UC Berkeley
UC Berkeley Previously Published Works

Title
What is the optimal way to prepare a Bell state using measurement and feedback?

Permalink

Journal

ISSN

Authors
Martin, Leigh
Sayrafi, Mahrud
Whaley, K Birgitta

Publication Date

DOI

Peer reviewed

What is the optimal way to prepare a Bell state using measurement and feedback?

Leigh Martin[1,2],[*] Mahrud Sayrafi[1], and K. Birgitta Whaley[1,3]

[1]*Berkeley Center for Quantum Information and Computation, Berkeley, California 94720 USA*
[2]*Department of Physics, University of California, Berkeley, California 94720 USA and*
[3]*Department of Chemistry, University of California, Berkeley, California 94720 USA*

(Dated: July 17, 2017)

Recent work has shown that use of quantum feedback can significantly enhance both the speed and success rate of measurement-based remote entanglement generation, but it is generally unknown what feedback protocols are optimal for these tasks. Here we consider two common measurements that are capable of projecting into pairwise entangled states, namely half- and full-parity measurements of two qubits, and determine in each case a globally optimal protocol for generation of entanglement. For the half-parity measurement, we rederive a previously described protocol using more general methods and prove that it is globally optimal for several figures of merit, including maximal concurrence or fidelity, and minimal time to reach a specified concurrence or fidelity. For the full-parity measurement, we derive a protocol for rapid entanglement generation related to that of (C. Hill, J. Ralph, Phys. Rev. A 77, 014305 (2008)), and then map the dynamics of the concurrence of the state to the Bloch vector length of an effective qubit. This mapping allows us to prove several optimality results for feedback protocols with full-parity measurements. We further show that our full-parity protocol transfers entanglement optimally from one qubit to the other amongst all measurement-based schemes. The methods developed here will be useful for deriving feedback protocols and determining their optimality properties in many other quantum systems subject to measurement and unitary operations.

1. INTRODUCTION

The non-local nature of quantum entanglement has played a major role in the theoretical development of quantum mechanics, and is anticipated to play an equally significant part in its practical applications. Remote entanglement is a necessary resource for many quantum technologies, such as quantum cryptography, distributed and blind [1] quantum computing, and general quantum networking. Such applications may take place over kilometer length scales, but intermediate ranges are also relevant. For example, fault-tolerant quantum computing may require hundreds of millions of qubits, and the resulting device could be extremely large in size [2]. To construct and characterize such a system, it is likely to be built in a modular way, much like a modern supercomputer [3, 4]. Efficient distribution of entanglement between spatially separate parts of such a system will then be an architectural requirement that presents a substantial engineering challenge.

Of the existing experimental schemes for entanglement generation between quantum objects that are too far apart to directly interact, many are based on measurement [5–11] and dissipation [12], rather than relying on quantum gates. Such schemes are more robust to loss, which tends to increase with separation distance. By combining signals from two separate objects and measuring the combined signal, these schemes can project a system into an entangled state while only being sensitive to loss from one pass through the connecting channel.

Although these methods are becoming ubiquitous in experimental quantum science, little is known about the optimal way to use them. Furthermore, deterministic as opposed to probabilistic generation of remote entanglement between stationary (*i.e.* non-flying) qubits remains an outstanding experimental challenge. However, theoretical work has demonstrated that measurement-based feedback can be used to enhance the success rate and fidelity of measurement-based entanglement generation [13–15]. The fact that more entanglement can be created during the time over which measurement collapse occurs indicates that quantum feedback provides a fundamental advantage, but it is unknown what measurement-based feedback protocols are optimal for such tasks or which schemes are most efficient in principle.

Such optimization lies in the domain of quantum optimal control. However, to date only a limited number of measurement-based quantum feedback protocols have been established to be globally optimal, owing in part to the non-linear nature of the problems. A feedback scheme for qubit purification was discovered by Jacobs [16] and its optimality properties were studied in a number of subsequent works [17–20]. A protocol for rapid purification of a qudit was shown to exist in [21], and constructed explicitly in [22]. Upper bounds on qudit and multiqubit purification speedups are known [23], but proving global optimality for all situations remains an open challenge except for the qutrit case, which was solved in [22]. A globally optimal protocol for rotating a monitored qubit to a desired state was given in [24], but the precise way in which it is optimal is somewhat unnatural and several basic questions remain open.

Regarding feedback for entanglement generation, some optimality results are known for linear quadratic Gaussian systems [25–27]. Feedback protocols for enhancing

*Electronic address: Leigh@Berkeley.edu

the rate of entanglement generation using half-parity and full-parity measurements were given in [15] and [14] respectively, but without proof of optimality. Both protocols generate entanglement faster than is possible with measurement alone. They are also of interest because they drive the system along a deterministic path in Hilbert space; feedback is chosen to exactly cancel the randomness introduced by measurement, allowing for deterministic entanglement generation. This is particularly significant for the half-parity measurement, which has only a 50% success rate in the absence of feedback.

In this work, we prove that the half-parity measurement protocol given in [15] is globally optimal for several practically relevant tasks listed in the following section. For the full-parity measurement, we derive a protocol which is related to that given in [14] but which is designed for initially pure states. Elaborating on the connection between rapid entanglement generation and rapid purification first discussed in [14], we derive a natural mapping between this two-qubit system and measurement of an effective qubit, in which the concurrence [28] exactly maps to this qubit's Bloch vector length. We then use this mapping to prove several optimality results for the proposed full-parity protocol. We also show that our full-parity protocol is optimal among all measurement-based entanglement generation schemes that acquire single qubit information at a fixed rate. This result sets a useful standard by which the effectiveness of an entanglement generation protocol may be assessed.

