qudits: Complementarity Is Not Enough

Joshua Combes[1,2] and Howard M. Wiseman[1]

[1]*Centre for Quantum Computation and Communication Technology (Australian Research Council), Centre for Quantum Dynamics, Griffith University, Brisbane, Queensland 4111, Australia*
[2]*Center for Quantum Information and Control, University of New Mexico, Albuquerque, New Mexico 87131-0001, USA*
(Received 4 May 2011; published 28 September 2011)

To maximize average information gain for a classical measurement, all outcomes of an observation must be equally likely. The condition of equally likely outcomes may be enforced in quantum theory by ensuring that the system's state ρ is maximally different, or complementary, to the measured observable. This requires the ability to perform unitary operations on the state, conditioned on the results of prior measurements. We consider the case of measurement of a component of angular momentum for a qudit (a D-dimensional system, with $D = 2J + 1$, where J is the total angular momentum). For weak or continuous-in-time (i.e., repeated weak) measurements, we show that the complementarity condition ensures an average improvement in the rate of purification of only 2. However, we show that, by choosing the locally optimal control protocol of this type, one can attain the best possible scaling, $O(D^2)$, for the average improvement. For this protocol, the acquisition of information is nearly deterministic. Finally, we contrast these results with those for complementarity-based protocols in a register of qubits.

Subject Areas: Atomic and Molecular Physics, Quantum Physics, Quantum Information

I. INTRODUCTION

In the classical world, predictability is associated with intimate knowledge of a system, while unpredictability implies surprise. In information theory, the surprisal of an outcome k is defined as $I(k) \equiv -\log_2(P_k)$, where P_k is the probability of that outcome [1]. The surprisal is an important quantity because it quantifies the amount of information one learns from an outcome of a measurement. When the average surprisal of an experiment is maximized, the observer maximizes the extraction of information from the experiment.

In quantum theory, the relationship between predictability and knowledge is not so clear cut. Even when an observer has maximal knowledge about a quantum system (that is, the observer's state of knowledge is pure), measurement outcomes may be unpredictable. This is the case if measurements on the system are performed in a basis which is not the eigenbasis of the state. (The eigenbasis is the basis in which the operator under consideration, e.g., ρ, has entries only on the diagonal.) Nevertheless, pure quantum states are the most predictable states available. It is for this reason, and for other reasons eloquently explained in Ref. [2], that we take the impurity to be a measure of information. These considerations have led to information [2], and they control [3] theoretic formulations of information–disturbance relations.

Inspired by this work, Jacobs asked the question: How quickly on average can information be extracted from a quantum system using continuous quantum measurements in a complementary basis [4]? (Other ways of quantifying the rate of information gain can also be considered [5,6].) He found that, for a widely applicable continuous-measurement model, measuring in a complementary basis led to a speeding-up in the purification by a factor of 2 for a two-level system [4]. (Practically, such a rapid purification protocol can be a primitive for a rapid state preparation protocol [6,7].) Later this was generalized to complementary bases for D-dimensional systems. In Ref. [8], the lower bound on the speeding-up for D-dimensional systems was found to be $S_{LB} = (2/3)(D + 1)$ when the monitoring was performed in a complementary basis. Recently, in Ref. [9], the upper bound S_{UB} was proven to be bounded above by $D^2/2$ for measurement in any complementary basis.

In this paper, we give an explicit construction of a complementary-measurement protocol for a qudit (that is, a D-dimensional system, with $D = 2J + 1$, where J is the total angular momentum) based on the quantum Fourier transform (QFT), which achieves a speeding-up of $0.2D^2$. Furthermore, we show that transforming the state's eigenbasis so that it is complementary with respect to the measurement eigenbasis is *not* sufficient to achieve this speeding-up; the speeding-up can be as low as 2. In addition, we show, numerically, that unbiased bases can achieve the upper bound, saturating the inequality $S \leq D^2/2$, when D is an even number. To define these speeding-ups rigorously, it is necessary to derive rigorous bounds on the information-extraction rate for commuting measurements without control, so we also do that in this

paper. (For the rate of commuting measurements with control, see Refs. [5,7].) Finally, we consider measurements in commuting and complementary bases for a register of qubits (where a qubit is a $D = 2$ qudit). The complementary-measurement scheme based on the QFT protocol appears to give a speeding-up of only 2 in the case of a register.

II. FURTHER BACKGROUND AND STRUCTURE OF THIS PAPER

In this section, we explain the structure of our paper by reviewing the relevant prior work on the subject and relating it to the work presented in this paper.

In the quantum-information context, information–disturbance relations were inspired by the *epistemic* interpretation of quantum states (that is, by the states of knowledge of an agent, i.e., an observer; see Ref. [10] and references therein) as applied to quantum cryptography [2]. The idea was that information gathering and disturbance should be grounded with respect to an observer's state of knowledge, not some preexisting property of the system. To formalize this intuition, Fuchs and Jacobs in Ref. [2] considered something like the following game. Two observers, Alice and Bob, initially agree that the state of some quantum system is ρ_0, an impure state. At some later time, Charlie performs a measurement along a randomly chosen axis and asks Alice and Bob to predict the outcome of the measurement. After repeating this many times, Charlie rewards the individual who makes the most correct guesses. Alice, being clever and resourceful, has the ability to perform one unitary operation followed by a two-outcome measurement. Her aim is to refine her state of knowledge about the quantum system, thus increasing its predictability and accordingly the likelihood of receiving the reward. However, Alice's morals prevent her from intentionally sabotaging Bob for the pecuniary reward. Thus Alice would like to increase the accuracy of her own predictions without affecting those of Bob.

This formulation suggests (intuitively) that weak measurements are the right way to probe the system. From the point of view of the weak-measurement formalism, a projective measurement corresponds to a measurement of infinite strength [2]. For finite-strength measurements on a single qubit, a nontrivial tradeoff is found between Alice's information-gathering actions (the refinement of her predictability) and the disturbance to Bob's predictability (his state of knowledge).

For a fixed-measurement rate, the interval of the nontrivial tradeoff may be characterized by the angle between the Bloch vector and the measurement axis. There are two extremes of this tradeoff. Minimal disturbance and minimal *average* information gain occur when Alice's axis is aligned with the Bloch vector, that is, when the measurement eigenbasis and the state's eigenbasis commute. On the other hand, maximal disturbance and maximal average information gain

are attained when Alice's axis is orthogonal to the Bloch vector. This strategy corresponds to making the outcomes equally likely, that is, to maximizing the *surprisal* [1] of the measurement result. In this situation, the bases are said to be complementary, maximally noncommuting, or unbiased. Thus we conclude that quantum theory does not allow Alice to succeed. The ability to measure in different bases is an essential feature of quantum mechanics that provides richness to the problem considered here.

It is easy to understand how noncommuting measurements provide maximal information gain and maximal disturbance through the following picture. Consider a mixed qubit state with its Bloch vector aligned along the $+|z\rangle$ direction. If we measure in the $|x\rangle$ basis, the Bloch vector will be elongated regardless of the measurement outcome. This point is illustrated in Fig. 1(a). Furthermore, both outcomes ($M_\pm$ in Fig. 1(a)) are equally likely, which means that the *average* information extraction is maximal. Unfortunately, the postmeasurement state is not pointing in the same direction as the premeasurement state, so it will not commute with the original state. Conversely, if the measurement basis and the state are both along the $|z\rangle$ direction, then, due to the stochastic nature of the measurement outcomes, the average elongation of the Bloch vector will be minimal [6,11]. In any particular measurement, the impurity may increase or decrease, but the disturbance to the state will be minimal [2]. In Fig. 1(b), we illustrate that one of the possible measurement outcomes causes the Bloch vector to shrink while the other outcome elongates it.

While information–disturbance relations are interesting for single measurements, the subject has been thoroughly explored. Less exploration has been undertaken in continuous-measurement settings because of the increased difficulty of the analysis. To make progress, the focus has
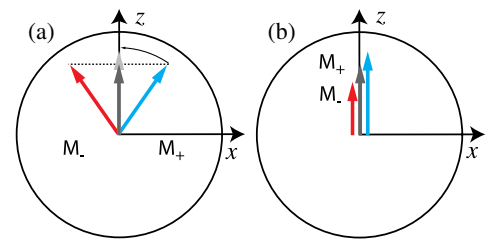

FIG. 1. A Bloch-sphere depiction of the extremes of information–disturbance tradeoffs. In both panels, the dark arrow represents the Bloch vector of the premeasurement state. The arrows labeled with $M_\pm$ correspond to the postmeasurement states conditioned on the random outcome of a weak measurement. (a) The measurement of a state in a complementary or unbiased basis; in this case, the measurement axis is along the x direction. When this occurs, the Bloch vector always elongates, as indicated by the light gray arrow (the plus outcome rotated for easy visual comparison). (b) The commuting measurement. In this case, if the measurement result M_- is obtained, the Bloch vector decreases in length.

mostly been on information gathering, while disturbance has largely been ignored (for good reason, as we shall see). This restriction has led to an interesting development in the theory of continuous quantum measurements [12–14].

To demonstrate an improvement in the extraction of information, the extraction rate of a standard continuous measurement must be defined. We follow Jacobs and define the standard measurement as a measurement in a commuting basis [4]. This is sensible for a number of reasons. First, the measurement will rapidly destroy the coherence of any state that does not commute with the measured observable [15]. This can be understood as the projection of the state to the eigenbasis of the observable by the measurement. To counter this effect, an adaptive measurement or closed-loop feedback control is required. Second, a continuous measurement in a commuting eigenbasis is essentially a classical continuous measurement.

The analysis of disturbance and information gain is significantly harder for continuous measurements of D-dimensional systems [8,16]. Minimal disturbance still occurs when measuring in the same eigenbasis as the state (a commuting basis). A full analysis of this situation would require characterization of the distribution of information-extraction rates, allowing calculation of the average information extraction, median, minimal case, and maximal case. In Sec. III, we derive the average purification rate for a commuting measurement of a D-dimensional system and provide an asymptotic expression. In the Appendix, we rigorously derive these expressions. Additionally, in Appendix A 4, we give exact results for bounds on the spread of the distribution of purities, which were heuristically derived in a previous work [16].

Now that we have defined what a standard measurement is, we can explore what it means to increase the rate of extraction of information. In the context of control theory, one might be interested in revealing or acquiring the value of some parameter; that is the domain of adaptive measurements [14,17]. But more typical control objectives are to drive the evolution of the state or to stabilize the state against noise [14,16,18–20]. Thus, disturbing the prior state is not a problem. This is why disturbance is not usually considered in feedback control. What becomes important is the noise introduced by the measurement to the system. Doherty, Jacobs, and Jungman in Ref. [3] showed that, for a continuously monitored qubit, the measurement noise is maximal when the eigenbases of the state and measurement axis are unbiased.

The study of Wiseman, Mancini, and Wang in Ref. [20] explored state stabilization in a qubit. The result relevant to the present discussion from Ref. [20] relates to the difficulty of stabilizing a state that is complementary to the measurement basis. They showed that stabilization can be achieved in the limit of infinite-strength ''Bayesian'' feedback, but not in simpler ''Markovian'' feedback procedures. This is a consequence of the increased measurement noise as predicted by Doherty *et al.* [3]. Later, Jacobs showed how continuously monitoring the state of a qubit in a complementary basis (assuming infinite-strength feedback) leads to a doubling rate of information extraction over monitoring in a commuting basis asymptotically [4].

Recently, a number of studies have examined the use of weak (and continuous) measurement in a complementary basis in systems of dimension D greater than two. The first study showed that the lower bound on the improvement in the rate of information extraction when measuring in any complementary basis is $(2/3)(D + 1)$ [8]. The second study showed that this speeding-up is bounded from above by a factor $D^2/2$ [9]. We summarize these results in Sec. IV A of this paper. In Secs. IV B, IV C, and IV D, we construct explicitly, for an arbitrary finite-dimensional system, an unbiased-basis protocol based on the quantum Fourier transform (QFT). We derive upper and lower bounds on the speeding-up of the QFT protocol. Surprisingly, the numerics show that the protocol saturates the upper bound—when D is even. For completeness, we mention the recent work of Ruskov *et al.* [21], which showed monitoring a qubit in three complementary bases and which claimed a speeding-up of a factor of approximately 3. We are currently exploring this relationship, so we will not discuss it further in this paper.

Motivated by some of the above investigations, Jacobs and Lund argued that the task of state stabilization in the regime of strong feedback is best accomplished by measurement in a complementary basis [19]. Foundational issues aside, this is one reason to continue investigations into maximal information extraction. In Ref. [19], the authors raised two interesting questions: First, are all complementary bases equally good for information extraction when $D \geq 4$? Second, is measurement in a complementary basis best for information extraction?

We reexamine these questions in Secs. V and VI. We show that complementarity of the measurement basis, i.e., the equal-likelihood condition, is not sufficient to guarantee maximum information extraction. Our analysis provides an insight into the mechanism for information extraction, which is quantum mechanical backaction. From this analysis, we explain how maximum information extraction is achieved by enforcing the complementary condition and arranging the eigenvalues in phase space so that measurement backaction can have maximum effect. We provide further evidence for the claim that not all bases are equally good. Because Jacobs and Lund did not examine permutations of the unbiased basis, it was still unclear from their analysis whether all unbiased bases are equally powerful for information extraction. So, in Sec. VI, we also consider the role of permutations. Our results show that, even after maximizing information extraction over permutations of the unbiased basis, not all unbiased bases are equally good at information extraction.

In Sec. VII, we examine complementary measurements in a register of qubits using the QFT. We start our analysis

in Sec. VII A, with the case of commuting measurements [9]. Then, in Sec. VII B, we review the results of Ref. [9], where it was shown that the speeding-up bounded by $2n(2^n - 1)^{-1} \leq S \leq 2n$. In Sec. VII C, we construct a protocol based on the QFT and present results which suggest that the speeding-up is at most 2. We conclude in Sec. VIII.

III. COMMUTING MEASUREMENTS

Later in this paper, we will investigate strategies that increase the rate at which information can be extracted from quantum systems. Before these strategies are examined, however, we must fully characterize this rate in the absence of feedback to provide a benchmark for assessing the performance of the strategies that aim to increase the rate of information extraction.

In this section, we restrict our attention to a *conditional* continuous measurement of a single D-dimensional quantum system, also called a qudit, monitored by a single output channel. This output channel is monitored in such a way that it results in a diffusive quantum trajectory [12,14]. This monitoring can be thought of as a sequence of short-duration, weak measurements. Following Fuchs and Jacobs [2], our notion of information is characterized by the impurity: $L[\rho(t)] \equiv 1 - \text{Tr}[\rho(t)^2]$. Accordingly, we derive an exact expression for the average impurity $\langle L(t) \rangle$, which must be integrated numerically. Then we derive a simpler approximate expression, which will be used in the remainder of this paper's analysis.

In the Appendix, we give a detailed derivation of these equations and a detailed discussion of the approximations that we use here in Sec. III. There is much new work in the Appendix, so we will briefly summarize it now. In Appendix A 1, we give a detailed derivation and analysis of the exact and approximate expressions for the average conditional impurity. In Appendixes A 2 and A 3, we examine the validity of our approximate expression through numerics. To completely characterize the rate of information extraction from commuting measurements, we look at bounding the distribution of purities in Appendix A 4. Finally, we summarize the key messages in Appendix A 5.

A. Qudit: Impurity analysis

The starting point for the current analysis is similar to that presented in Refs. [8,9,22]. We use a widely applicable model that describes the evolution of the state ρ of a quantum system due to repeated weak measurements of a dimensionless observable X. This model is described by the stochastic master equation (SME) [14,23–25],

$$d\rho = dt\mathcal{L}_c[X]\rho \equiv 2\gamma dt\mathcal{D}[X]\rho + \sqrt{2\gamma}dW\mathcal{H}[X]\rho, \quad (1)$$

where $\mathcal{D}[A]\rho \equiv A\rho A^\dagger - \frac{1}{2}(A^\dagger A\rho + \rho A^\dagger A)$, $\mathcal{H}[A]\rho \equiv A\rho + \rho A^\dagger - \text{Tr}[(A^\dagger + A)\rho]\rho$ [14], and dW is the increment of a Wiener noise process [26]. It should be noted that

we have moved to a frame that has enabled us to factor out the Hamiltonian evolution. The *measurement rate*, γ, determines the rate at which information is extracted, and thus the rate at which the system is projected onto a single eigenstate of X [16,27,28]. (This means that, for times $\tau \gg \gamma^{-1}$, we may say that we have performed a projective measurement of the observable X.) The measurement result in a small time interval $[t, t + dt)$ is denoted

$$dR(t) = \sqrt{4\gamma}\langle X(t) \rangle dt + dW(t), \quad (2)$$

where dW is the same Wiener noise process that appears in Eq. (1) and $\langle X(t) \rangle = \text{Tr}[X\rho(t)]$. A quantum trajectory is a continuous-in-time description of the state conditioned on the measurement result. Thus, to specify a trajectory over the interval $[0, T)$, we must be given the measurement results over that interval. Then the Itō definition $\rho(t + dt) = \rho(t) + d\rho(t)$ may be used recursively to generate a trajectory. Clearly, a single weak measurement is described by one application of this relation.

We will denote the continuous-measurement record obtained by the observer integrated up until time t as

$$R(t) = \int_0^t dR(t'). \quad (3)$$

The unnormalized version of Eq. (1) is known as the *linear SME*. We denote this linear SME as [14,29–31]

$$d\bar{\rho} = dt\bar{\mathcal{L}}_c[X]\bar{\rho} \equiv 2\gamma dt\mathcal{D}[X]\bar{\rho} + \sqrt{2\gamma}dR\bar{\mathcal{H}}[X]\bar{\rho}. \quad (4)$$

Here, $\bar{\mathcal{H}}[A]\rho \equiv A\rho + \rho A^\dagger$, and the bar over ρ denotes the lack of normalization at all but the initial time. We take the initial state to be $\bar{\rho} = \rho(0) = \mathbf{I}/D$, where $\mathbf{I}$ is the $D \times D$ identity matrix. Because the initial-state matrix commutes with X at all times, this makes obtaining the solution to Eq. (4) simple. This solution is [24]

$$\bar{\rho}(R, t) = \exp(-4\gamma X^2 t) \exp[2\sqrt{2\gamma}XR(t)]\mathbf{I}/D. \quad (5)$$

This state is normalized by $\mathcal{N}(R, t) = \text{Tr}[\bar{\rho}(R, t)]$, so that $\rho = \bar{\rho}/\mathcal{N}$. Thus the impurity may be written as

$$\Lambda(R, t) = 1 - \text{Tr}[\rho(R, t)^2]. \quad (6)$$

The impurity averaged over all possible measurement records R is given by

$$\langle L(t) \rangle = \int_{-\infty}^{\infty} \Lambda(t, R)\mathcal{P}(R, t)dR, \quad (7)$$

where $\mathcal{P}(R)$ is the probability to obtain the state $\rho(R, t)$ at time t. (This is explained in great detail in Appendix A 1.) To calculate this integral, one must perform numerical integration. It would be useful to have a fully analytic expression for $\langle L(t) \rangle$. Now we will make some approximations to obtain a simple analytic expression for Eq. (7).

The approximate expression is arrived at by noting that, in the long-time limit, $\Lambda(t, R)\mathcal{P}(R, t)$ is sharply peaked in $(D - 1)$ places, which means that the integral in Eq. (7) can

be simplified into a sum of $(D-1)$ integrals. At long times, these integrals are essentially the same; thus the integral is now $(D-1)$ times the integral for one region:

$$(D-1)\left[\frac{e^{-\gamma t}}{D\sqrt{8\pi t}}\int_M \frac{e^{-(R^2/2t)}}{\cosh(\sqrt{2\gamma}R)}dR\right],\tag{8}$$

where $M\in[-\infty,\infty]$ is the region of integration. Furthermore, the distribution in the numerator of Eq. (8) is broad compared to the distribution in the denominator for $t\gg\gamma^{-1}$. Thus we can make the approximation:

$$\int_M dRe^{-R^2/2t}/\cosh(\sqrt{2\gamma}R)\approx\int_M dR/\cosh(\sqrt{2\gamma}R)$$

$$=\pi/\sqrt{2\gamma}.$$

The final expression for the impurity is thus

$$\langle L_2(t)\rangle_{LT}=\frac{2(D-1)}{D}\frac{\pi e^{-\gamma t}}{\sqrt{16\gamma t\pi}},\tag{9}$$

where the subscript 2 denotes the particular approximation used to derive this expression (see Appendix A 1 for details) and the subscript LT denotes that it is only valid in the long-time limit. This is the analytic expression for $\langle L(t)\rangle$ that we had set out to find, and it is the key result of this section.

IV. INFORMATION ACQUISITION IN DIMENSION D AND ITS RELATIONSHIP TO COMPLEMENTARITY

A. Previous results

In this subsection, we briefly summarize the results of previous studies on information acquisition by measurement in a complementary basis. The original articles on the subject are found in Refs. [2,4,8,9]. The basic idea is to use quantum feedback control to continuously keep the eigenbasis of the state complementary to the measurement basis. By analogy with the intuition presented in Fig. 1, this procedure is expected to enhance information extraction.

From Eq. (1), it is easy to show using Itō calculus that the information acquired [2], as characterized by the change in impurity, about the state due to measurement of infinitesimal duration is [9]

$$dL=-8\gamma dt(\text{Tr}[\rho X\rho X]-2\,\text{Tr}[X\rho]\text{Tr}[X\rho^2]$$
$$+\text{Tr}[\rho X]^2\text{Tr}[\rho^2])-4\sqrt{2\gamma}dw\{\text{Tr}[X\rho^2]$$
$$-\text{Tr}[\rho X]\text{Tr}[\rho^2]\}.\tag{10}$$

We wish to study the effect of measuring in a complementary basis. In order to measure in a complementary basis throughout the measurement process, one must use quantum feedback control [14] to continually adjust the basis. We will not labor on those details here, as they have been adequately discussed before [4,8,9,14]. The required transformation of the measurement basis is $X(t)\to\check{X}(t)$, such

that $|\langle\check{x}|i\rangle|=1/\sqrt{D}$, for all $\check{x}$, and i where $|i\rangle$ is an eigenstate of ρ and $|\check{x}\rangle$ is an eigenstate of $\check{X}$. Replacing all the X's by $\check{X}$ in Eq. (10) considerably simplifies the expression to [8,9]

$$dL=-8\gamma\text{Tr}[\check{X}\rho\check{X}\rho]dt,\tag{11}$$

or equivalently

$$dL=-8\gamma dt\sum_{r,c}|\check{X}_{r,c}|^2\lambda_r\lambda_c.\tag{12}$$

Observe that the unbiased condition that simplified Eq. (10) to Eq. (11) leaves a permutational degree of freedom in Eqs. (11) and (12) because a permutation of the eigenstates of ρ does not affect $|\langle\check{x}|i\rangle|=1/\sqrt{D}$. This means that we should optimize over permutations in the expression $dL=-8\gamma\text{Tr}[\check{X}P_m^\dagger\rho P_m\check{X}P_m^\dagger\rho P_m]dt$, where P_m is a permutation matrix, to maximize the decrease in impurity. To obtain a lower bound $|dL|$ of the optimal permutation, we average over all possible permutations: $dL=-8\gamma\sum_m\text{Tr}[\check{X}P_m^\dagger\rho P_m\check{X}P_m^\dagger\rho P_m]dt$. The average gives a lower bound because the sum of a sequence is always less than or equal to the greatest term in the sum.

It was shown in Refs. [8,9] that a tight lower bound on Eq. (11), for the optimal permutation, is

$$dL\leq-\frac{2}{3}\gamma t(D+1)L(t),\tag{13}$$

while the upper bound was recently found to be

$$dL\geq-\frac{1}{2}\gamma tD^2L(t)\tag{14}$$

for $D\gg1$ [9]. Eqs. (13) and (14) hold for measurement in any complementary basis. Using these bounds, one may compare the time that it takes for a commuting measurement [i.e., Eq. (9)] to extract a certain amount of information (i.e., to attain $\langle L\rangle=\epsilon$) to how long a complementary measurement may take to extract that same amount of information. That is, we equate Eq. (9) and the solutions of Eqs. (13) and (14), and solve for the ratio of $t_{\text{complementary}}$ and t_{commute}. We call this ratio the speeding-up in information acquisition. For $t_{\text{commute}}\gg\gamma^{-1}$, the speeding-up is bounded by

$$\frac{2}{3}(D+1)\leq S\leq\frac{D^2}{2}\tag{15}$$

for $D\gg1$ [9].

While these results show the usefulness of feedback in rapid purification, some important issues have not yet been addressed. First, does an explicit construction of such a complementary-measurement strategy exist? Second, is actual performance of such a strategy $O(D)$ or $O(D^2)$? Third, what is the mechanism of the speeding-up? The first and second questions are answered in Secs. IV B and IV C respectively. We defer the third question until Sec. V.

B. Complementarity via quantum Fourier transform for D=2, 3, 4

In this subsection, we explore implementing a continuous-complementary measurement in systems of dimension three and four. There are many unbiased bases, so to make our analysis concrete we must choose one. It has been noted elsewhere [19] that, for $D > 3$, not all unbiased bases are equally good at reducing the impurity. Nevertheless, we choose a particular complementary basis, the one generated by the quantum Fourier transform of the logical basis. The construction of the continuous-complementary measurement then consists of: (a) calculationally diagonalizing the state matrix; (b) calculationally ordering (permuting) those eigenvalues to maximize $|\langle dL \rangle|$ after a measurement; and (c) applying the appropriate permutation and the QFT unitary to the state. This whole procedure is an example of quantum feedback control [14]. The D-dimensional QFT can be represented by this matrix:

$$
T = \frac{1}{\sqrt{D}}
\begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
1 & q & q^2 & \cdots & q^{1(D-1)} \\
1 & q^2 & q^4 & \cdots & q^{2(D-1)} \\
\vdots & \vdots & \vdots & \ddots & \\
1 & q^{(D-1)} & q^{2(D-1)} & & q^{(D-1)^2}
\end{pmatrix}, \quad (16)
$$

where $q = \exp(2\pi i/D)$. The matrix elements of T are

$$
T_{rc} = \frac{1}{\sqrt{D}} \exp\left[\frac{2\pi i}{D}(r \times c)\right], \quad (17)
$$

where $r, c \in (0, \ldots, D-1)$.

For $D = 2$, Jacobs' protocol [4] turns out to be equivalent to the QFT protocol. The transformation unitaries in both cases are

$$
T_{\mathrm{Jac}} = \exp\left(i\frac{\pi}{2}J_y\right) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad (18)
$$

$$
T_{\mathrm{QFT}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (19)
$$

The effect of these transformations on dL is identical: $\check{X} = T_{\mathrm{Jac}} J_z T_{\mathrm{Jac}}^\dagger = T_{\mathrm{QFT}} J_z T_{\mathrm{QFT}}^\dagger = J_x$. Because this scheme is equivalent to Jacobs', we will not analyze this $D = 2$ case any further.

For a system of dimension $D = 3$, the transformed measurement is

$$
\check{X} = TJ_zT^\dagger = \frac{1}{3}\begin{pmatrix} 0 & 1-q & 1-q^* \\ 1-q^* & 0 & 1-q \\ 1-q & 1-q^* & 0 \end{pmatrix}, \quad (20)
$$

where $q = \exp(2\pi i/3)$. Taking the state to be $\rho = \mathrm{diag}(\lambda_0, \lambda_1, \lambda_2)$, where $\lambda_r > \lambda_c$ when $r < c$, and solving the equation for dL, using the fact that $L = \sum_{r \neq c, c \neq r} \lambda_r \lambda_c$, we find
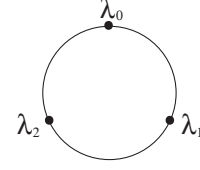

FIG. 2. The only possible permutation for $D = 3$. The solid lines between the eigenvalues represent the weighting of the multiplying factors in Eq. (21). That is, $\frac{2}{3} = |\check{X}_{r,c}|^2$, $\forall r, c$.

$$
\begin{aligned}
dL &= -8\gamma dt \sum_{r,c} |\check{X}_{r,c}|^2 \lambda_r \lambda_c \\
&= -8\gamma dt\left[\frac{2}{3}(\lambda_0\lambda_1 + \lambda_0\lambda_2 + \lambda_1\lambda_2)\right] \quad (21)
\end{aligned}
$$

$$
= -8\gamma dt \frac{1}{3}L(t), \quad (22)
$$

which coincides with the bound of Eq. (13). The above expression can be easily integrated to give

$$
L(t) = e^{-(8/3)\gamma t}L(0). \quad (23)
$$

In Fig. 2, the weights $|\check{X}_{r,c}|^2$ between the $\lambda_r\lambda_c$'s are plotted. The weights in this case are all equal, so it is obvious that permuting the eigenvalues will not change the decrease of impurity. This explains why we did not have to find an optimal permutation.

Now we compare the above calculation to numerics. The average impurity for the continuous-complementary-measurement protocol and the continuous-commuting-measurement protocol are calculated by stochastic simulations and Eq. (7), respectively. Figure 3 shows that
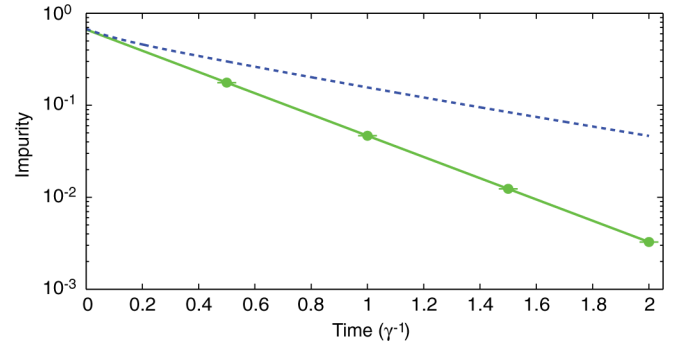

FIG. 3. Trajectory simulations of the impurity for a $D = 3$ system subject to feedback every $\delta t = 10^{-3}\gamma^{-1}$ in order to keep the state and measurement eigenbases complementary. The ensemble size is 100. The dashed line is $\langle L \rangle$ for a continuous-commuting measurement ($D = 3$), i.e., Eq. (7). The solid line is the numerically calculated ensemble average of $\langle L \rangle$ for the complementary measurement; it is plotted in green, and error bars are plotted at times $(0.5, 1, 1.5, 2) \times \gamma^{-1}$. The analytic expression for the average impurity, Eq. (23), is indistinguishable from the results of the simulation.

the stochastic simulations of the complementary-measurement protocol give a deterministically decreasing $\langle L \rangle$, as predicted by Eq. (23).

We numerically investigate the effect of varying the frequency, $1/\delta t$, of the operation applied (feedback) to keep the state and measurement eigenbases complementary in Fig. 4. The stochastic fluctuations in the simulations arise from the finite size of δt in the simulations. This effect can also be found in simulations of the Jacobs' qubit feedback protocol [32]. As $\delta t \to 0$, the numerically calculated average impurity for complementary measurement approaches Eq. (23). For $\delta t \ll 10^{-3}\gamma^{-1}$, numerical simulations are indistinguishable from the analytically calculated (and then interpolated) speeding-up, which is the solid line in Fig. 4.

We now consider complementary measurement using the QFT for the case when $D = 4$. The increment for the impurity is

$$
\begin{aligned}
dL &= -8\gamma dt \sum_{r,c} |\check{X}_{r,c}|^2 \lambda_r \lambda_c \\
&= -8\gamma dt \Big(\lambda_0 \lambda_1 + \frac{1}{2}\lambda_0 \lambda_2 + \frac{1}{2}\lambda_1 \lambda_3 + \lambda_0 \lambda_3 \\
&\quad + \lambda_2 \lambda_3 + \lambda_1 \lambda_2 \Big).
\end{aligned}
\tag{24}
$$

Unfortunately, it is not possible to factor out an expression for the impurity on the right-hand side of the above equation. It is obvious from Eq. (24) that permuting the eigenvalues will affect $|dL|$. Our task is to maximize the decrease in dL. From Eq. (24) it is possible to intuit the form of the best and worst permutations, which are depicted in Fig. 5. An optimal permutation is $\rho_{\text{opt}} = \text{diag}(\lambda_0, \lambda_1, \lambda_3, \lambda_2)$. This permutation is not uniquely optimal; the permutation $\text{diag}(\lambda_1, \lambda_0, \lambda_2, \lambda_3)$ is also optimal.