We begin in Sec. 2 with an introduction to continuous measurements and list the optimality results to be proved. Section 3 rederives the half-parity feedback protocol in a more general way and presents the full-parity feedback protocol. The main section of the paper is Sec. 4, where we prove our optimality results. Sec. 5 discusses potential applications and a number of future directions. Details for implementation of the full-parity feedback protocol and its relation to that of Ref. [14] are given in appendix B, where we also elaborate on the connection between rapid purification and rapid entanglement generation.

2. NON-LOCAL AND CONTINUOUS MEASUREMENTS

The measurements we will consider are the full-parity measurement $X_F = \sigma_{z1}\sigma_{z2}/2$ which distinguishes odd-from even-parity states, and the half-parity measurement $X_H = (\sigma_{z1} + \sigma_{z2})/2$. The latter is so called because it does not distinguish the odd parity states $|01\rangle$ and $|10\rangle$, but does distinguish the even parity states $|00\rangle$ and $|11\rangle$. If two qubits are initialized in the separable state $(|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle)$ (ignoring normalization), then a full-parity measurement collapses the state into either $|01\rangle + |10\rangle$ or $|00\rangle + |11\rangle$, both of which are entangled. In contrast, the half-parity measurement only creates entanglement with a 50% success rate, because the sep-

arable outcomes $|00\rangle$ or $|11\rangle$ occur half the time. Despite this disadvantage, half-parity measurements are often easier to implement in practice. These operators have both been implemented as continuous measurements in superconducting qubits in references [29] and [6] respectively. Although the latter did not involve remote qubits, full-parity measurements can be implemented on remote qubits as well using the same experimental system [5].

In the idealized limit of instantaneous projective measurements, the notion of applying feedback during a measurement is ill-defined and full-parity measurement alone therefore already creates perfect, deterministic entanglement with no time delay. However realistic measurements occur over a finite duration as a continuous process which we describe in detail below. This provides the opportunity to apply feedback by performing additional operations to a system conditional on the measurement outcome up to that point in time. As these measurements are often resource-intensive, it is desirable to optimize the rate of entanglement generation.

In this work, we assume a controller has the ability to apply feedback only in the form of local unitary (LU) rotations on the measured qubits, so that measurement is the only source for entanglement generation. This restriction substantially simplifies experimental implementation and is essential when the qubits are remotely separated. The protocols for X_F and X_H are given in Sec. 3. We name these as P_F and P_H respectively. Each protocol involves applying proportional feedback, *i.e.*, local unitary rotations with a rotation angle that is proportional to the most recent measurement outcome. In Sec. 4, we prove the following global optimality properties:

Half-parity measurement $X_H = (\sigma_{z1} + \sigma_{z2})/2$:
P_H is globally optimal for the following tasks.

1. Max. concurrence goal: P_H maximizes the expectation value of the entanglement (quantified using the concurrence $\mathcal{C}$ [28]) reached at a chosen stopping time T.

2. Max. fidelity goal: Same as above but for fidelity $\mathcal{F}$ [30] with respect to any target state which is pure and maximally entangled.

3. Min. time goal for concurrence: P_H minimizes the expected time to reach a desired concurrence $\mathcal{C}_{\text{Threshold}}$, so long as $\mathcal{C}_{\text{Threshold}} \leq 1/\sqrt{2}$.

4. Min. time goal for fidelity: Same as above but for a desired fidelity $\mathcal{F}_{\text{Threshold}}$, so long as $\mathcal{F}_{\text{Threshold}} \leq (1 + \sqrt{2})/\sqrt{8}$

Full-parity measurement $X_F = \sigma_{z1}\sigma_{z2}/2$:

1. Max. concurrence goal: P_F is globally optimal.

2. Max. Fidelity goal: P_F is globally optimal

3. Min. time goal ($\mathcal{C}$ or $\mathcal{F}$): Measuring without feedback is globally optimal.

The min. time results for concurrence also apply to any monotonic function of $\mathcal{C}$, such as entanglement of formation. Finally, we show that P_F is optimal for the max. concurrence goal among all protocols that acquire single qubit information at a fixed rate. A precise statement of this result is given with its proof in Sec. 4.

We now provide a brief introduction to continuous measurement. In both applications mentioned above, the dynamics of the monitored system can be modeled as continuous measurements. These are described by a stochastic master equation [31–33]:

$$\rho(t + dt) = \rho(t) + \mathcal{D}[M]\rho(t)dt + \mathcal{H}[M]\rho(t)\sqrt{\eta}dW(t)$$

$$M \equiv \sqrt{\frac{\Gamma}{2}}X, \tag{1}$$

where X is either the full or half-parity measurement operator. Following convention, we have defined the Lindblad dissipator $\mathcal{D}[M] \equiv M\rho M^\dagger - 1/2(M^\dagger M\rho + \rho M^\dagger M)$ and the measurement superoperator $\mathcal{H}[M]\rho \equiv M\rho + \rho M^\dagger - \langle M + M^\dagger \rangle \rho$. Γ is the measurement rate, which sets the timescale for collapse onto an eigenstate of X. Lastly, η is the measurement efficiency, which characterizes how much of the signal is collected and observed. η may range from zero to one. Because measurement introduces randomness into the evolution of a wave function, the equations of motion must have a stochastic component, which in this case is given by dW, a Gaussian distributed zero mean random variable with variance dt. In this work, we treat this quantity according to the rules of Ito calculus, although the Stratonovic formalism may also be used [31, 34].

To study the dynamics of Eq. (1), one may use Monte Carlo simulation to sample dW and generate possible realizations of the measurement outcome, usually termed quantum trajectories [31]. In an actual experiment, $dW(t)$ is derived from the signal generated by the measurement apparatus, dV_t. The relation between this signal and dW is [32, 33]

$$dV_t = \langle X \rangle(t)dt + \frac{dW(t)}{\sqrt{2\eta\Gamma}}. \tag{2}$$

Feedback then consists of applying an additional Hamiltonian which depends on both the measurement record $dV(t')$ over the interval $0 \le t' \le t$, and the initial state. Equivalently, one can use Eq. (1) and $dV(t')$ to estimate the current state and calculate feedback accordingly.

In what follows, we make the simplifying assumptions that $\eta = 1$ and that the initial state is pure. In the absence of additional decoherence, the state remains pure at all times and one may then replace Eq. (1) with a simpler stochastic Schödinger equation

$$\psi(t + dt) = \left[-\frac{1}{2}(X - \langle X \rangle)^2 dt + (X - \langle X \rangle)dW(t) \right]|\psi(t)\rangle, \tag{3}$$

where we have set the measurement rate $\Gamma = 2$ for simplicity, a convention that we will retain in what follows.

3. FEEDBACK PROTOCOLS FOR ENTANGLEMENT GENERATION

Our objective is to quantify how entanglement changes under arbitrary feedback protocols for binary systems, so that we can identify which are optimal. We characterize entanglement using the concurrence, which for pure states is defined as [28]

$$\mathcal{C} \equiv |\langle \psi^* | \sigma_y \otimes \sigma_y |\psi\rangle| \tag{4}$$

with $\langle \psi^* |$ the complex conjugate of $\langle \psi |$, or equivalently the transpose of $|\psi\rangle$. Like all valid entanglement measures, $\mathcal{C}$ is invariant under local unitary rotations, so our allowed feedback operations leave it unchanged. For the time being, we take $\mathcal{C}$ to be our figure of merit.

Since all bipartite pure states with the same concurrence are equivalent up to LU operations [35], it is possible to parameterize any pure state in terms of $\mathcal{C}$ and single qubit rotations. Such a parameterization is useful in this context because feedback can directly control the latter quantities. Therefore we can model feedback to directly set these qubit rotation parameters to desired values without specifying the Hamiltonian necessary to prepare the resulting state. The Schmidt decomposition provides an explicit example of such a parameterization [35]. By expressing the Schmidt coefficients in terms of the concurrence, we can write a general two-qubit state as

$$\psi(\mathcal{C}, \theta_1, \theta_2, \phi_1, \phi_2, \gamma_1, \gamma_2) \tag{5}$$

$$= U_1 \otimes U_2 \left[\sqrt{\frac{1 + \sqrt{1 - \mathcal{C}^2}}{2}}|00\rangle - \sqrt{\frac{1 - \sqrt{1 - \mathcal{C}^2}}{2}}|11\rangle \right]$$

$$U_i \equiv \exp(-i\gamma_i\sigma_z/2)\exp(-i\sigma_y\theta_i/2)\exp(-i\phi_i\sigma_z/2),$$

where we have written U_i in terms of the Euler angles $\{\phi_i, \theta_i, \gamma_i\}$. For convenience in subsequent calculations, we break the local unitaries into symmetric and antisymmetric rotations by defining $\theta \equiv (\theta_1 + \theta_2)/2$, $\Delta\theta \equiv (\theta_1 - \theta_2)/2$, and likewise for ϕ and γ. The final expression does not depend on $\Delta\phi$ because the state in brackets is invariant under antisymmetric rotations about σ_z. This separation of variables into an LU-invariant quantity depending on $\mathcal{C}$ and LU rotations depending on $\{\theta_i, \phi_i, \gamma_i\}$ motivates an analogy with parameterization of a qubit in terms of a Bloch vector $\vec{r}$ in spherical coordinates; $\mathcal{C}$ is analogous to $|\vec{r}|$, which is invariant under unitary operations, while $\{\theta_i, \phi_i, \gamma_i\}$ can be set arbitrarily using Hamiltonian feedback like θ and ϕ. When considering the full-parity measurement in Sec. 4, we will find that this analogy admits an explicit mapping in which $\mathcal{C}$ and r obey the same equations of motion under measurement.

To study how entanglement changes under measurement, we substitute Eq. (5) into Eq. (3) and compute the concurrence of the resulting state $\psi(t + dt)$. It is not necessary to compute how the five angles parameterizing

ψ evolve, since we will model the controller to set them according to some feedback protocol. The computation is further simplified by the fact that σ_z rotations commute with X_F and X_H, so that the resulting equations of motion do not depend on γ or $\Delta\gamma$. We first focus on the half-parity measurement X_H. As is detailed in appendix A, application of Ito's lemma yields

$$dC = \begin{cases} 2C\sqrt{1-C^2}uv\ dW \\ \quad + [(v^2-u^2)w - C(v^2+u^2)]dt & |\ C > 0 \\ |v^2-u^2|\ dt & |\ C = 0 \end{cases} \quad (6)$$

where we have defined the control variables $u = \cos(\theta)$, $v = \cos(\Delta\theta)$ and $w = \cos(2\phi)$.

In what follows, we implicitly assume that at any time, the controller can instantaneously set the angles $\theta, \Delta\theta$ and ϕ to any desired value. This is equivalent to assuming that the controller can implement any local unitary on both qubits with no time delay. For many systems such as superconducting qubits, single qubit rotations can be performed much faster than the measurement time scales, which makes this approximation appropriate. The assumption of zero time delay can be satisfied as long as the propagation delay between the qubits is small relative to the inverse measurement rate. However, this restriction can be relaxed for the full-parity protocol, as it will turn out that the protocol can be implemented by applying feedback on only one of the two qubits. Thus, Eq. (6) is an equation of motion for the concurrence of the state under arbitrary feedback protocols that may be specified by choosing a particular, set $u(t, C)$, $v(t, C)$ and $w(t, C)$.