FIG. 5. (a) The weights $|X_{r,c}|^2$ in Eq. (24) after diagonalization and in descending ordering. (b) The optimal permutation for $D = 4$. (c) The worst permutation for $D = 4$. The solid line represents the maximum weighting of multiplication in Eq. (24); the dashed lines represent minimum weights corresponding to $\frac{1}{2}$.

One of the worst permutations is $\rho_{\text{worst}} = \text{diag}(\lambda_0, \lambda_3, \lambda_1, \lambda_2)$. Unfortunately, knowing the optimal permutation does not help us simplify the expression. Because we cannot solve this case analytically, we invoke the procedure developed in Refs. [5,9] to find the bounds on $|\langle dL \rangle|$ (and hence S) for the QFT protocol for all D. These bounds are calculated in Sec. IV C.

Figure 6 depicts the numerically calculated speeding-up for $D = 4$. The speeding-up is larger than that of the $D = 3$ case [33]. Furthermore, for reasonably frequent feedback, the numerically calculated speeding-up lies between the analytical bounds predicted by Eqs. (13) and (14).

We also briefly compare the complementary-measurement protocol to the rapid-measurement (RM) algorithm of Ref. [5] in Fig. 7. The decrease in the ensemble average of the impurity seems to be of the same order. However, the trajectories in the RM case have a large variance. This is because the trajectories are not differentiable. By comparison, even for a finite δt the L in the complementary measurement is very close to *deterministic*. Although L is still stochastic, it is also differentiable for $\delta t \to dt$ so the noise in the simulations is reduced.
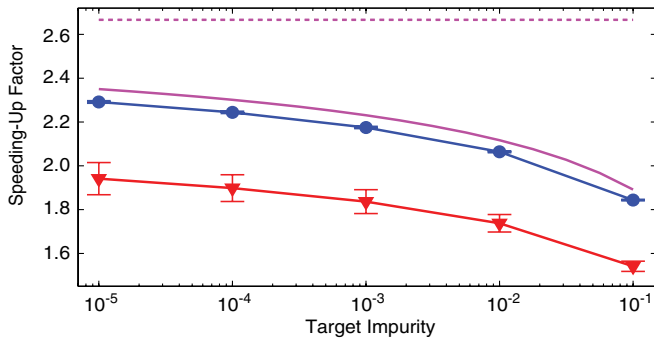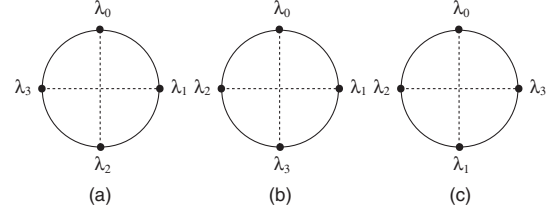

FIG. 4. Short-time numerically calculated speeding-up for $D = 3$. The dashed line is the analytically calculated asymptotic speeding-up $S = 8/3 = 2.\dot{6}$. The solid magenta line is the ideal speeding-up, calculated numerically from Eqs. (7) and (23). The circles and triangles are numerical calculations of the speeding-up with finite $\delta t = 10^{-3}\gamma^{-1}$ and $10^{-2}\gamma^{-1}$, respectively; the ensemble size is 100. As $\delta t \to 0$, the numerically calculated speeding-up approaches the solid line and the error bars vanish.
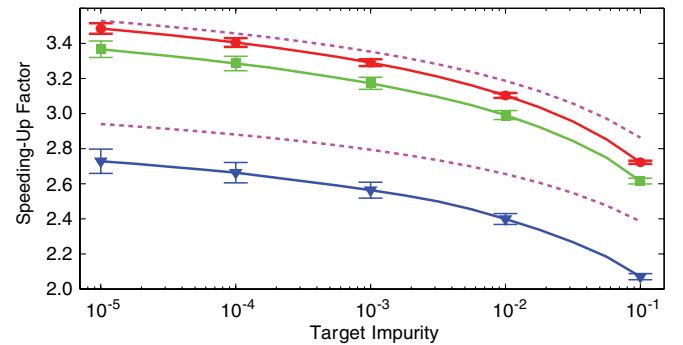

FIG. 6. Short-time numerically calculated speeding-up for $D = 4$. The dashed lines are the bounds on the speeding-up given by Eqs. (27) and (31). The circles, squares, and triangles are numerical calculations of the speeding-up with finite $\delta t = 10^{-3}\gamma^{-1}$, $10^{-2}\gamma^{-1}$, $10^{-1}\gamma^{-1}$, respectively. In order to keep the error bars roughly the same size, the ensemble size was varied. The ensemble sizes for circles, squares, and triangles are 100, 400, and 6400, respectively.
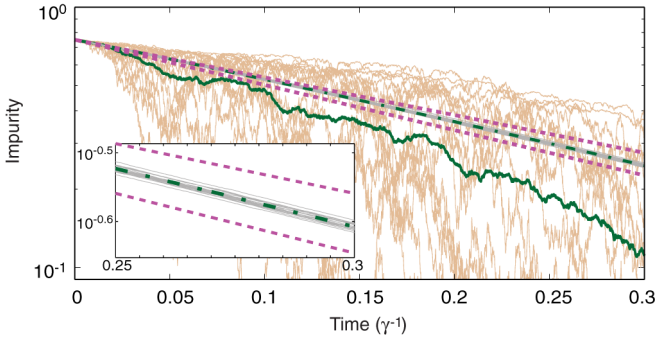
FIG. 7. A comparison between the complementary-measurement protocol and RM feedback (of Ref. [5]) for $D = 4$, $\delta t = 10^{-4}\gamma^{-1}$, and an ensemble of 20 trajectories in each case. The dashed lines are the analytically calculated bounds on dL for complementary feedback [Eqs. (27) and (30)]. The dash-dotted green line (inset) is the ensemble average for L with the QFT feedback. The solid dark green line is the ensemble average for the RM feedback. The light brown lines are the trajectories for the RM feedback. The analytical bounds suggest that both feedback strategies have the same order of speeding-up. Inset: All of the complementary-measurement trajectories (gray lines) lie within the predicted bounds. The trajectories are very nearly deterministic.

C. Bounds on the QFT complementary-measurement protocol for all D

1. Lower bound

From Eq. (15) it is clear that the lower bound on the optimal speeding-up for the measurement of any complementary basis is $S = \frac{2}{3}(D + 1)$. Nevertheless, it is interesting to work this out explicitly for a particular complementary basis. The method we use to find the lower bound was first presented in Ref. [9]. This method uses a fictitious state ρ_F with the same purity as the state under consideration, i.e., $L[\rho] = L[\rho_F]$. The fictitious state ρ_F is defined as $\rho_F = \mathrm{diag}(1 - \Delta, \delta, ..., \delta)$, where $1 - \Delta$ is the largest eigenvalue of ρ_F and $\delta = \Delta/(D - 1)$. The equation for dL is invariant under permutations of the eigenvalues ρ_F. Substituting ρ_F into the right-hand side of Eq. (12) gives $2(1 - \Delta)\frac{\Delta}{(D-1)}\sum_{r \neq 1}|\check{X}_{r1}|^2 + \frac{\Delta^2}{(D-1)^2} \times \sum_{r \neq 1c, c \neq r}|\check{X}_{rc}|^2$. This expression is further simplified by explicitly calculating the matrix elements. For the QFT, we have

$$\sum_{r \neq 1}|\check{X}_{r1}|^2 = \sum_{m=1}^{D-1}\frac{1}{4\sin^2(\pi m/D)} = (D^2 - 1)/12, \quad (25)$$

using the identity $\sum_{m=1}^{D-1}\csc^2(\pi m/D) = (D^2 - 1)/3$ [34]. Also, we have

$$\sum_{r \neq 1, c \neq r}|\check{X}_{r,c}|^2 = \frac{D + 1}{12}(D - 1)(D - 2), \quad (26)$$

using the above identity and the fact that $\csc^2[\pi(D-1)/D] = \csc^2(\pi/D)$. After some simplification,

the final expression is an upper bound on the decrease in impurity:

$$dL_F \leq -\frac{2}{3}(D + 1)\gamma dt L(t), \quad (27)$$

which implies that the lower bound on the asymptotic speeding-up is

$$S_F \geq \frac{2}{3}(D + 1), \quad (28)$$

as found in Refs. [8,9] for a general complementary measurement.

2. Upper bound

To find the upper bound on the speeding-up for the QFT protocol we again use the method of Ref. [9]. This instance requires a different fictitious state ρ_2, again with $L[\rho] = L[\rho_2]$, which was termed the binary distribution. The binary distribution is defined as $\rho_2 = \mathrm{diag}(1 - \Delta', \Delta', 0, ..., 0)$, where $1 - \Delta'$ is the largest eigenvalue. It is known that dL is most sensitive to permutations of the eigenvalues of ρ_2 [9].

Substituting ρ_2 into Eq. (12) gives only two terms. Now there is a choice about where to situate the two eigenvalues so that dL is maximized. By calculating the element $|\check{X}_{r,c}|^2$, using the QFT, we find that

$$|\check{X}_{r,c}|^2 = \frac{1}{2[1 - \cos\frac{2\pi}{D}(r - c)]}. \quad (29)$$

This allows us to optimize the change in impurity over the permutations by finding $\max|P_m^\dagger \check{X}_{r,c} P_m|^2$. The largest element of the matrix is $|\check{X}_{01}|^2 = 1/2(1 - \cos\frac{2\pi}{D})$. From this, it is clear that dL will be maximized provided the two eigenvalues from ρ_2 are in succession. The change in the impurity for the binary distribution is

$$dL_2 = \frac{-8\gamma dt(2\lambda_0\lambda_1)}{2(1 - \cos\frac{2\pi}{D})} = \frac{-4\gamma dt}{(1 - \cos\frac{2\pi}{D})}L_2(t). \quad (30)$$

Thus, the asymptotic upper bound for the speeding-up is

$$S_2 \leq \frac{4}{1 - \cos\frac{2\pi}{D}}. \quad (31)$$

For $D \gg 1$, one finds

$$S_2 = \frac{2D^2}{\pi^2} \approx 0.2D^2. \quad (32)$$

In Fig. 8 we depict the conjectured optimal and worst permutations of eigenvalues of a D-dimensional state ρ. The optimal permutation is $\rho = \mathrm{diag}(\lambda_0, \lambda_1, \lambda_3, \lambda_5, ..., \lambda_6, \lambda_4, \lambda_2)$.

In Fig. 9 the numerically calculated asymptotic speeding-up as a function of D is plotted. As expected, it is within the lower bound of Eq. (28) and the upper bound of Eq. (31). Further, it confirms a very nearly quadratic speeding-up: the fit shown is $S = 0.189D^2 + 0.109D + 0.248$. If, however, we choose to fit only the quadratic term, the fit is
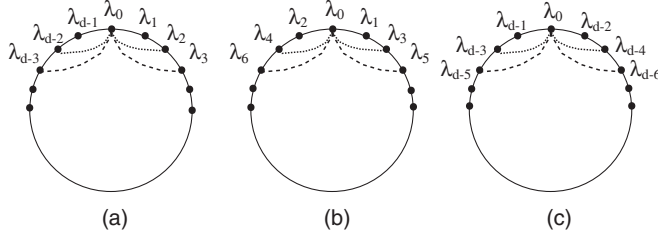
$$S = 0.19D^2. \quad (33)$$

FIG. 8. (a) Weighting factors in Eq. (12) after the feedback has diagonalized and ordered the eigenvalues in ρ. (b) The conjectured general optimal permutation. (c) The conjectured worst permutation. The weights are shown only from the λ_0 perspective.
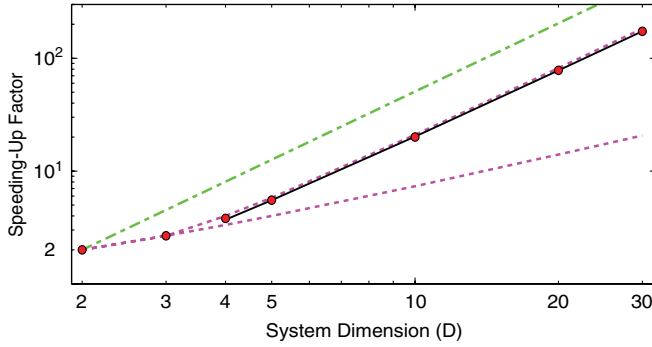

FIG. 9. The asymptotic speeding-up for reaching a given level of purity as a function of the system dimension. The solid black line indicates the fit in Eq. (33). The dashed magenta lines are the upper, Eq. (31), and lower, Eq. (28), bounds derived in the text for the QFT protocol. The dash-dotted green line is the upper bound on all unbiased-basis purification strategies, which is shown on the right-hand side of Eq. (15).

For $D \geq 10$ this is a good approximation and is very close to Eq. (32), which was derived for long times and large D. The numerically calculated asymptotic speeding-up is only a multiplicative constant away from the ultimate upper bound on all complementary-basis purification protocols, the dashed line ($S = 0.5D^2$), i.e., Eq. (15).

D. Discussion

The results presented in Secs. IV A, IV B, and IV C raise a number of questions which we now elucidate further.

Because of symmetries, for dimensions two and three, permuting the basis of ρ was unnecessary to maximize dL. The intuitive thinking is that, for maximizing information gain, the complementary measurement marginalizes the significance of which eigenvalue is the largest. For $D > 3$, we must permute the basis of ρ, after diagonalization and before applying T_{QFT}, to maximize dL. The permutations applied do not increase the signal-to-noise ratio of the measurement, unlike the rapid-measurement protocol of Ref. [5]. In Secs. IV B and IV C 2 we argued that the permutations should be chosen to maximize the product of the two largest eigenvalues in Eq. (12). While this recipe

is operationally sound, it lacks physical insight. For example, this recipe does not explain what the mechanism of the purification is and why the permutations can dramatically affect the rate of purification. These questions are addressed in Sec. V.

The final open question concerns the efficacy of different complementary bases to change the rate of purification. In a previous work, Lund and Jacobs suggested that not all complementary bases were equally good at entropy reduction [19]. That is, they claimed that one particular unbiased basis can reduce entropy more quickly than another. If this were true, then the $S = 0.2D^2$ upper bound we derived from the QFT basis would not necessarily be the true upper bound, which could be closer to $S = 0.5D^2$. We numerically investigate this question in Sec. VI.

V. WHY COMPLEMENTARITY IS NOT ENOUGH

How is it that rotating to an unbiased basis provides any speeding-up? Naïvely following the reasoning applied to qubits in Refs. [2–4] suggests that, by making each outcome equally likely (i.e., maximizing the surprisal [1]), one maximizes the average amount of information a measurement extracts. From this one might argue that the speeding-up observed can be explained by the D-dimensional version of this argument. However, it is not clear how this argument would explain why the arrangement of eigenvalues in ρ (permutations) is important to attain the best speeding-up. To address this, we examine a phase-space representation of the optimal and worst permutations that are schematically depicted in Fig. 8.

Defining a phase-space picture for a discrete variable is not trivial. In Ref. [35] it was shown that a spin Wigner function $W(\theta, \phi)$ can be defined in terms of Clebsch-Gordan coefficients and spherical harmonic functions. This spin Wigner function is a pseudoprobability distribution on the Bloch sphere, with θ and ϕ the usual Euler angles. The spin Wigner function is a little counterintuitive; for example, unlike the original Wigner function $W(x, p)$, for position and momentum, the marginal distribution for ϕ is not the true phase distribution $P(\phi)$. However, for large D, the marginals are a good approximation to the true phase distribution [36]. We plot the Wigner function using the equal-area projection (described by coordinates ϕ and $J \cos\theta$).

The conjectured best and worst permutations were depicted on a ring in Fig. 8 and explained in terms of the weighting factors $|X_{r,c}|^2$. We now have an intuitive understanding of the angle around this ring as being the phase ϕ in the angular representation of spin states. The eigenstates of ρ in the measurement (J_z) basis are $|r\rangle$ (Dicke states [37]). These are transformed by the QFT to the states $|\phi_r\rangle$, where $\phi_r := \frac{2\pi}{D}(J - r)$. The states $|\phi_r\rangle = \frac{1}{\sqrt{D}}\sum_{m=-J}^{J} \times \exp(-im\phi_r)|J, m\rangle$ are equivalent to the Pegg-Barnett phase states [38,39]. If the QFT were an easy operation in some physical system, then our protocol would be a procedure for

rapidly preparing a Pegg-Barnett phase state. We note that in some spin systems it should be possible to construct the desired unitary [40].