Although there is no reason *a priori* that a locally optimal strategy is also globally optimal, it often turns out to be so in practice. This will be the case in the results that we prove here. The locally optimal protocol maximizes the expectation value of the concurrence at time $t + dt$. To find it, one simply chooses the values of u, v and w to maximize the dt term of Eq. (6). This occurs for $\{u = 0, v = 1, w = 1\}$ with all values of γ and $\Delta\gamma$ allowed.[1] We henceforth refer to this protocol as P_H. The resulting equations of motion under this set of control parameters may be easily solved:

$$dC = (1-C)dt \implies C(t) = 1 - (1-C(0))e^{-t}. \quad (7)$$

The evolution of the state under this feedback protocol may be computed by substituting this solution and the above control values into Eq. (5). Setting $C(0) = 0$ for

----

[1] $\{u = 1, v = 0, w = -1\}$ is also a solution, but since it is actually equivalent to the first if one makes the transformation $\gamma \to \gamma - \pi/2$, $\Delta\gamma \to \Delta\gamma + \pi/2$, we ignore it. Solutions corresponding to $u = -1$ or $v = -1$ are similarly redundant.

simplicity, the state evolution under feedback is thus

$$\psi(t) = \frac{1}{2}\begin{bmatrix} e^{-i\gamma}\sqrt{e^{-t}} \\ e^{-i\Delta\gamma}\sqrt{2-e^{-t}} \\ e^{i\Delta\gamma}\sqrt{2-e^{-t}} \\ e^{i\gamma}\sqrt{e^{-t}} \end{bmatrix}. \quad (8)$$

For $\gamma = \Delta\gamma = 0$, this state evolution exactly matches the solution to the feedback protocol given in [15], indicating that they are the same feedback protocol.

Equation (8) can be more easily understood by writing down the states for $t = 0$ and $t \to \infty$. Up to a global phase, these states are

$$\psi(0) = \frac{1}{2}\big(|0\rangle + e^{i\gamma_1}|1\rangle\big) \otimes \big(|0\rangle + e^{i\gamma_2}|1\rangle\big)$$
$$\psi(\infty) = \frac{1}{\sqrt{2}}\big(|01\rangle + e^{2i\Delta\gamma}|10\rangle\big). \quad (9)$$

Thus this protocol involves preparing both qubits in a separable state polarized along some axis of the equator of their respective Bloch spheres. Feedback deterministically projects the qubits into a Bell pair with the relative phase determined by the relative equatorial angle of the initial preparation. As shown in Ref. [15], this protocol may be implemented with proportional feedback, *i.e.*, by continuously applying local qubit rotations with an angle proportional to the measurement signal dV.

Before proving global optimality results for the above protocol, we first repeat the above analysis for the full-parity measurement. Following similar steps to those above but now for X_F, the equations of motion for the concurrence can be shown to be

$$dC = \begin{cases} (1-C^2)\big[(u^2-v^2)w\ C\ dW \\ \quad + (u^2-v^2)^2(1-w^2)dt/2C\big] & |\ C \neq 0 \\ (u^2-v^2)dW & |\ C = 0, \end{cases} \quad (10)$$

A minor technicality in the derivation given in appendix B has forced us to allow C to take negative on values in general. However since $\psi(C)$ only depends on C^2, this fact presents no further difficulty. We shall simply interpret $|C|$ as the concurrence instead of C.

For the full-parity measurement, two distinct sets of locally optimal parameters emerge: $\{u = 0, v = 1, w = 0\}$ and $\{u = 1, v = 0, w = 0\}$. The resulting state evolutions are equivalent up to a complex conjugation of ψ, so the underlying dynamics are therefore essentially identical. We may therefore only consider the protocol $P_F : \{u = 0, v = 1, w = 0\}$ and ignore the other solution. As for the half-parity case, the equations of motion for $C(t)$ are deterministic and again yield an analytic expression with easy solution:

$$dC = \frac{1-C^2}{2C}dt \implies C(t) = \pm\sqrt{1-[1-C(0)^2]e^{-t}}. \quad (11)$$

Validity of this solution at $\mathcal{C}(0) = 0$ is most easily established by deriving equations of motion for $\mathcal{C}(t)^2$ at $\mathcal{C} = 0$ and showing that the solutions coincide. State evolution under this protocol is given by

$$\psi(t) = \frac{1}{\sqrt{8}} \begin{bmatrix} e^{-i(\gamma+\pi/4)}\left(\sqrt{1-e^{-t/2}} - i\sqrt{1+e^{-t/2}}\right) \\ e^{-i(\Delta\gamma+\pi/4)}\left(\sqrt{1-e^{-t/2}} + i\sqrt{1+e^{-t/2}}\right) \\ e^{i(\Delta\gamma-\pi/4)}\left(\sqrt{1-e^{-t/2}} + i\sqrt{1+e^{-t/2}}\right) \\ e^{i(\gamma-\pi/4)}\left(\sqrt{1-e^{-t/2}} - i\sqrt{1+e^{-t/2}}\right) \end{bmatrix}. \tag{12}$$