Consider the long-time-limit state. The worst case, for purification purposes, is when the eigenvalues are equal. Under this two-eigenvalue approximation, the worst permutation corresponds to putting the second largest eigenvalue, λ_1, the maximal distance away from λ_0 in phase space ($\phi = \pm\pi$). The spin Wigner function for this configuration of the mixture is plotted in Fig. 10(a) for $D = 10$. The optimal permutation is when the two largest eigenvalues are next to each other in the phase space; see Fig. 10(b). It is now apparent that the schematic diagrams in Fig. 8 represent a slice through the unwrapped Bloch sphere, and the positions of the eigenvalues are their arrangements of the phase states corresponding to the original eigenvalues. This phase-space picture suggests that it might be possible to explain why these are the best and worst cases and how the speeding-up is generated. To explain these two features, we move away from exact calculation of Wigner functions and move to a schematic representation of the Wigner function.

In Fig. 11(a) we represent schematically the bulk of the Wigner function by a rectangle of width $2\pi/D$ (corresponding to the ϕ coordinate) and height $D - 1 = 2J$ (which corresponds to the $J\cos\theta$ coordinate). In what follows, on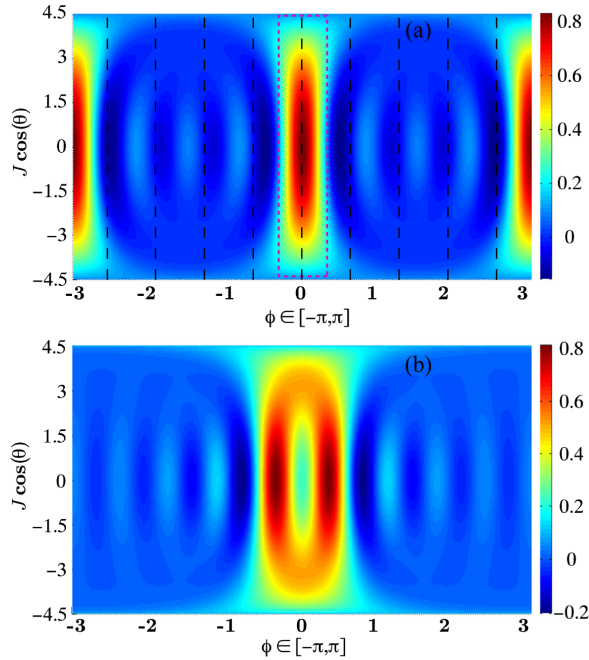e may loosely think of the rectangle as representing the uncertainties in an observer's knowledge about two conjugate variables; according to the Heisenberg uncertainty relation, the area of the rectangle must be constant for pure states. Now, consider the effect of a weak measurement of J_z. In our protocol this is of infinitesimal duration, but here we exaggerate the effect to illustrate our point. In Fig. 11(b) we have taken the result of the measurement to be positive. Because the result is positive, the Wigner function does not have much support on the lower part of the plot. As the positive result contains information about the distribution of j, the uncertainty in this variable is reduced. In keeping with the Heisenberg uncertainty relation, the conjugate variable (ϕ) suffers an increase in variance. For Wigner functions, the purity is proportional to the inverse of the area under the function [41]. Here there has been no change in the total area and hence no change in the total purity. Because this is only a heuristic for understanding protocol, the previous statement is not entirely true; the final paragraph of this section will explain the actual result.

In Fig. 11(c) the optimal permutation in the QFT basis is depicted. The total area of the two rectangles before the measurement is $\sim 4\pi$. After a positive measurement result, the phase distributions for the two eigenvalues significantly overlap. The total area is now $\sim 3\pi$; this reduction in area leads to an increase in purity; see Fig. 11(d). Now the intuitive understanding for the speeding-up and the permutation sensitivity is apparent. The permutations are
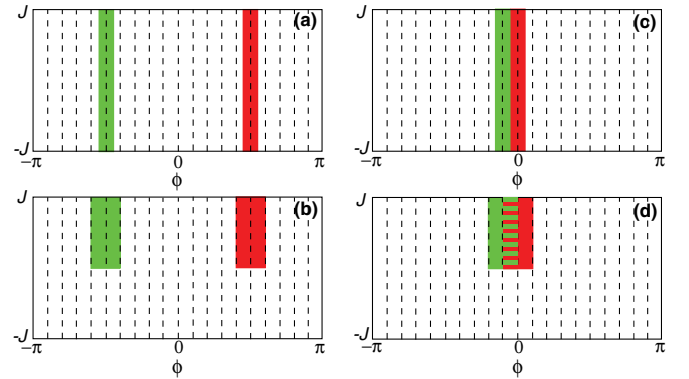

FIG. 10. An equal-area plot of the Wigner function of an equal mixture of two eigenstates after the QFT, $D = 10$. (a) The worst permutation; cf. Fig. 8(c). The dashed vertical lines correspond to the locations of the peaks of the D-transformed eigenstates. The dotted box around the peak at $\phi = 0$ corresponds to the bulk of the Wigner function for this state. (b) The conjectured general optimal permutation; cf. Fig. 8(b).

FIG. 11. Schematic Wigner functions for the mixtures of two-phase eigenstates for $D = 20$. (a) The worst permutation; cf. Fig. 8(c). We have rotated the entire sphere by $\pi/2$ about z to simplify the explanation of the speeding-up. (b) The worst permutation after a positive measurement result (modulo $\pi/2$). The total area is constant in (a) and (b), and it is equal to 4π. (c) The optimal permutation; cf. Fig. 8(b). (d) The optimal permutation after a positive measurement result. The total area is reduced in (d) to 3π, which gives rise to a purifying effect. (The purity is inversely proportional to the area underneath the Wigner function.) The striped region $\phi \in (-\pi/D, 0]$ denotes that the red and green rectangles are overlapping in this region. From these figures, it is apparent that the circles in Fig. 8 can be thought of as showing ϕ.

important so that the large eigenvalues in ρ may "bleed" into each other after a measurement. The "bleeding" is due to measurement backaction in the variable conjugate to J. It is the reduction in area that this bleeding effects that causes the purification; however, it only works if the largest two eigenvalues are adjacent in phase space. Although this picture is crude, it captures the essence of the protocol.

So far we have given an intuitive explanation of the mechanism underlying the purification. Now we provide an intuitive explanation for the speeding-up. Consider the integrated measurement result

$$\int_0^t dR(t) = \int_0^t \sqrt{4\gamma} X(s) ds + \int_0^t dW(s). \qquad (34)$$

This expression contains a term corresponding to the signal of interest (the first term) and a term representing the noise (the second term). We may define the signal-to-noise ratio as $\text{SNR} = \text{signal}^2/\text{noise}^2$. In this case, we have $\text{signal}^2 = 4\gamma\Delta t^2 \langle X^2 \rangle$ and $\text{noise}^2 = \langle(\Delta W(t))^2\rangle = \Delta t$. Thus, the ratio becomes $\text{SNR} = 4\gamma\Delta t\langle X^2 \rangle$. Now we wish to estimate the time taken to evolve from the Wigner functions in the first row of Fig. 11 to those in the second row. Note that an observer gains 1 bit of information about J_z between the first and second rows. From the Shannon-Hartley theorem, $\text{bits} = \log_2(1 + S/N)$, we infer that this implies $\text{SNR} = 1$. Now we solve for time and take $X = J_z$ consequently $\langle J_z^2 \rangle = (D^2 - 1)/12$ (see Appendix C in Ref. [9]) to find

$$\Delta t = \frac{3}{\gamma(D^2 - 1)}. \qquad (35)$$

Recall that the purity is approximately inversely proportional to the phase-space area. The initial purity in Fig. 11(c) is $P_i \propto 1/4\pi$, and we know that the state is an equal mixture of two eigenstates, which means that the purity equals one-half. This allows us to determine the proportionality constant to be 2π, so $P_i = \frac{1}{2}$. The final purity is thus $P_f = 2/3$, which means the change in purity is $\Delta P = 1/6$. Consequently,

$$\frac{1}{P}\frac{\Delta P}{\gamma\Delta t} = \frac{1}{2}\frac{(1/6)}{3/(D^2 - 1)} = \frac{1}{9}(D^2 - 1). \qquad (36)$$

This can be compared to Eq. (32), that is,

$$\frac{1}{L}\frac{dL}{\gamma dt} = -\frac{2D^2}{\pi^2} \approx -\frac{1}{5}D^2. \qquad (37)$$

Considering the crudeness of the arguments we have employed, the agreement between Eqs. (36) and (37) is surprisingly good. This gives weight to the intuition that the effect of purification for the QFT protocol comes from measurement backaction, and the role of the permutations is to maximize the backaction by placing the eigenvalues close in phase space.

We now return to the worst permutations for purification. All one can conclude from the equal likelihood of all outcomes is that there should be a speeding-up of at least 2 (in a discrete-outcome measurement model, which is

equivalent to our current measurement model). Let us confirm this intuition now. Returning to the two-eigenvalue approximation, consider the worst permutation. It is possible to determine the upper bound on the speeding-up resulting from this permutation by substituting $r - c = D/2$ into Eq. (29). We find that $S = 2$, which confirms the intuition above. This is why complementarity is not enough to guarantee maximal information extraction.

VI. ARE ALL COMPLEMENTARY BASES EQUALLY GOOD FOR ENTROPY REDUCTION?

We now address the question raised by Jacobs and Lund [19]: Are all unbiased bases equally good for rapid purification? Their answer was no. However, at the time of their analysis, the role of permutations was not clearly understood. Thus, it is worthwhile to reexamine this question. Because of rotational symmetry of the unbiased bases and the permutational symmetry of the density operator, the answer in the cases $D = 2$ and 3 is yes. However, for $D = 4$, it is easy to find a counterexample to this trend. For example, measuring in any of the four *mutually unbiased bases* (MUBs) [42] gives

$$\begin{aligned} dL &= -8\gamma dt \sum_{r,c} |\check{X}_{r,c}|^2 \lambda_r \lambda_c \\ &= -8\gamma dt \Big[2\lambda_0\lambda_1 + \frac{1}{2}(\lambda_1\lambda_2 + \lambda_0\lambda_3) + 2\lambda_2\lambda_3 \\ &\quad + 0(\lambda_0\lambda_2 + \lambda_1\lambda_3) \Big], \end{aligned} \qquad (38)$$

where $\check{X} = M_i^\dagger J_z M_i$, where M_i is a transformation to one of the MUBs. (Explicit expressions of the five MUBs when $D = 4$ are given in Ref. [42], for example. For convenience we have reproduced them in endnote [43].) As before, we may use ρ_2 to obtain a lower bound on dL (and hence an upper bound on the speeding-up). Doing so gives $dL = -8\gamma dtL(t)$, which implies a speeding-up of $S = 8$. This saturates the upper bound on the speeding-up found in Eq. (15). The factor of 2 improvement over the QFT can be understood from the difference in weights of the largest terms $\lambda_0\lambda_1$ between Eqs. (24) and (38).

To get more intuition about the purification process using the MUB transform, we look at the Wigner function of the four states in one of the four MUBs, which, for convenience, we denote $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$ in Fig. 12. The states plotted in Figs. 12(a) and 12(d) look like the states in the QFT basis (Pegg-Barnett phase states). The Wigner functions plotted in Figs. 12(b) and 12(c) are, however, quite different—they contain a hole which is more negative than the dips in Figs. 12(a) and 12(d). The states plotted in Figs. 12(b) and 12(c) can be said to be highly nonclassical because of this.

We now compare the MUB transform to the QFT using the Wigner function in an attempt to get an intuitive explanation of the advantage of the MUB transform. To make the
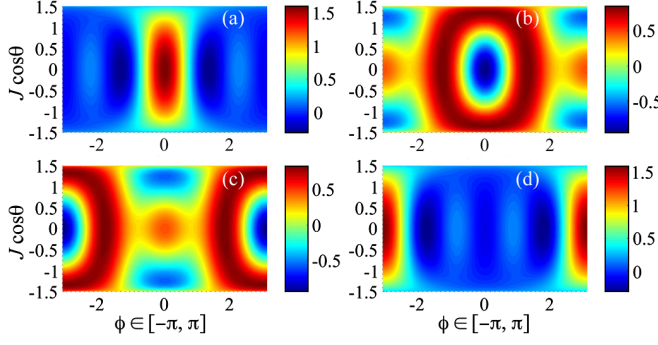
FIG. 12. Equal-area plots of the Wigner function resulting from transforming four different states to the M_1 unbiased basis ($D=4$). (a) The transform of $\rho_0 = \text{diag}(1, 0, 0, 0)$, (b) $\rho_0 = \text{diag}(0, 1, 0, 0)$, (c) $\rho_0 = \text{diag}(0, 0, 1, 0)$, and (d) $\rho_0 = \text{diag}(0, 0, 0, 1)$. Unlike the unbiased basis generated by the QFT, not all of the resulting states are like the Pegg-Barnett phase states; see, for example, (b) and (c).

analysis simple, we restrict our consideration to states that have two large eigenvalues only (as we did above). From Eq. (38), we see that permuting these states in the logical basis so that they correspond to $|0\rangle$ and $|1\rangle$ (or $|2\rangle$ and $|3\rangle$) is the optimal thing to do. Recall that permuting the two largest eigenvalues to $|0\rangle$ and $|1\rangle$ was also optimal for the QFT protocol; see Fig. 13(a). There we found that doing so created two peaks close in phase space, so the purification could be explained by measurement backaction. In Fig. 13(c) the Wigner function of the optimally permuted state for the MUB transform is depicted. [It is an equal mixture of Figs. 12(a) and 12(b).] Here we can no longer attribute the purification mechanism to backaction. In fact, it seems as though the purification effect comes from distinguishing the two peaks in the J_z distribution. The worst QFT permutation also corresponds to a poor MUB transform, as evidenced in Figs. 13(b) and 13(d).

Interestingly there is a permutation that is worse for the MUB transform than the permutation depicted in Fig. 13(d). It is $\rho = \text{diag}(1, 0, 1, 0)/2$; the transformed state of this permutation is depicted in Fig. 13(e). This permutation results in no purification at all. We can see this by substituting $\lambda_0 = \frac{1}{2}$ and $\lambda_2 = \frac{1}{2}$ into Eq. (38) to get $dL = 0$. This complete lack of purification may have application in the state stabilization of states with an arbitrary purity [19].

We would like to know if the upper bound on the speeding-up is saturated in all dimensions. To answer this question, we resort to a numerical search for $D \in [2, 10]$. We do not claim that our search is exhaustive. Recall that the stochastic simulations show that the speeding-up predicted by ρ_2 is close to the achievable amount. It seems reasonable to assume this to be true in other complementary bases. This greatly simplifies the analysis by making the optimization of permutations superfluous. For this reason, we may numerically search over all unbiased bases for the basis which has the largest element $|\check{X}_{r,c}|^2$.

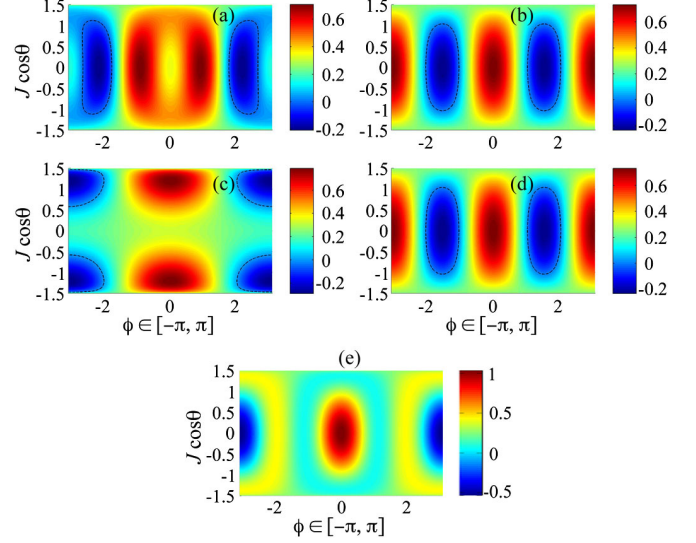

FIG. 13. A Wigner function comparison between the QFT (top row) and MUB transform (bottom row) for $D = 4$. We plot an equal mixture of two eigenstates with the optimal permutation [(a) and (c)] and a poor permutation [(b) and (d)] for each transform. In (a) through (d), the dashed black line is the zero contour. (a) The optimally permuted state for the QFT $\rho = \text{diag}(1, 1, 0, 0)/2$. (b) The worst permutation for the QFT $\rho = \text{diag}(1, 0, 1, 0)/2$. (c) The optimally permuted state for the MUB transform $\rho = \text{diag}(1, 1, 0, 0)/2$. (d) A poor permutation for the MUB transform $\rho = \text{diag}(1, 0, 0, 1)/2$. Note that cases (b) and (d) are actually identical. Unlike the QFT case, the mixture of optimally permuted MUB states has two peaks which could be distinguished by a measurement of J_z. (e) The worst permutation, i.e., $\rho = \text{diag}(1, 0, 1, 0)/2$, after the transform to the M_1 unbiased basis, which results in *no* purification.

Converting the element $|\check{X}_{r,c}|^2$ to a speeding-up gives the following trend. The speeding-up for even D attains the upper bound, i.e., $D^2/2$, and the speeding-up for odd D equals $(D-1)^2/2$, as seen in Fig. 14. In either case, this is
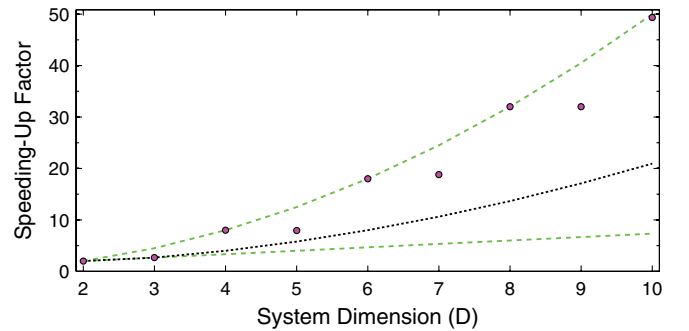

FIG. 14. The asymptotic speeding-up for reaching a given level of purity as a function of the system dimension. The dashed green lines are the upper and lower bounds on all unbiased-basis purification strategies, i.e., Eq. (15). The dotted black line is the maximum achievable speeding-up for the QFT protocol, i.e., Eq. (31). The solid circles are the results of a numerical search over unbiased bases to maximize the asymptotic speeding-up. See the text for more details.

much larger than the speeding-up of the QFT protocol, which was $S = 0.2D^2$. For $D \gg 1$, it is reasonable to believe that the achievable speeding-up does indeed scale like the predicted $S = 0.5D^2$ [9].

VII. INFORMATION ACQUISITION IN A REGISTER OF QUBITS AND ITS RELATIONSHIP TO COMPLEMENTARITY

Finally, we consider complementary measurements of a register of n qubits, where each qubit is independently and weakly (or continuously) measured, as introduced in Ref. [5]. Instead of one observable X, there are now n, given by $X^{(r)} = I^{(1)} \otimes I^{(2)} \otimes \ldots \sigma_z^{(r)} \ldots \otimes I^{(n)}$, where r labels the r-th qubit. The SME describing such a measurement is

$$d\rho = \sum_r 2\kappa dt \mathcal{D}[X^{(r)}]\rho + \sqrt{2\kappa}dW^{(r)}\mathcal{H}[X^{(r)}]\rho. \quad (39)$$

The combined state of the n qubits exists in a $D = 2^n$-dimensional Hilbert space.

A. Commuting-measurements analysis

Here we will not analyze the commuting measurements of a register of qubits with the same detail as we did for the qudit in Sec. III. Instead, we rely on a result from Ref. [9] where an expression for the evolution of the average impurity of a register undergoing a continuous-commuting measurement was found. In the long-time (LT) limit, the impurity is

$$\langle L^{(n)}(t)\rangle_{\mathrm{LT}} = \frac{n\pi e^{-4\kappa t}}{8\sqrt{\pi\kappa t}}. \quad (40)$$

We will use this expression in this section to calculate the speeding-up. For this purpose, only the asymptotic scaling is important: $\langle L^{(n)}(t)\rangle \sim e^{-4\kappa t}$.

B. Complementary measurements

It was shown in Ref. [9] that the change in impurity for a register of qubits monitored in an unbiased basis is

$$dL = -8\kappa dt \sum_{r=1}^{n} \mathrm{Tr}[\check{X}^{(r)}\rho\check{X}^{(r)}\rho], \quad (41)$$

$$= -8\kappa dt \sum_{r=1}^{n} \sum_{i,j=0}^{(D-1)} |\check{X}_{i,j}^{(r)}|^2 \lambda_i \lambda_j. \quad (42)$$

Here we define the complementary observable to be $\check{X}^{(r,m)} = P_m T X^{(r)} T^\dagger P_m^\dagger$. As before, the T's are conditional unitarities that introduce the unbiasedness [between ρ and $X^{(r)}$], and the P_m's are the permutation operators.

It was found that the upper bound on the impurity was

$$L_{\mathrm{UB}}^{(n)}(t) = \exp[-8\kappa nt/(D-1)]L(0), \quad (43)$$

while the lower bound was

$$L_{\mathrm{LB}}^{(n)}(t) = \exp(-8\kappa nt)L(0). \quad (44)$$

From these equations, we can infer the following bounds on the asymptotic speeding-up factor:

$$\frac{2n}{D-1} \leq S \leq 2n. \quad (45)$$

For all $n \geq 3$, the lower bound on the speeding-up becomes less than unity. For large n, the slowing-down implicit in the lower bound is $\sim n2^{-n+1}$. (The interested reader should also see the related study of Hill and Ralph [44].)

C. Complementarity via the quantum Fourier transform

From Eq. (45) it is not clear if measurements in a complementary basis provide any benefit in information extraction. In this subsection, we present some progress toward answering this question.

We begin our analysis by considering the speeding-up when $n = 2$. The bounds on the speeding-up are $4/3 \leq S \leq 4$. In order to see if the upper bound is achievable, we perform stochastic simulations. In Fig. 15 we numerically determine the advantage of the QFT feedback protocol in a
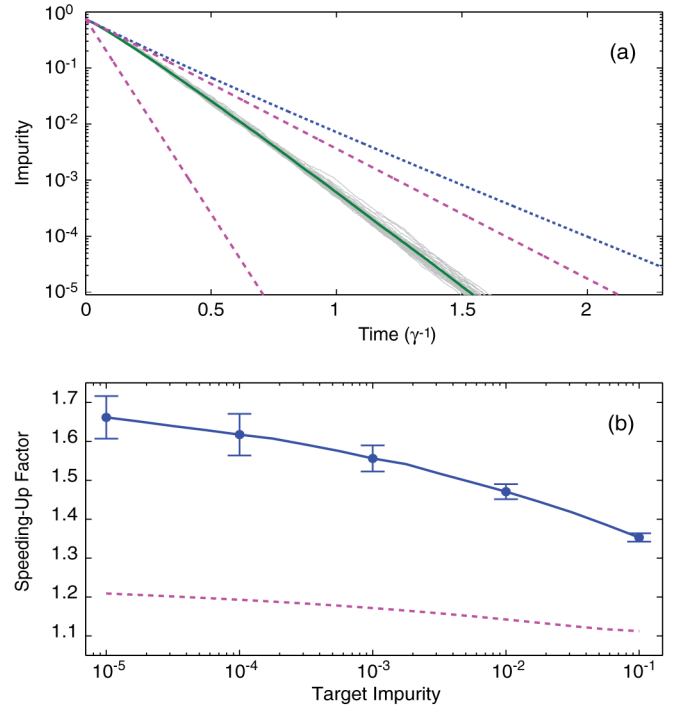

FIG. 15. (a) Impurity for a two-qubit register. The small-dashed blue line indicates a continuous measurement; the solid green line, the average impurity for feedback; the solid gray lines, individual trajectories; the dashed magenta lines, the upper [Eq. (43)] and lower [Eq. (44)] bounds, respectively. (b) Speeding-up is indicated by the solid blue line, and the lower bound by the dashed magenta line. The upper bound is not shown, as it is well above the numerical result. The ensemble size is 20, and feedback was applied at times $\delta t = 1 \times 10^{-3}\gamma^{-1}$.
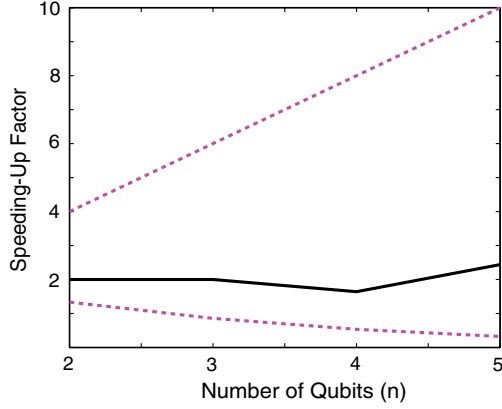
FIG. 16. Bounds on the asymptotic speeding-up for a register of qubits. The dashed magenta lines are the upper and lower bounds derived in the main text, Eq. (45). The solid line is the largest element of $X_{\max} = \max_{i,j}\sum_r |X_{i,j}^{(r)}|$. The element $X_{\max}$ is conjectured to be a good indicator for maximum achievable speeding-up with the QFT in a register of qubits.

register of two qubits. The permutation used for simulating the evolution was $\rho = (\lambda_0, \lambda_4, \lambda_3, \lambda_1)$.

In Fig. 15(a), the QFT protocol clearly does not saturate the bound on $\langle L \rangle$ given in Eq. (44). One may randomly permute the eigenvalue arrangement at times δt before applying the QFT feedback to deterministically achieve the lower bound of Eq. (45). From Fig. 15(b), it seems as though the speeding-up is approaching asymptotically toward 2 rather than the upper bound specified by Eq. (45), i.e., 4. For comparison, we note that the speeding-up found for the locally optimal rapid-measurement protocol in Ref. [5] was $S_{\mathrm{RM}} \approx 1.4$ in the long-time limit.

The structure of unbiased bases for a register of qubits is quite complicated [45–47]. Even for the QFT, the optimal permutation is not obvious for $n > 2$. It is for this reason that we plot the two-eigenvalue approximation to the asymptotic speeding-up for the QFT in Fig. 16 as a function of n. The values were obtained by finding the largest term in Eq. (42). This term will be denoted by $X_{\max} := \max_{i,j}\sum_r |X_{i,j}^{(r)}|$. The feedback places the two eigenvalues at $i_{\max}$ and $j_{\max}$. Thus, $X_{\max}$ is proportional to the speeding-up. Curiously, the speeding-up,

$$S \approx 2, \tag{46}$$

appears to be independent of the size of the register. It is not clear if this is true for any unbiased basis in a register. Our arguments from Sec. VI indicate that it might be higher by a multiplicative constant.

VIII. DISCUSSION

Prior to this work, it has been shown that it is possible to speed up the extraction of information from a quantum system using a continuous-complementary measurement [4,8,9]. In this paper, we have given an explicit method for

constructing such a protocol, using the QFT, that achieves a speeding-up of $S = 0.2D^2$ for a qudit. This explicit construction allowed us to show that complementarity between the state and measurement observable is not enough to guarantee maximal information extraction. Choosing the right permutation of eigenstates, before Fourier transforming the state—see Fig. 8—is needed to guarantee maximal information extraction. This finding is contrary to what one might expect from classical information theory [1]. The effect of the permutation can change the speeding-up in information extraction from its minimal value 2 to $0.2D^2$. We also argued that, in the case of the qudit, the maximum possible speeding-up, predicted in Ref. [9], $S \sim D^2/2$ should be achievable. One interesting upshot of our investigation (see Sec. VI) is an explicit example of using an unbiased basis for state stabilization [19].

In this paper, we demonstrated that measurement back-action is fundamental to the purification process for the QFT protocol. We did not, unfortunately, find an information theoretic explanation of the measurement process as was found in Ref. [48]. It is not clear, at present, that the bounds presented in Sec. IV C could be obtained from more elementary reasoning about uncertainty relations [49].

In a related work, Shabani and Jacobs [50,51] found the globally optimal (in time) protocol for $D = 3$ and the locally optimal protocol for all dimensions for the reduction of a quantity related to impurity. Their calculated bounds on the speeding-up were $2(D-1) \le S_{\mathrm{SJ}} \le 2(D-1)^2$. Naïvely comparing their bounds to ours suggests that an unbiased protocol is not optimal. However, it is not yet clear if their upper bound is achievable for $D > 4$.

It is possible to perform a simple calculation, independent of the one presented in Ref. [50], for the absolute upper bound on the speeding-up. First, we calculate a lower bound on $\mathrm{E}[dL]$,

$$\mathrm{E}[dL] = -8\gamma dt\{\mathrm{Tr}[\rho\check{X}\rho\check{X}] - 2\,\mathrm{Tr}[\check{X}\rho]\mathrm{Tr}[\check{X}\rho^2]$$
$$+ \mathrm{Tr}[\rho\check{X}]^2\mathrm{Tr}[\rho^2]\}, \tag{47}$$

by using the trace inequality $\mathrm{Tr}[ABAB] \le \mathrm{Tr}[A^2B^2]$, where A and B are Hermitian matrices. The breve on X denotes the unitary transformation $\check{X} = U^\dagger XU$, where U is *any* unitary. Note that all but the first term, $\mathrm{Tr}[X\rho X\rho]$, are $O(1 - L^2)$ [or $O(L^2)$]. Consequently, the first term dominates $\mathrm{E}[dL]$ as it is $O(L)$. Next we make a two-eigenvalue approximation to ρ. We find

$$\mathrm{E}[dL_2] = -8\gamma dt\mathrm{Tr}[\rho_2\check{X}\rho_2\check{X}] \tag{48}$$

$$\le -8\gamma dt\mathrm{Tr}[X^2U\rho_2^2U^\dagger] \le 2(D-1)^2. \tag{49}$$

This implies an upper bound on the speeding-up of $S = 2(D-1)^2$, as found by Shabani and Jacobs [50]. The results of a numerical search, like the one performed in Fig. 14, indicates that this bound is indeed achievable. Furthermore, we can confidently say the speeding-up

of the time-optimal control strategy for impurity reduction is bounded above by $S = 2(D - 1)^2$. This is justified by making a two-eigenvalue approximation to ρ_2 and then using the proof that Jacobs' protocol [4] is optimal [11,50–52].

It is interesting that the rapid-measurement protocol (considered in Refs. [5,7]) and the complementary-measurement protocols both afford *at most* a speeding-up $O(D^2)$. They work in very different ways. The rapid-measurement protocol is essentially classical in nature. It uses operations that ensure that the state and measurement commute at all times. In particular, the operations are permutations of the state in the measurement basis. The phase-space picture for the rapid-measurement protocol, in the long-time limit (i.e., under the two-eigenvalue approximation), would be two J_z eigenstates (Dicke states) placed at $\pm J$. As the measurements are in the z basis, the mechanism for the rapid-measurement speeding-up is an increase in signal-to-noise ratio in the measurement record. In this paper, we have shown that the complementary-measurement protocol is essentially quantum mechanical in nature. The mechanism for the purification and speeding-up is measurement backaction (at least for the QFT protocol). The two approaches also have different advantages. Rapid measurement has the advantage of enabling one to obtain information about the initial state. The complementary-measurement protocol presented here has the advantage of providing a nearly deterministic improvement in information gathering of the current system state. In future work, we will explore if it is possible to obtain information about the initial state for a continuous-complementary measurement.