Again, taking early- and late-time limits of the solution gives some insight into the induced dynamics. Here

$$\psi(0) = \frac{1}{2}\left(|0\rangle + e^{i\gamma_1}|1\rangle\right) \otimes \left(|0\rangle - ie^{i\gamma_2}|1\rangle\right) \tag{13}$$

$$\psi(\infty) = \frac{1}{2}\left(e^{-i\gamma}|00\rangle + e^{-i\Delta\gamma}|01\rangle + ie^{i\Delta\gamma}|10\rangle - ie^{i\gamma}|11\rangle\right),$$

where as before we have dropped a global phase from the states. The optimal initial state is again to prepare both qubits in an equal superposition of $|0\rangle$ and $|1\rangle$. Somewhat counterintuitively, the final state produced by feedback is not an eigenstate of the measurement operator. This is somewhat analogous to Jacobs' purification speedup protocol, in which the state is maintained to be in an equal superposition of the measurement eigenstates [16]. Application of Jacobs' protocol to an encoded qubit led the authors of [14] to a method for converting a classically correlated mixed state into a maximally entangled state. The protocol presented here performs the analogous task for an initially pure state. In appendix B, we show that $\mathcal{C}(t)$ under application of P_F coincides with that given in [14]. In Sec. 4, we will establish an even more precise connection between P_F and Jacobs' rapid qubit purification protocol. More details are given in appendix B, where we derive the feedback Hamiltonian that implements P_F and outline P_F's relation to [14].

The performance of the half-parity and full-parity feedback protocols are plotted in Fig. 1. We also plot the average concurrence without feedback for both measurement operators. We have normalized the measurement operators X_F and X_H so that the dephasing rate on one of the two qubits (tracing out the other qubit and assuming no feedback) are the same under measurement, allowing quantitative comparison the half- and full-parity performance. It is evident that the full-parity measurement with feedback strictly outperforms all other protocols.

4. GLOBAL OPTIMALITY IN ENTANGLEMENT GENERATION

We now set out to prove the optimality results stated in Sec. 2. For the half-parity feedback protocol, we apply the standard verification theorems often used in control theory to check whether a protocol is globally optimal
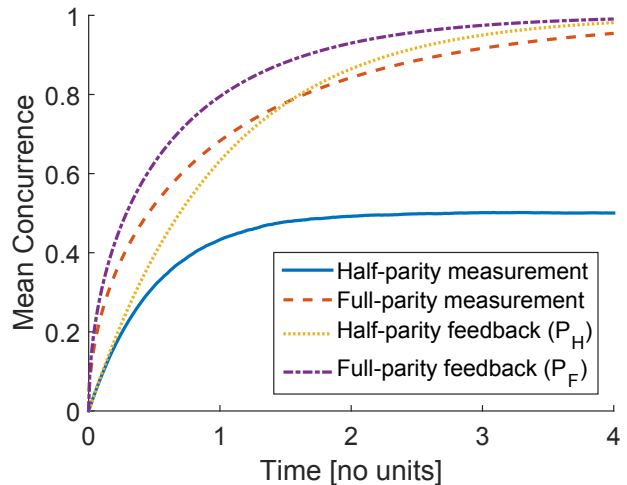

FIG. 1: (Color online) Expectation value of the concurrence while applying feedback or by applying measurement alone. Using the method of linear quantum trajectories [36], concurrence under full-parity measurement without feedback can be shown to take the form $\mathcal{C}(t) = \mathrm{erf}(\sqrt{t/2})$. Concurrence under half-parity measurement without feedback is calculated numerically by averaging 10000 trajectories.

[18, 37]. While one might expect that proving such a result would require knowledge about all allowed strategies, it turns out that knowing the performance of the trial protocol is sufficient, so long as one knows how it behaves for any initial state.

Before using the verification theorems, we first provide a brief summary following Ref. [37]. A common feedback goal is to minimize the expectation value of some cost function after applying feedback for a fixed time interval T. The cost function can depend on the final state of the system as well as the resources used to apply feedback. From the cost function, one defines the cost-to-go $c(P, \mathbf{x}, t)$ where the vector $\mathbf{x}$ represents the system state at time t. Here c is defined to be the expectation value of the cost function at time T, assuming that the system was in the state $\mathbf{x}$ at time t and that the controller used feedback protocol P from t to T. To check optimality, one first generically writes the equations of motion for $\mathbf{x}$ in terms of their deterministic and stochastic parts as

$$d\mathbf{x} = \mathbf{A}(t, \mathbf{v}_P(t), \mathbf{x}(t))\ dt + \mathbf{B}(t, \mathbf{v}_P(t), \mathbf{x}(t))\ dW \tag{14}$$

where $\mathbf{v}_P$ are the feedback control settings and parameters specified by protocol P. In the context of quantum feedback, $\mathbf{v}$ can represent Hamiltonian, measurement and dissipation parameters. The procedure can easily be generalized to include multiple noise processes if needed. If $\partial c(P)/\partial t$ and $\partial^2 c(P)/\partial \mathbf{x}^2$ are continuous, then global optimality of P follows from the following two conditions. Firstly, assuming the cost function only depends on the final state, the Hamilton-Jacobi-Bellman equation for the

cost-to-go $c(P, \mathbf{x}, t)$, given by

$$G(t, \mathbf{u}, \mathbf{x}, P) \equiv \tag{15}$$
$$-\frac{1}{2}\mathbf{B}(t, \mathbf{u}, \mathbf{x})^\mathsf{T} \frac{\partial^2 c(P)}{\partial \mathbf{x}^2} \mathbf{B}(t, \mathbf{u}, \mathbf{x}) - \mathbf{A}(t, \mathbf{u}, \mathbf{x})^\mathsf{T} \frac{\partial c(P)}{\partial \mathbf{x}}$$

and

$$\frac{\partial c(P)}{\partial t} = \max_{\mathbf{u}}[G(t, \mathbf{u}, \mathbf{x}, P)]. \tag{16}$$

must be satisfied for all time 0 to T and for all states $\mathbf{x}$. Secondly, $\mathbf{u} = \mathbf{v}_P$ must maximize G for all time and states.

In order to apply the verification theorems, we parameterize the state $|\psi\rangle$ as a vector $\mathbf{x} = \{\mathcal{C}, u, v, w\}$ using Eq. (5) and the definitions of u, v and w given below Eq. (6). Applying a specific local unitary feedback protocol is equivalent to setting $\mathbf{v} = \{u, v, w\}$ equal to specific functions of t and $\mathcal{C}$.

Half-parity measurement: For the max. concurrence goal (maximizing $\langle \mathcal{C}(T)\rangle$), we define our cost function to be $1-\mathcal{C}(t)$. Choosing the locally optimal protocol for the half-parity measurement $P_H : \{u = 0, v = 1, w = 1\}$, we can use the analytic solution Eq. (7) to find the cost-to-go when evolving according to P_H as

$$c(P_H, \mathbf{x}, t) = (1 - \mathcal{C}(t))e^{t-T}. \tag{17}$$

The Hamilton-Jacobi-Bellman equation uses the following derivatives of c

$$\frac{\partial c}{\partial t} = (1 - \mathcal{C}(t))e^{t-T}, \quad \frac{\partial c}{\partial \mathbf{x}} = \{-e^{t-T}, 0, 0, 0\}, \quad \frac{\partial^2 c}{\partial \mathbf{x}^2} = 0 \tag{18}$$

which satisfy the continuity conditions required by the verification theorem. We divide the equations of motion under feedback and half-parity measurement Eq. (6) into their deterministic and stochastic parts as

$$d\mathcal{C} = A_\mathcal{C} \ dt + B_\mathcal{C} \ dW$$

$$A_\mathcal{C} = \begin{cases} [(v^2 - u^2)w - \mathcal{C}(v^2 + u^2)]dt & | \ \mathcal{C} > 0 \\ |v^2 - u^2| \ dt & | \ \mathcal{C} = 0 \end{cases} \tag{19}$$

$$B_\mathcal{C} = \begin{cases} 2\mathcal{C}\sqrt{1-\mathcal{C}^2}uv \ dW & | \ \mathcal{C} > 0 \\ 0 & | \ \mathcal{C} = 0 \end{cases}$$

where we have only calculated the $\mathcal{C}$ components of $\mathbf{A}$ and $\mathbf{B}$. Due to the form of G in Eq. (15), the other components are unnecessary, since $\partial c / \partial \mathbf{v} = 0$, *i.e.*, our cost function is invariant under the feedback control settings. Substituting Eqs. (18) and (19) into Eqs. (15) and (16), we find the condition for P_H to be globally optimal is that it satisfies the maximization condition

$$(1 - \mathcal{C}(t))e^{t-T} \tag{20}$$
$$= \max_{\{u,v,w\}}[(v^2 - u^2)w - \mathcal{C}(t)(v^2 + u^2)]e^{t-T}.$$

The maximum occurs for the values of u, v and w specified by P_H, and the equation then is satisfied. This proves that the half-parity protocol P_H is globally optimal for maximizing the concurrence at fixed time T. Because $c(P_H, \mathbf{x}, t)$ is linear in $\mathcal{C}$, one can also show that global optimality follows directly from local optimality in this case [19], but this proof method is not applicable in general.

A similar calculation can be performed in which we test for global optimality with respect to the min. time goal [37] (minimizing the expected time at which $\mathcal{C}$ reaches some desired value $\mathcal{C}_{\text{Threshold}}$, called the expected hitting time). One finds that P_H maximizes G only when $\mathcal{C} \leq 1/\sqrt{2}$, which implies that P_H is not globally optimal in general. If one chooses $\mathcal{C}_{\text{Threshold}} \leq 1/\sqrt{2}$ however, then only the dynamics of the system when $\mathcal{C} \leq 1/\sqrt{2}$ are relevant for determining the expected hitting time. Thus we can instead ask if P_H is globally optimal within the constraint that $\mathcal{C} \leq 1/\sqrt{2}$. Restricted to this parameter space, the Hamilton-Jacobi-Bellman equation Eqs. 15-16 is satisfied for all allowed values of $\mathbf{x}$, and P_H maximizes G. This proves that P_H is globally optimal for the min. time goal when $\mathcal{C}_{\text{Threshold}} \leq 1/\sqrt{2}$. As the expected hitting time is the same whether one considers $\mathcal{C}$ or some arbitrary monotonic function $f(\mathcal{C})$ and the corresponding threshold $f(\mathcal{C}_{\text{Threshold}})$, the min. time proof applies to all monotonic functions of $\mathcal{C}$, such as entanglement of formation.

Concurrence has the favorable property of being invariant under the control parameters, which makes it amenable to methods for proving global optimality. For many tasks however, a specific target state is desired, in which case fidelity is a more relevant figure of merit. We now show that the global optimality proofs given above extend to the corresponding fidelity goals.

We take any maximally entangled state $|\Psi\rangle$ to be the target state and define P_H^Ψ to be a variant of P_H which rotates ψ to have the maximal fidelity with respect to Ψ at the final time. For the max. fidelity goal, the final time is simply T. For the fidelity min. time goal, the final time is the earliest time at which $\mathcal{F}_{\text{Threshold}}$ can be reached. For both of these goals, we need to know the fidelity of a given state ψ maximized over all local unitaries. For pure states, this maximal fidelity is uniquely determined by the concurrence of ψ, and is given by [38]

$$\mathcal{F}_{\max} = \frac{\mathcal{C} + 1}{2}. \tag{21}$$

In the context of entanglement distillation, $\mathcal{F}_{\max}$ is often called the singlet fraction, and measures the usefulness of a general quantum state for the task.