Finally, we will speak to the practicality of implementing QFT-based protocols. In a recent work, we showed how to analyze imperfections in a purification protocol using a feedback master equation [53], so we will not discuss such issues further here. Instead, we will focus on the implementation the QFT. For qubits, the QFT is a fundamental quantum logic gate known as the Hadamard gate [54] and is typically easy to implement. For $D > 2$, implementing the QFT is more difficult, as it involves multiple logic gates (for n-qubit systems) [55] or multilevel coherent operations (for a qudit). Earlier we mentioned that in some atomic spin systems it should be possible to construct the desired QFT unitary [40]. We think, however, that solid-state systems are the most likely candidate for which it would be useful to implement the ideas presented in this paper. This is because, in solid-state systems, the measurement strength γ is typically much smaller (by an order of magnitude or more) than the maximum control strength $|\alpha|$ (e.g., in Ref. [56], $1/\gamma$ is on the order of microseconds and $1/|\alpha|$ is of order nanoseconds). Consequently, it is possible to imagine applying feedback continuously through the measurement process so that the eigenbasis of the state and observable are QFT pairs. Alternatively, if the maximum control strength is large enough, the feedback could

be applied impulsively at discrete times, which can gives results surprisingly close to the continuous version (as noted in Fig. 6 and Ref. [7]). Further, the exciting results in the field of superconducting qudits ($D \in [3, 5]$), where the measurement and full unitary control have been demonstrated for a number of systems [57,58], lead us to speculate that our proposals could be experimentally tested within the next decade.

APPENDIX: DETAILS OF COMMUTING MEASUREMENTS

1. The exact expression for impurity and the approximate expression

Using the linear trajectory (see Refs. [29–31]) solution, i.e., Eq. (5), we now explicitly calculate the form of the D-dimensional state matrix. We take the observable X to be the z component of the angular momentum (represented by the operator J_z); thus $D = 2J + 1$. In matrix form, the J_z operator is $J_z = \mathrm{diag}(J, J - 1, \ldots, -J + 1, -J)$. Using this representation of J_z, the unnormalized solution of the state evolution can be written as matrix elements exponentiated:

$$\bar{\rho}(R, t) = \frac{1}{D} \, \mathrm{diag}[e^{-4\gamma J^2 t} e^{2\sqrt{2\gamma} JR}, \, e^{-4\gamma (J-1)^2 t}$$
$$\times \, e^{2\sqrt{2\gamma}(J-1)R}, \ldots, e^{-4\gamma(-J)^2 t} e^{-2\sqrt{2\gamma} JR}]. \quad (A1)$$

The normalization factor can be written with a nonsymmetric sum or a symmetric sum:

$$\mathcal{N} \equiv \mathrm{Tr}[\bar{\rho}] = \frac{1}{D} \sum_{m=0}^{2J} e^{-4\gamma(J-m)^2 t} e^{2\sqrt{2\gamma}(J-m)R} \quad (A2)$$

$$= \frac{1}{D} \sum_{s=-J}^{J} e^{-4\gamma s^2 t} e^{2\sqrt{2\gamma} sR}. \quad (A3)$$

For later discussions, we also calculate the probability distribution of the result R and choose the symmetric sum for the normalization

$$dR\mathcal{P}(R) = dR\mathrm{Tr}[\bar{\rho}(R, t)]P(R) \quad (A4)$$

$$= dR \frac{1}{D\sqrt{2\pi t}} \sum_{s=-J}^{J} e^{-(R-2\sqrt{2\gamma} st)^2/2t}. \quad (A5)$$

Here $P(R)$ has been called the ostensible distribution for R [14,30]. It is given by $P(R, t) = \exp(-R^2/2t)/\sqrt{2\pi t}$. By

choosing the symmetric summation for the normalization, one can see that the probability distribution for R contains D peaks centered around the values $\propto s$, where $s \in [J, J-1, \ldots, -J+1, -J]$. For times $t \gg \gamma^{-1}$, the distribution is sharply peaked about these D values. Each of these peaks has a FWHM of $2\sqrt{2\ln2}\sqrt{t} \approx 2.35\sqrt{t}$. This seems paradoxical; one would expect that as time increases the distribution would get narrower, reflecting the fact that the observer is more confident about which eigenstate they actually have. This paradox is resolved when one realizes that the distance between the peaks increases at a rate greater than the increase in the width of the peaks. This time-dependent scaling of the distance between the peaks can be removed by changing variables to $V \equiv R/(2\sqrt{2\gamma t})$. Under this transformation, the probability distribution is

$$dV\mathcal{P}(V) = dV \frac{1}{D}\sqrt{\frac{4\gamma t}{\pi}} \sum_{s=-J}^{J} e^{-(V-s)^2/(1/4\gamma t)}. \qquad (A6)$$

Now the FWHM of each peak scales as $\sqrt{\ln2/(\gamma t)} \approx 0.83/\sqrt{\gamma t}$, which clearly illustrates that the probability distribution becomes sharply peaked about the D values for long times. In Fig. 17 the dashed curve is a plot of Eq. (A6) for $D = 5$ and $t = 4\gamma^{-1}$. All of the peaks are clearly distinguishable with a measured FWHM ≈ 0.418. This agrees with the prediction given by $0.83/\sqrt{\gamma t}$.

Returning to the nonsymmetric sums, i.e., Eq. (A2), for the remainder of this calculation, the purity is

$$\mathrm{Tr}[\rho^2] = \frac{1}{D^2\mathcal{N}^2} \sum_{m=0}^{2J} e^{-8\gamma(J-m)^2 t} e^{16\gamma t(J-m)V}. \qquad (A7)$$

The impurity, for a particular value of V, is given by

$$\Lambda(V,t) = \left[1 - \frac{\sum_{m=0}^{2J} e^{-8\gamma(J-m)^2 t} e^{16\gamma t(J-m)V}}{D^2\mathcal{N}^2}\right] \qquad (A8)$$
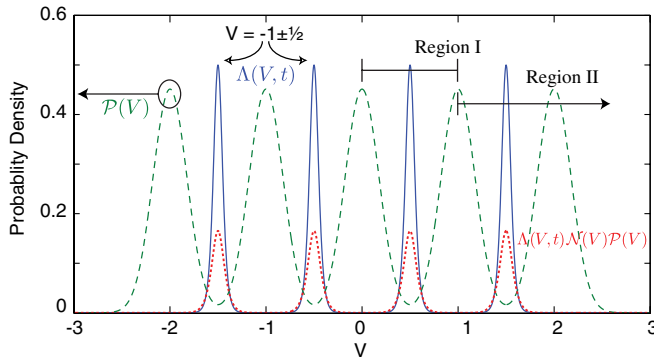

FIG. 17. The dashed green curve is the probability distribution $\mathcal{P}(V)$, Eq. (A6); the solid blue line is $\Lambda(V,t)$, Eq. (A9), shown for $D = 5$ at $t = 4\gamma^{-1}$. The dotted red curve is $\Lambda(V,t)\mathcal{N}(V)P(V)$ (multiplied by 20 so that it can be viewed on the same scale). Also plotted are regions for the two-eigenvalue expression for the impurity integral.

$$= \left[\frac{D^2\mathcal{N}^2 - \sum_{m=0}^{2J} e^{-8\gamma(J-m)^2 t} e^{16\gamma t(J-m)V}}{D^2\mathcal{N}^2}\right]. \qquad (A9)$$

The solid curve in Fig. 17 is a plot of Eq. (A9) for $D = 5$ and $t = 4\gamma^{-1}$. To find the average value of impurity for a continuous measurement of J_z, one must integrate this function over the scaled measurement record V, weighted by $P(V)$ [8]:

$$\langle L(t)\rangle = \int_{-\infty}^{\infty} \Lambda(t,V)\mathcal{N}(V)P(V)dV. \qquad (A10)$$

When using this result to calculate the measurement-induced evolution of the mean linear entropy, one should use the kernel in Eq. (A9) rather than the kernel in Eq. (A8). This is because it will reduce the numerical error that occurs when subtracting from one a number very close to 1.

To calculate this integral, one must perform numerical integration. It would be useful to have a fully analytic expression for $\langle L(t)\rangle$ in the D-dimensional case. Now we will make some approximations to obtain a simple analytic expression for Eq. (A10). We will show that the approximations allow one to place a bound on the full expression. Then we will confirm our approximations with numerical simulations in Appendices A 2 and A 3.

The first approximation involves truncating the state matrix to two eigenvalues and then renormalizing. We call this the two-eigenvalue approximation. The motivation for the truncation stems from the following observations. In the long-time limit ($t \gg \gamma^{-1}$), the true probability distribution, $\mathcal{P}(V) = \mathcal{N}(V)P(V)$, is sharply peaked in D places; see Fig. 17. Also in this limit one finds that $\Lambda(V,t)$, Eq. (A9), is sharply peaked in $(D-1)$ places; the peaks are between peaks of $\mathcal{P}(V)$. Wiseman and Ralph have pointed out that, for a qubit, this arrangement leads to poorly purifying trajectories dominating $\langle L(t)\rangle$ [6]. In a qubit, the poorly purifying trajectories are those trajectories with eigenvalues of the same magnitude [$V(t) \approx 0$] in the long-time limit. Physically, this means that the measurement has not projected, or the filter has not decided, if the state is in $z = \pm\frac{1}{2}$. In qudits, there are $(D-1)$ regions like this, where $V(t) \approx r$ for $r \in [J-\frac{1}{2}, J-\frac{3}{2}, \ldots, J+\frac{3}{2}, -J+\frac{1}{2}]$. In any of these regions only two eigenvalues contribute significantly to the integrand. Consequently, it is reasonable to truncate the state to the two eigenvalues in a particular region and renormalize.

The effect of the two-eigenvalue approximation is two-fold: it increases the purity of the state, and it makes it possible to derive a fully analytic expression for the impurity. The eigenvalues of the truncated state matrix are $(\lambda_0, \lambda_1)/(\lambda_0 + \lambda_1)$ [59], where λ_0 is the largest eigenvalue of ρ and $\lambda_a > \lambda_b$ when $a < b$. It is easy to show that ρ is majorized [54] by ρ_2 (that is, $\rho \prec \rho_2$), meaning that the original state is more mixed than our two-eigenvalue approximation to it. From the fact that the purity is Schur convex [60], it follows that $\mathrm{Tr}[\rho^2] \leq \mathrm{Tr}[\rho_2^2]$. This means

that the impurity of the two-eigenvalue approximation is a lower bound on the true impurity: $L(\rho) \geq L(\rho_2)$. In the long-time limit, it is reasonable to expect that $L(\rho) \approx L(\rho_2)$ from the arguments above.

Now we calculate the impurity of the truncated state matrix $\langle L_2(t) \rangle$. We split the integral in Eq. (A10) into regions such that the same two eigenvalues are the largest two eigenvalues for all values of V in each region. The integration is then performed in each of these regions, and then the regions are summed. Figure 17 illustrates how the regions should be split for $D = 5$. After a coordinate transform (for example, in region I it would be $V' = V - 1/2$), the integral in region I becomes

$$\mathfrak{R}_\text{I} = \frac{e^{-\gamma t}}{D}\sqrt{\frac{4\pi t}{\pi}}\int_{-(1/2)}^{1/2}\frac{e^{-4\gamma t V'^2}}{\cosh(4\gamma t V')}dV'. \quad \text{(A11)}$$

After a similar transformation, the integral in region II becomes

$$\mathfrak{R}_\text{II} = \frac{e^{-\gamma t}}{D}\sqrt{\frac{4\pi t}{\pi}}\int_{-(1/2)}^{\infty}\frac{e^{-4\gamma t V'^2}}{\cosh(4\gamma t V')}dV'. \quad \text{(A12)}$$

The total integral is thus

$$\langle L_2(t) \rangle = 2\mathfrak{R}_\text{II} + (D - 3)\mathfrak{R}_\text{I}. \quad \text{(A13)}$$

For $t \gg \gamma^{-1}$, the integrands become sharply peaked about $V' = 0$ with negligible contributions to the integrals from the region outside the interval $V' \in [-\frac{1}{2}, \frac{1}{2}]$, so that $\mathfrak{R}_\text{II} \simeq \mathfrak{R}_\text{I}$. In this limit, we have

$$\langle L_2(t) \rangle \simeq (D-1)\frac{e^{-\gamma t}}{D}\sqrt{\frac{4\pi t}{\pi}}\int_{-(1/2)}^{1/2}\frac{e^{-4\gamma t V'^2}}{\cosh(4\gamma V')}dV'. \quad \text{(A14)}$$

The final approximation is arrived at by noting that, in the long-time limit, the distribution in the numerator is broad compared to the distribution in the denominator for $t \gg \gamma^{-1}$. Thus, we can make the approximation:

$$\int_{-\infty}^{\infty}dVe^{-4\gamma t V'^2}/\cosh(\sqrt{2\gamma}V) \approx \int_{-\infty}^{\infty}dV/\cosh(\sqrt{2\gamma}V)$$
$$= \pi/\sqrt{2\gamma}.$$

The final expression for the impurity is thus

$$\langle L_2(t) \rangle_\text{LT} = \frac{2(D-1)}{D}\frac{\pi e^{-\gamma t}}{\sqrt{16\gamma t \pi}}. \quad \text{(A15)}$$

This is the analytic expression for $\langle L(t) \rangle$ that we set out to find, and is the key result of this section.

Because previous work has used the exact qubit results, we write it out for future reference:

$$\langle L(t) \rangle_\text{qubit} = \frac{e^{-\gamma t}}{2}\sqrt{\frac{4\pi t}{\pi}}\int_{-\infty}^{\infty}\frac{e^{-4\gamma t V'^2}}{\cosh(4\gamma t V')}dV'. \quad \text{(A16)}$$

The long-time limit of this expression is

$$\langle L_2(t) \rangle_\text{qubit-LT} = \frac{\pi e^{-\gamma t}}{\sqrt{16\gamma t \pi}}. \quad \text{(A17)}$$

As could reasonably be expected for a qudit under the two-eigenvalue approximation (effectively a two-level system), the resulting average impurity is proportional to the qubit impurity with the proportionality depending on the qudit dimensionality. In Appendix A 2, we examine the accuracy and validity of this approximation.

2. Comparison with numerics

In Fig. 18(a), one can see that, for $t \gtrsim \gamma^{-1}$, the impurity, found by numerically evaluating Eq. (A10), is better approximated by Eq. (A13) than by Eq. (A16). Moreover, for all times, the qubit impurity equation (A16) is a lower bound on the true impurity. However, when calculating the speeding-up, one is interested in times where $\gamma t \gg 1$, which are plotted in Fig. 18(b).

Equation (A13) is not plotted in Fig. 18(b) because on this scale it is practically indistinguishable from $\langle L \rangle$ at long times. (It approaches $\langle L \rangle$ from below.) But $\langle L_2(t) \rangle_\text{LT}$, that is, Eq. (A15), approaches $\langle L \rangle$ from above. Thus, the approximations made to obtain Eq. (A15) from Eq. (A13) increase the impurity. This effect is easily understood once one realizes what the approximations amount to: throwing away parts of the integral, Eq. (A14), and overestimating the integral by approximating $\exp(-4\gamma t V'^2) = 1$. The next closest curve is that of the long-time-limit expression for a qubit, Eq. (A17), and below that is the exact result, Eq. (A16).

From these curves, we infer the following: the long-time expression for the two-eigenvalue approximation, $\langle L_2(t) \rangle_\text{LT}$, is closer to $\langle L \rangle$ than the qubit long-time limit; $\langle L_2(t) \rangle_\text{LT}$ is approaching $\langle L \rangle$ in the same way that $\langle L(t) \rangle_\text{qubit-LT}$ is approaching $\langle L(t) \rangle_\text{qbit}$. This is important,
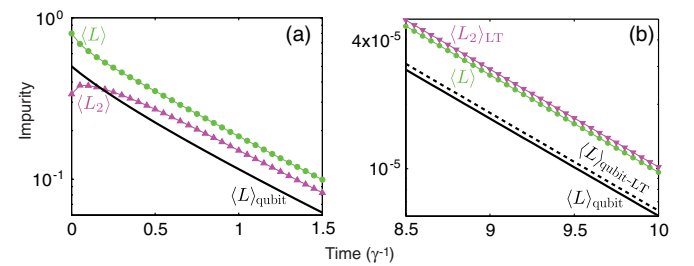

FIG. 18. A comparison of approximations to the evolution of $\langle L \rangle$, for $D = 3$. (a) $\langle L \rangle$, i.e., Eq. (A10), is shown by green circles; $\langle L_2 \rangle$, i.e., Eq. (A16) with the integrals evaluated numerically, is shown by magenta triangles; $\langle L \rangle_\text{qubit}$, i.e., Eq. (A10), is shown by the solid black line. (b) $\langle L_2 \rangle_\text{LT}$, i.e., Eq. (A15), is shown by magenta triangles; $\langle L \rangle$, i.e., Eq. (A10), is shown by green circles; $\langle L \rangle_\text{qubit-LT}$ i.e., Eq. (A17), is shown by the dashed black line; and $\langle L \rangle_\text{qubit}$, i.e., Eq. (A16), is shown by the solid black line. Equation (A13) is not plotted in (b) because, on this scale, it is indistinguishable from $\langle L \rangle$.

as previously in the literature it has been common to use the qubit long-time-limit expression as a lower bound for $\langle L \rangle$ any dimension [5,8]. It is clear that, for all times $t \gg \gamma^{-1}$, $\langle L_2(t) \rangle_{\text{LT}}$ is a better approximation to $\langle L \rangle$ than the qubit expression. Of course, asymptotically, the expressions decay at the same rate. Therefore, there is nothing in this work to suggest that the speeding-up calculated in previous works is incorrect.

3. Stochastic simulations

We now compare the analytic solutions from Sec. III A and the exact numerics from Appendix A 2 to stochastic nonlinear trajectory simulations of Eq. (1). We used an Euler integration method with the following parameters: $\delta t = 1 \times 10^{-4} \gamma^{-1}$, $D = 5$, ensemble size $= 20$. The ensemble size is small because we will plot all of the trajectories so that we may gain some qualitative understanding from them.

Figure 19(a) shows that the L of all of the trajectories is bounded from above by $\frac{1}{2}$ for $t \gg \gamma^{-1}$. Intuitively, this effect can be understood from the two-eigenvalue approximation; in the long-time limit a poorly purifying trajectory has two eigenvalues and thus these trajectories are bounded from above by $L = \frac{1}{2}$. A more rigorous explanation is given in Appendix A 4. Figure 19(a) also plots the numerically calculated ensemble average impurity and the linear trajectories solution, Eq. (A10).

There is some difficulty in obtaining convergence for $\langle L \rangle$ in the stochastic simulations, even with very large ensemble sizes. The convergence problem is due to a small number of poorly purifying trajectories first noted for a qubit in Ref. [6]. These poorly purifying trajectories are evident in Fig. 19(b) where many trajectories touch the dash-dotted line at $L = \frac{1}{2}$. For moderate ensemble sizes, the whole region of possible impurities is filled by the ensemble even at $t = 2\gamma^{-1}$. As noted above, the major contributions to $\langle L \rangle$ are from the regions where $V = r$ (see Fig. 17), that is, for results V that are as far as possible from the most likely values of V [the peaks of $\mathcal{P}(V)$]. Thankfully, one can use Eq. (A10) to calculate the evolution of the mean impurity. The advantages of this method are twofold: the computation time is greatly reduced, and it is exact to numerical precision.

Figure 19(b) shows that many trajectories also touch the dashed line corresponding to a scaling $L \sim \exp(-4\gamma t)$. The reason for this clustering, and that at $L = \frac{1}{2}$, will be explained in Appendix A 4.

4. The distribution of impurities

In Ref. [16], Stockton, van Handel, and Mabuchi gave reasonable but nonrigorous arguments for a bound [their Eq. (41)] on the degree of the spreading of trajectories (that is, on the width of the distribution of trajectories) for the model of Eq. (1). We now use linear trajectories to rigorously derive the bounds that they found. In all cases, it is easy to obtain an analytic expression, but the exact form of
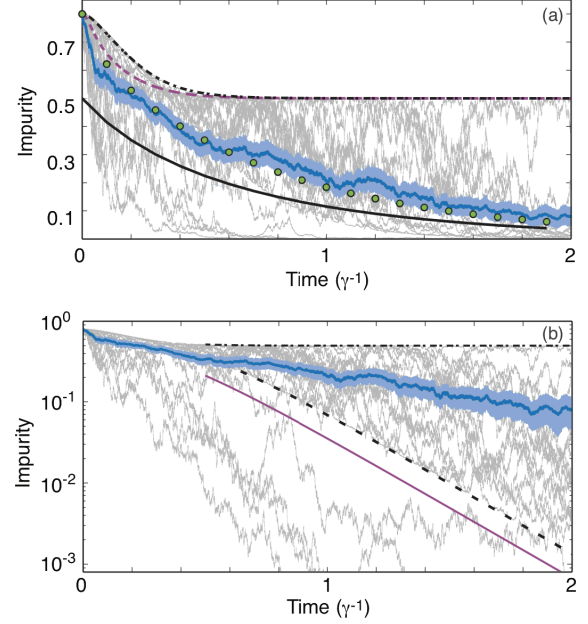

FIG. 19. Short-time impurity dynamics for $D = 5$. We plot 20 impurity trajectories (gray) and the ensemble average of the trajectories (dark blue). The error bars for the ensemble average, $\pm \sigma_{\text{mean}}$, are smeared around the mean (light blue). Time is measured in units of γ^{-1}. (a) Linear abscissa. The dashed magenta and dash-dotted black lines are approximate upper bounds on the L trajectories; see Appendix A 4 for further explanation. The solid black line is the exact mean impurity for the case of a qubit, i.e., Eq. (A16). The circles are $\langle L \rangle$ for $D = 5$, i.e., Eq. (A10). (b) The impurity is plotted on a logarithmic axis; this helps to show the approximately exponential decrease in impurity predicted by Eq. (A10). The approximate upper bound on the spread of L trajectories is the dash-dotted black line. The dashed black line corresponds to the most likely trajectories. Below the solid magenta line, there is only $1/D$ of the ensemble of trajectories at long times. These lines are explained in Appendix A 4.

the expression depends on D. Therefore, we quote only the long-time asymptotic scaling of the bounds.

Central to our explanation of the bounds of the distribution of trajectories is the intuition gained by examining Fig. 20, where we have replotted portions of Fig. 17 on two separate axes. In Fig. 20(a) $\mathcal{P}(V, t)$ is plotted. This probability-density plot is significant in the current analysis because it shows which records V are likely (the peaks of $\mathcal{P}$) and which are unlikely (the troughs). In Fig. 20(b) the kernel $\Lambda(V, t)$, that is, Eq. (A9), is plotted on a log scale. The key features of this plot are (1) the peaks at $V = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$, which represent very impure trajectories and correspond to unlikely records, and (2) the troughs at $V = -1, 0, 1$, which coincide with some of the purest trajectories and correspond to minima of $\mathcal{P}(V, t)$. We will elaborate on other relevant features of this graph below as each bound is explained.

(1) *The upper bound.*—This is obtained by substituting $V(t) = r$ (recall that $r \in [J - \frac{1}{2}, J - \frac{3}{2}, ..., J + \frac{3}{2}, -J + \frac{1}{2}]$)
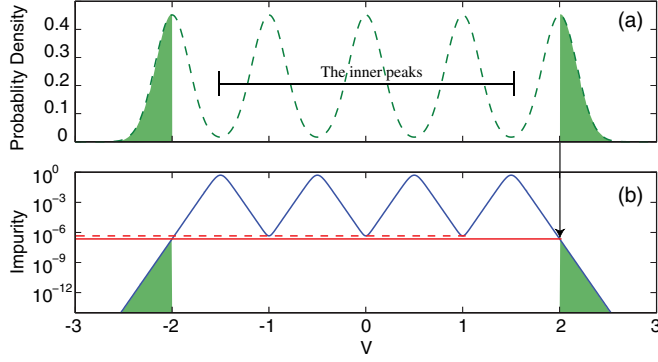
FIG. 20. (a) Probability distribution $\mathcal{P}(V, t)$, Eq. (A6), for $D = 5$ and $t = 4\gamma^{-1}$. The shaded regions of the probability distribution contain $1/D$ of the total probability at long times. (b) The solid blue line is $\Lambda(V, t)$, Eq. (A9), plotted with a logarithmic scale on the abscissa for $D = 5$ and $t = 4\gamma^{-1}$. The dashed red line is a projection onto the abscissa of the minima of $\Lambda(V, t)$ at $V = -1, 0, 1$. The solid red line is a projection onto the abscissa of the $\Lambda(V, t)$ evaluated at $V = \pm 2$; notice that this projection is purer than the dashed line. The shaded regions in (b) depict the impurity of records that have $|V| \geq J$. As these regions concur with the regions in (a), they can be interposed as saying that only a proportion $1/D$ of the trajectories will have impurities smaller than those given by the projection $\Lambda(\pm J, t)$.

into Eq. (A1), normalizing, and then calculating the impurity. This procedure gives the upper bound because these records correspond to the most impure trajectories, that is, peaks in Fig. 20(b). Because this procedure also corresponds to a worst-case scenario of the two-eigenvalue approximation, it can also be interpreted as the filter being unable to decided between the eigenvalues $\lambda_{r-(1/2)}$ and $\lambda_{r+(1/2)}$. For $D = 5$, the records of interest correspond to $V = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$ (see Figs. 17 and 20). In the long-time limit, all solutions give

$$L_{\mathrm{LT}} = \frac{1}{2}, \tag{A18}$$

which is also the bound predicted by Stockton *et al.* [16]. For small times, there is some variation between the bounds given by $V = r = \pm\frac{1}{2}$ and $V = r = \pm\frac{3}{2}$. This is because $\mathcal{P}(V)$ is broad at short times so that more eigenvalues contribute to the impurity when, for example, $V = \frac{1}{2}$. Conversely, when $V = \frac{3}{2}$, there is only one eigenvalue to the left of this record. Accordingly, it is reasonable to expect that at short times the bound given by $V = \pm\frac{1}{2}$ will produce a bound that is more impure than the bound given by $V = \pm\frac{3}{2}$. This behavior is confirmed in Fig. 19(a), where the dashed line is $V = \pm\frac{3}{2}$ and the dash-dotted line is $V = \pm\frac{1}{2}$.

(2) *A pseudolower bound.*—When $V(t) = s$ for $s \in [J - 1, \ldots, -J + 1]$, the record corresponds to the inner peaks of the probability distribution, as depicted in Fig. 20(a) (the peaks at $\pm J$ are *not* included in this analysis). These peaks of $\mathcal{P}(V)$ are the most likely records and coincide with some of the purest trajectories, the minima in Fig. 20(b). For $D = 5$, the peaks of interest are those corresponding to $V = -1, 0, 1$ in Fig. 20. By substituting $V(t) = s$ into Eq. (A1), normalizing, and then calculating the impurity, one obtains the lower bound without any approximations. These solutions can be obtained for any dimension D. By making a two-eigenvalue approximation, one can solve the evolution of the impurity analytically, giving

$$L_{\mathrm{LT}} \sim \exp(-4\gamma t). \tag{A19}$$

Equation (A19) corresponds to the lower bound quoted by Stockton *et al.* [16]. In their case, their bound appeared to be a true lower bound as their initial state was a coherent spin state (a collection of $D/2$ spin one-half particles where $D = 2J + 1$ with a mean spin vector, J, aligned along the x axis in this case). This has a small population in $|\pm J\rangle$, and it is therefore unlikely in their case that $|V| > |J|$. For our analysis, we see clearly that the bound is not a true lower bound [see Fig. 19(a)], because a maximally mixed state has equal populations in all eigenstates. However, because these records $V(t) = j$ are the most likely records, many of the trajectories touch or cluster around this bound. This will be further explored in point 4. below.

(3) *Physically likely trajectories.*—In our case there is no lower bound, although a trajectory that purifies infinitely fast is infinitely unlikely [61]. It is possible to give a natural bound on the physically likely trajectories based on the probability distribution $\mathcal{P}(V)$. In the lower-bound section above, we did not include the peaks of the probability distribution found at $V = \pm J$. It is clear from Fig. 20(b) that there is no corresponding minimum of the kernel $\Lambda(V, t)$ in this region. When $\Lambda(V, t)$ is evaluated at $V(t) = \pm J$, one finds that it gives impurities smaller [see the solid red line in Fig. 20(b)] than those which correspond to the inner peaks [the dashed red line in Fig. 20(b)]. From the shaded regions in Fig. 20(a), it is apparent that only $1/D$ of the probability density is found for $|V| \geq J$. This means that, in the long-time limit, only a proportion $1/D$ of the trajectories will have purities greater than the bound we now present. One obtains the bound using the linear trajectory solution with $V(t) = \pm J$ and solving for the impurity. By making a two-eigenvalue approximation, one finds that the bound scales asymptotically as

$$L_{\mathrm{LT}} \sim \exp(-4\gamma t). \tag{A20}$$

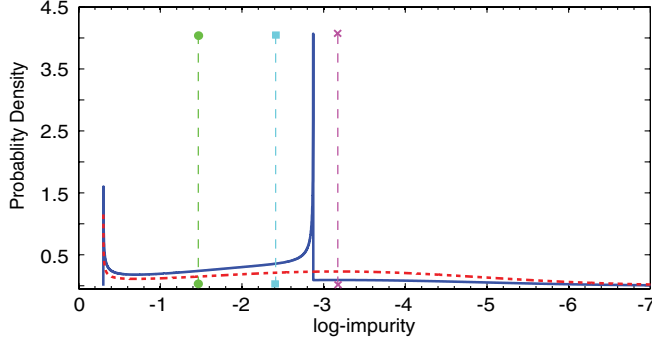The scaling in Eqs. (A19) and (A20) is precisely the same. The difference in the bounds is only apparent

FIG. 21. The distribution of log-impurities at $t = 2\gamma^{-1}$. The solid blue curve is the distribution for $D = 5$. The dashed red curve is for $D = 2$, i.e., Eq. (A21), and is plotted for comparison only. The dashed green line with the circles is $\log_{10}(\langle L \rangle) = -1.46$ at $t = 2\gamma^{-1}$, according to Eq. (A10). The dashed blue line with the squares is $\langle \ell \rangle = -2.41$ at $t = 2\gamma^{-1}$. Only a proportion $1/D$ of the trajectories has purity lower than (to the right of) the dashed magenta line with the crosses (plotted at $\ell = -3.1736$), as explained in the main text. The two features on the graph at $\ell = -0.3010$ and $\ell = -2.87$ correspond to the upper, Eq. (A18), and the pseudolower, Eq. (A19), bounds on the distribution of impurities, as described in the main text.

when the calculation is exact; see Fig. 19(b). The dashed black line is Eq. (A19), while the thin solid magenta line is Eq. (A20) (also see Fig. 21). Obviously, for a qubit ($D = 2$), the bound given by $V = \pm \frac{1}{2}$ corresponds to the median of the distribution of impurities, denoted by $\wp(L, t)dL$, at all times, and the mode at long times.

(4) *Distribution of impurities*—The distribution of impurities can be calculated analytically for $D = 2$ [6]. However, we will change variables to the log-impurity so that we may clearly see the feature discussed above. The required change of variables from Eq. (24) in Ref. [6] is $\ell \equiv \log_{10} L = \log_{10}[\frac{1}{2}(1 - z^2)]$ so that $z = \sqrt{1 - 2 \times 10^\ell}$. The distribution of log-impurities is thus

$$d\ell \wp(\ell, t)$$

$$= d\ell \frac{e^{-\gamma t} \ln(10) \cosh[\operatorname{arctanh}^2(\sqrt{1 - 2 \times 10^\ell})]}{2\sqrt{8\pi\gamma t}\sqrt{1 - 2 \times 10^\ell}}$$

$$\times \exp(-\operatorname{arctanh}^2(\sqrt{1 - 2 \times 10^\ell})/4\gamma t). \quad (A21)$$

In Fig. 21, this distribution is plotted at $t = 2\gamma^{-1}$ (the dashed red curve). The peak of this distribution at $\ell = -0.301$ becomes less prominent at long times. This is when, as Wiseman and Ralph pointed out [6], considering the average log-impurity is a good way to find an approximation to the mean time to a fixed L.

The poorly purifying trajectories affect $\langle L \rangle$, and a large number of these trajectories cluster around the upper

bound (as seen in Fig. 19). This motivates the consideration of another measure of mixedness which deemphasizes the peaks of the distribution [5,6]. However, in this paper, we will remain focused on $\langle L \rangle$.