To prove that P_H is globally optimal for the fidelity min. time goal, we note that any optimal protocol must have $\mathcal{F} = \mathcal{F}_{\max}(\mathcal{C})$ at the hitting time. If this were not the case, the protocol could have applied local unitaries to ψ at an earlier time that increase $\mathcal{F}(\psi)$ to $\mathcal{F}_{\max}(\mathcal{C}(\psi))$, and hence achieve an earlier hitting time. As $\mathcal{F}_{\max}$ is a monotonic function of $\mathcal{C}$, this relation suffices to prove

optimality of P_H with respect to the fidelity min. time goal when $\mathcal{F}_{\text{Threshold}} \le \mathcal{F}_{\max}(1/\sqrt{2}) = (1+\sqrt{2})/\sqrt{8}$.

To prove global optimality of the max. fidelity goal, we assume the evolution of the system is well-approximated by a discrete protocol in which one measures for a small but finite duration before applying feedback [19]. As realistic implementations of feedback inevitably suffer from feedback delay and finite bandwidth effects at sufficiently short timescales, the continuum limit may not be a good physical model, and we shall therefore not focus on this here. In the discrete approximation, we can express the expectation value of the fidelity at the stopping time T as an integral over all possible measurement outcomes $\{\mathbf{V}\}$. Suppose that some hypothetical feedback protocol P' is globally optimal for the max. fidelity goal. We write

$$\langle \mathcal{F}(P',T) \rangle = \int \mathcal{F}_{\mathbf{V}}(P',T) p(\mathbf{V}) d\mathbf{V}$$

$$= \int \mathcal{F}_{\max}(\mathcal{C}_{\mathbf{V}}(P',T)) p(\mathbf{V}) d\mathbf{V} \qquad (22)$$

where $\mathcal{F}_{\mathbf{V}}(P',T)$ and $\mathcal{C}_{\mathbf{V}}(P',T)$ are respectively the fidelity and concurrence at time T, assuming protocol P' was applied and measurement outcome $\mathbf{V}$ occurred. $p(\mathbf{V})$ is the probability of $\mathbf{V}$ occurring (note that p implicitly depends on P' and the state evolution, but this is of no consequence for the proof). The last equality follows from the fact that the globally optimal protocol could perform at least as well if it maximizes the fidelity of each possible final state. If it does not do so on a set of non-zero measure, then a better protocol exists which does. We continue by relating the performance of P' to that of P_H^Ψ:

$$= \int \frac{\mathcal{C}_{\mathbf{V}}(P',T)+1}{2} p(\mathbf{V}) d\mathbf{V}$$

$$\le \int \frac{\mathcal{C}_{\mathbf{V}}(P_H^\Psi,T)+1}{2} p(\mathbf{V}) d\mathbf{V}$$

$$= \int \mathcal{F}_{\max}(\mathcal{C}_{\mathbf{V}}(P_H^\Psi,T)) p(\mathbf{V}) d\mathbf{V} = \int \mathcal{F}_{\mathbf{V}}(P_H^\Psi,T) p(\mathbf{V}) d\mathbf{V}$$

$$= \langle \mathcal{F}(P_H^\Psi,T) \rangle. \qquad (23)$$

The inequality follows from global optimality of P_H with respect to the max. concurrence goal. Note that relating the performance of P' to that of P_H relies on the fact that $\mathcal{F}_{\max}$ is linear in $\mathcal{C}$. The second-to-last equality follows from the fact that at time T, P_H^Ψ rotates ψ to have maximum fidelity with respect to the target state. Together Eqs. (22) and (23) imply that P_H^Ψ performs at least as well as any potential protocol P', and therefore that P_H^Ψ is globally optimal for the task.

Full-parity measurement: A connection between Jacobs' rapid purification protocol, and rapid entanglement using full-parity measurement was first established in Ref. [14]. To prove optimality of the full-parity protocol P_F, we observe that the dynamics of the concurrence of a two-qubit pure state under full-parity measurement and local feedback are precisely those of the Bloch vector length of a single continuously monitored qubit with

feedback. We then use existing optimality results regarding rapid qubit purification [18] to prove the analogous two-qubit results.

To see the correspondence between 2-qubit concurrence $\mathcal{C}$ and the Bloch vector length of a single qubit, consider a qubit undergoing a continuous measurement of $\tilde{\sigma}_z$ at a rate $\tilde{\Gamma}$. We assume that some feedback controller can instantly apply any unitary operation at any time, as we have assumed for the two-qubit case. Parameterizing the qubit as a Bloch vector in spherical coordinates $\{\tilde{r}, \tilde{\theta}, \tilde{\phi}\}$, one can use Eq. (1) to derive an equation of motion for the Bloch vector length $\tilde{r}$ as a function of $\tilde{\theta}$ and $\tilde{\phi}$ [20]

$$d\tilde{r} = (1-\tilde{r}^2)\left[\frac{\tilde{\Gamma}}{4\tilde{r}}(1-\tilde{u}^2)dt + \sqrt{\frac{\tilde{\Gamma}}{2}}\tilde{u}\,dW \right] \qquad (24)$$

where $\tilde{u} = \cos(\tilde{\theta})$ and $\tilde{\theta}$ is the angle the Bloch vector makes with the measurement axis $\tilde{\sigma}_z$. We do not derive equations of motion for $\tilde{\theta}$ and $\tilde{\phi}$ because we assume feedback can set them to their desired values at any time. Making the following identifications

$$\mathcal{C} \to \tilde{r}$$
$$w \to \tilde{u}$$
$$2(u^2-v^2)^2 \to \tilde{\Gamma}, \qquad (25)$$

the equation of motion for concurrence under full-parity measurement and feedback, Eq. (10) becomes Eq. (24). This mapping reveals several interesting features of the dynamics. As observed in Ref. [14], entanglement generation can be turned on and off by rotating the system into one of the decoherence free subspaces of the measurement operator X_F. This is evident from the dependence of $\tilde{\Gamma}$ on u and v, as $u = v = 1$ and $u = v = 0$ correspond to states fully localized to the even and odd parity subspaces, respectively. Of particular interest is the direct mapping between $\mathcal{C}$ and $\tilde{r}$; the concurrence coincides exactly with Bloch vector length of the effective qubit. We provide more details on this effective qubit in the appendix.