For $D > 2$, the distribution of ℓ can be numerically calculated from the probability distribution $\mathcal{P}(V)$ and $\Lambda(V, t)$ as follows. First, one splits the distribution $\Lambda(V, t)$ into regions $\{(-\infty, -j), [-j, -j + \frac{1}{2}), \ldots, [j, \infty)\}$. These regions are chosen because ℓ is always monotonically increasing or decreasing in these regions; see Fig. 20(b). This is required in order to perform the numerical interpolation required by the following procedure. In each region one discretizes the range of ℓ and then finds the values of V that correspond to this ℓ. The probability distribution $\mathcal{P}(V)$ is then integrated over this region to give $d\ell \wp(\ell, t)$. This is then summed over all regions.

In Fig. 21, the solid curve is the probability distribution for $D = 5$ at $t = 2\gamma^{-1}$; this distribution was obtained using the above method. The upper bound and pseudo-lower bound described in the text above are quite apparent in distribution; they correspond to the two sharp features at $\ell = -0.301$ and $\ell = -2.87$. It is also apparent that $\langle \ell \rangle$ is a more faithful central-tendency measure than $\log_{10}\langle L \rangle$, as the line corresponding to $\langle \ell \rangle$ is closer to the bulk of the distribution.

5. Conclusion

We conclude our study of information extraction from commuting measurements by summarizing the key points. It is possible to characterize the average amount of information extracted using the average impurity. We found a simple expression for the average impurity, $\langle L_2(t) \rangle_{\mathrm{LT}} \sim \exp(-\gamma t)$, which is valid when $t \gg \gamma^{-1}$. Additionally, we found that in this asymptotic limit a portion $(1 - 1/D)$ of the trajectories would be bounded by $c \exp(-4\gamma t) \leq L \leq \frac{1}{2}$ for a constant c that we can determine for a particular D. It is the existence of rare but poorly purifying trajectories which explains both the upper bound of $\frac{1}{2}$ and the difference between the scaling of the mean and that of the lower bound, $c \exp(-4\gamma t)$. Indeed, other measures of central tendency, such as the median, or the exponential of the mean log, exhibit the same scaling as the lower bound, $\sim \exp(-4\gamma t)$, which is also reflected in the mean time to attain a given purity [7].

Despite its limitations, we have used $\langle L \rangle$, which scales as $\exp(-\gamma t)$, as our measure of average information in the main text of the paper. Here, we briefly comment on what would change if we were to have used a more faithful measure of central tendency such as the median. In Sec. IV B we found that the spread of trajectories for the QFT protocol was small for sufficiently frequent feedback. This feature is true, in general, for measurement in an unbiased basis [9]. The main consequence of this, for complementary-continuous measurements, is that the mean impurity is approximately equal to the median

impurity. Thus, we can directly compare the results obtained for $\langle L \rangle$ in the main part of the paper to the scaling for the median impurity, $\sim \exp(-4\gamma t)$, for the commuting-measurement protocols. The result, obviously, is a diminution in the speeding-up offered by the latter by a factor of 4: $\frac{1}{6}(D+1) \leq S^{\text{median}} \leq \frac{1}{8}D^2$. For small systems ($D = 2$ or 3), this implies a slowing-down, or at best a modest speeding-up. But for large D, the scaling with D implies that using feedback to construct suitable measurements in an unbiased basis will beat the no-feedback protocol even in terms of the median impurity or mean time.

———

[1] M. Tribus, *Thermostatics and Thermodynamics: An Introduction to Energy, Information and States of Matter, with Engineering Applications* (Van Nostrand, Princeton, 1961), p. 64.

[2] C. A. Fuchs and K. Jacobs, *Information-Tradeoff Relations for Finite-Strength Quantum Measurements*, Phys. Rev. A **63**, 062305 (2001).

[3] A. C. Doherty, K. Jacobs, and G. Jungman, *Information, Disturbance, and Hamiltonian Quantum Feedback Control*, Phys. Rev. A **63**, 062306 (2001).

[4] K. Jacobs, *How to Project Qubits Faster Using Quantum Feedback*, Phys. Rev. A **67**, 030301(R) (2003).

[5] J. Combes, H. M. Wiseman, and K. Jacobs, *Rapid Measurement of Quantum Systems Using Feedback Control*, Phys. Rev. Lett. **100**, 160503 (2008).

[6] H. M. Wiseman and J. F. Ralph, *Reconsidering Rapid Qubit Purification by Feedback*, New J. Phys. **8**, 90 (2006).

[7] J. Combes, H. M. Wiseman, and A. J. Scott, *Replacing Quantum Feedback with Open-Loop Control and Quantum Filtering*, Phys. Rev. A **81**, 020301(R) (2010).

[8] J. Combes and K. Jacobs, *Rapid State Reduction of Quantum Systems Using Feedback Control*, Phys. Rev. Lett. **96**, 010504 (2006).

[9] J. Combes, H. M. Wiseman, K. Jacobs, and A. J. O'Connor, *Rapid Purification of Quantum Systems by Measuring in a Feedback-Controlled Unbiased Basis*, Phys. Rev. A **82**, 022307 (2010).

[10] C. M. Caves, C. A. Fuchs, and R. Schack, *Unknown Quantum States: The Quantum de Finetti Representation*, J. Math. Phys. (N.Y.) **43**, 4537 (2002); **49**, 019902 (2008).R. W. Spekkens, *Evidence for the Epistemic View of Quantum States: A Toy Theory*, Phys. Rev. A **75**, 032110 (2007).

[11] H. M. Wiseman and L. Bouten, *Optimality of Feedback Control Strategies for Qubit Purification*, Quant. Info. Proc. **7**, 71 (2008).

[12] H. J. Carmichael, *An Open Systems Approach to Quantum Optics*, Lecture Notes in Physics, Vol. 18 (Springer-Verlag, Berlin, 1993).

[13] N. Gisin, *Quantum Measurements and Stochastic Processes*, Phys. Rev. Lett. **52**, 1657 (1984); C. M. Caves and G. J. Milburn, *Quantum-Mechanical Model for Continuous Position Measurements*, Phys. Rev. A **36**, 5543 (1987); L. Diosi, *Continuous Quantum Measurement and Itō Formalism*, Phys. Lett. A **129**, 419 (1988); V. P. Belavkin, *Nondemolition Measurement and Nonlinear Filtering of Quantum Stochastic Processes* (Springer-Verlag, Berlin, 1988); J. Dalibard, Y. Castin, and K. Mølmer, *Wave-Function Approach to Dissipative Processes in Quantum Optics*, Phys. Rev. Lett. **68**, 580 (1992); A. Barchielli and M. Gregoratti, *Quantum Trajectories and Measurements in Continuous Time: The Diffusive Case* (Springer-Verlag, Berlin, 2009).

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

[15] W. H. Zurek, *Decoherence and the Transition from Quantum to Classical—Revisited*, Progress in Mathematical Physics **48**, 1 (2006).

[16] J. K. Stockton, R. van Handel, and H. Mabuchi, *Deterministic Dicke-State Preparation with Continuous Measurement and Control*, Phys. Rev. A **70**, 022106 (2004).

[17] S. J. Dolinar, *An Optimum Receiver for the Binary Coherent State Quantum Channel*, in *MIT Research Laboratory of Electronics Quarterly Program Report* No. 111, 115120 (1973); H. M. Wiseman, *Adaptive Phase Measurements of Optical Modes: Going Beyond the Marginal Q Distribution*, Phys. Rev. Lett. **75**, 4587 (1995); H. M. Wiseman and R. B. Killip, *Adaptive Single-Shot Phase Measurements: The Full Quantum Theory*, Phys. Rev. A **57**, 2169 (1998); B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, *Entanglement-Free Heisenberg-Limited Phase Estimation*, Nature (London) **450**, 393 (2007).

[18] V. P. Belavkin, *Measurement, Filtering and Control in Quantum Open Dynamical Systems*, Rep. Math. Phys. **43**, A405 (1999); H. M. Wiseman, *Quantum Theory of Continuous Feedback*, Phys. Rev. A **49**, 2133 (1994); A. N. Korotkov, *Selective Quantum Evolution of a Qubit State Due to Continuous Measurement*, Phys. Rev. B **63**, 115403 (2001); H. M. Wiseman and A. C. Doherty, *Optimal Unravellings for Feedback Control in Linear Quantum Systems*, Phys. Rev. Lett. **94**, 070405 (2005).

[19] K. Jacobs and A. P. Lund, *Feedback Control of Nonlinear Quantum Systems: A Rule of Thumb*, Phys. Rev. Lett. **99**, 020501 (2007).

[20] H. M. Wiseman, S. Mancini, and J. Wang, *Bayesian Feedback versus Markovian Feedback in a Two-Level Atom*, Phys. Rev. A **66**, 013807 (2002).

[21] R. Ruskov, A. Korotkov, and K. Mølmer, *Qubit State Monitoring by Measurement of Three Complementary Observables*, Phys. Rev. Lett. **105**, 100506 (2010).

[22] K. Jacobs, *Optimal Feedback Control for Rapid Preparation of a Qubit*, Proc. SPIE-Int. Soc. Opt. Eng. **5468**, 355 (2004).

[23] H. M. Wiseman and G. J. Milburn, *Quantum Theory of Field-Quadrature Measurements*, Phys. Rev. A **47**, 642 (1993).

[24] K. Jacobs and D. Steck, *A Straightforward Introduction to Continuous Quantum Measurement*, Contemp. Phys. **47**, 279 (2006).

[25] T. A. Brun, *A Simple Model of Quantum Trajectories*, Am. J. Phys. **70**, 719 (2002).

[26] K. Jacobs, *Stochastic Processes for Physicists* (Cambridge University Press, Cambridge, England, 2010).

[27] R. van Handel, J. K. Stockton, and H. Mabuchi, *Feedback Control of Quantum State Reduction*, IEEE Trans. Autom. Control **50**, 768 (2005).

[28] S. L. Adler, D. C. Brody, T. A. Brun, and L. P. Hughston, *Martingale Models for Quantum State Reduction*, J. Phys. A **34**, 8795 (2001).

[29] P. Goetsch and R. Graham, *Linear Stochastic Wave Equations for Continuously Measured Quantum Systems*, Phys. Rev. A **50**, 5242 (1994).

[30] H. M. Wiseman, *Quantum Trajectories and Quantum Measurement Theory*, Quantum Semiclass. Opt. **8**, 205 (1996).

[31] K. Jacobs and P. L. Knight, *Linear Quantum Trajectories: Applications to Continuous Projection Measurements*, Phys. Rev. A **57**, 2301 (1998).

[32] J. Combes and K. Jacobs (unpublished).

[33] In the simulations, we used a linear trajectory type of simulation so that positivity of the state is ensured.

[34] Wolfram Research, Inc., Finite Summation of Csc, http://functions.wolfram.com/01.10.23.0002.01, accessed 14/04/2009.

[35] J. C. Varilly and J. M. Gracia-Bondia, *The Moyal Representation for Spin*, Ann. Phys. (N.Y.) **190**, 107 (1989).

[36] J. Combes and H. M. Wiseman, *States for Phase Estimation in Quantum Interferometry*, J. Opt. B: Quantum Semiclass. Opt. **7**, 14 (2005).

[37] F. T. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, *Atomic Coherent States in Quantum Optics*, Phys. Rev. A **6**, 2211 (1972).

[38] D. T. Pegg and S. M. Barnett, *Unitary Phase Operator in Quantum Mechanics*, Europhys. Lett. **6**, 483 (1988).

[39] D. T. Pegg and S. M. Barnett, *Phase Properties of the Quantized Single-Mode Electromagnetic Field*, Phys. Rev. A **39**, 1665 (1989).

[40] S. T. Merkel, G. Brennen, P. S. Jessen, and I. H. Deutsch, *Constructing General Unitary Maps from State Preparations*, Phys. Rev. A **80**, 023424 (2009).

[41] W. Schleich, *Quantum Optics in Phase Space* (Wiley-VCH Verlag, Berlin, 2001).

[42] A. Klappenecker and M. Rötteler, "Constructions of Mutually Unbiased Bases" in *Finite Fields and Applications: 7th International Conference, Fq7, Toulouse, France, 2003*, Lecture Notes in Computer Sciences Vol. 2948 (Springer-Verlag, Berlin, 2004), p. 262.

[43]

$$M_0 = [(1, 0, 0, 0)', (0, 1, 0, 0)', (0, 0, 1, 0)', (0, 0, 0, 1)'],$$

$$M_1 = [(1, 1, 1, 1)', (1, 1, -1, -1)', (1, -1, -1, 1)',$$
$$(1, -1, 1, -1)']/2,$$

$$M_2 = [(1, -1, -i, -i)', (1, -1, i, i)', (1, 1, i, -i)', (1, 1, -i, i)']/2,$$

$$M_3 = [(1, -i, 1, i)', (1, i, 1, -i)', (1, -i, -1, -i)', (1, i, -1, i)']/2,$$

$$M_4 = [(1, -i, -i, -1)', (1, -i, i, 1)', (1, i, i, -1)', (1, i, -i, 1)']/2.$$

[44] C. Hill and J. F. Ralph, *Weak Measurement and Rapid State Reduction in Entangled Bipartite Quantum Systems*, New J. Phys. **9**, 151 (2007).

[45] A. B. Klimov, L. L. Sánchez-Soto, and H. Guise, *Multicomplementary Operators via Finite Fourier Transform*, J. Phys. A **38**, 2747 (2005).

[46] J. Lawrence, Č. Brukner, and A. Zeilinger, *Mutually Unbiased Binary Observable Sets on N Qubits*, Phys. Rev. A **65**, 032320 (2002).

[47] J. L. Romero, G. Björk, A. B. Klimov, and L. L. Sánchez-Soto, *Structure of the Sets of Mutually Unbiased Bases for N Qubits*, Phys. Rev. A **72**, 062310 (2005).

[48] P. Kurzyński, W. Kaszub, and M. Czechlewski, *Mutually Unbiased Bases and Complementary Spin 1 Observables*, J. Phys. A **43**, 265303 (2010).

[49] S. Massar and P. Spindel, *Uncertainty Relation for the Discrete Fourier Transform*, Phys. Rev. Lett. **100**, 190401 (2008).

[50] A. Shabani and K. Jacobs, *Locally Optimal Control of Quantum Systems with Strong Feedback*, Phys. Rev. Lett. **101**, 230403 (2008).

[51] K. Jacobs and A. Shabani, *Quantum Feedback Control: How to Use Verification Theorems and Viscosity Solutions to Find Optimal Protocols*, Contemp. Phys. **49**, 435 (2008).

[52] V. P. Belavkin, A. Negretti, and K. Mølmer, *Dynamical Programming of Continuously Observed Quantum Systems*, Phys. Rev. A **79**, 022123 (2009).

[53] J. Combes and H. Wiseman, *Quantum Feedback for Rapid State Preparation in the Presence of Control Imperfections*, J. Phys. B, At. Mol. Opt. Phys. **44**, 154008 (2011).

[54] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).

[55] J. Chiaverini, J. Britton, D. Leibfried, E. Knill, M. D. Barrett, R. B. Blakestad, W. M. Itano, J. D. Jost, C. Langer, R. Ozeri *et al.*, *Implementation of the Semiclassical Quantum Fourier Transform in a Scalable System*, Science **308**, 997 (2005).

[56] J. Petta, A. C. Johnson, J. M. Taylor, E. A. Laird, A. Yacoby, M. D. Lukin, C. M. Marcus, M. P. Hanson, and A. C. Gossard, *Coherent Manipulation of Coupled Electron Spins in Semiconductor Quantum Dots*, Science **309**, 2180 (2005).

[57] R. Bianchetti, S. Filipp, M. Baur, J. M. Fink, C. Lang, L. Steffen, M. Boissonneault, A. Blais, and A. Wallraff, *Control and Tomography of a Three Level Superconducting Artificial Atom*, Phys. Rev. Lett. **105**, 223601 (2010).

[58] M. Neeley, Markus Ansmann, Radoslaw C. Bialczak, Max Hofheinz, Erik Lucero, Aaron D. O'Connell, Daniel Sank, Haohua Wang, James Wenner, Andrew N. Cleland *et al.*, *Emulation of a Quantum Spin with a Superconducting Phase Qudit*, Science **325**, 722 (2009).

[59] The order and placement of the eigenvalues on the diagonal will depend on the initial state.

[60] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics (Springer, New York, 1996).

[61] P. E. F. Mendonça and J. Combes (unpublished).