Although the effective qubit we consider here is different from that discussed in Ref. [14], rapid entanglement generation corresponds to applying Jacobs' protocol to the effective qubit in both mappings. $\tilde{u} = 0$ is globally optimal for maximizing the linear entropy $\langle \tilde{r}(T) \rangle$ as shown in [18]. Although that work did not consider allowing the measurement rate $\tilde{\Gamma} \le 2$ to vary, it is straightforward to extend the proofs of [18] to this case by repeating their calculation with $\tilde{\Gamma}$ as a control variable bounded from 0 to 2. Thus the mapping between equations of motion for $\tilde{r}$ and $\mathcal{C}$ implies that P_F is globally optimal for the max. concurrence goal.

It is also the case that not applying feedback to the effective qubit yields the same equations of motion as not applying feedback to the two-qubits. Since it was proved in [18] that not applying feedback is globally optimal for

the min. time goal for linear entropy, the analogous result applies to the full-parity measurement of two qubits for the concurrence min. time goal. The arguments used to extend the max. concurrence and the concurrence min. time results of P_H to the corresponding fidelity results (*i.e.* Eqs. 21-23) also apply to P_F without modification. This completes the proof of the enumerated results given in Sec. 2.

Upper-bound on measurement-based protocols: So far, we have focused on optimality given a fixed measurement operator. However, one may ask whether a different measurement operator could offer superior performance using similar resources. In the context of remote entanglement generation, when entanglement is created using some signal degree of freedom as an intermediary (see Fig. 2), one could ask whether entanglement is transferred with unit efficiency. Motivated by these questions, we now prove a more general result which sets an upper limit on the entanglement entropy achievable generally under a much larger class of measurement-based protocols. We will find that the bound is only saturated by the full-parity feedback protocol P_F.

We consider any system in which the action of the measurement on qubit 1 is of the form

$$d\rho_1 = (\Gamma_{\text{deph.}}/2)\mathcal{D}[\vec{\sigma} \cdot \hat{n}]\rho_1 \ dt \qquad (26)$$

where ρ_1 is the state of qubit 1 unconditioned on the measurement outcome and tracing out qubit 2. Physically, the dephasing rate $\Gamma_{\text{deph.}}$ sets an upper bound on the amount of information that can be extracted from the measurement [39], so this restriction fixes the rate at which information about qubit 1 is transferred to the rest of the system. Thus $\Gamma_{\text{deph.}}$ defines a physically meaningful reference that lets us compare the performance of P_H and P_F to more general protocols. By tracing out qubit 2 in Eq. (1), one arrives at Eq. (26) with $\Gamma_{\text{deph.}} = 1/2$ for measurement of both X_H and X_F.

The entanglement entropy of a state is given by $E_1 = \text{Tr}[\rho_1 \log_2(\rho_1)]$, the Von Neumann entropy of qubit 1. Note that if the joint state is pure, then this expression also equals the entanglement of formation. In order for the entanglement to increase, the entropy of the subsystem must increase, as it does under the action of Eq. (26). As measurement constitutes the only available non-local interaction, Eq. (26) fully determines how entanglement may change. At time t, the increase in entropy of ρ_1 is maximized if the state is unbiased with respect to the measurement axis $\vec{\sigma} \cdot \hat{n}$. For example, if $\vec{\sigma} \cdot \hat{n} = \sigma_z$, then optimal states would be of the form $\rho_1 = (x\sigma_x + y\sigma_y + \sigma_0)/2$ where σ_0 is the identity matrix. Assuming this condition is satisfied at all times, the entropy as a function of time may be derived by solving Eq. (26) with the initial condition $\rho_1 = (|0\rangle + |1\rangle)(\langle 0| + \langle 1|)/2$. Taking $\Gamma_{\text{deph.}} = 1/2$,
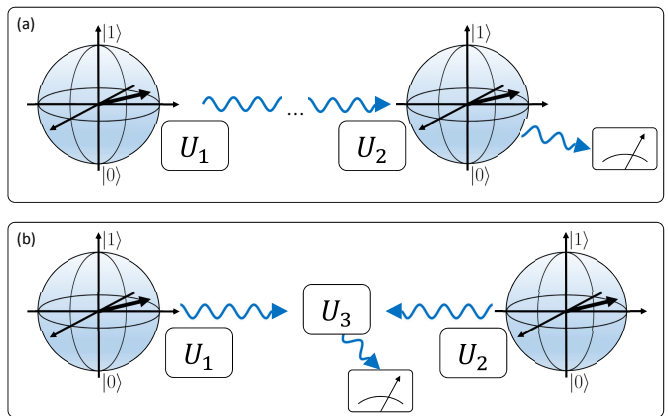

FIG. 2: (Color online) Remote 2-qubit measurement implementations. (a) Remote 2-qubit measurement in which the signal degrees of freedom (wavey line) propagate from qubit 1 to qubit 2. U_1 and U_2 represent the interactions between the signal and each qubit, while the dial represents projective measurement of the signal after qubit interactions. These three operations determine the effective 2-qubit measurement operator. (b) Remote measurement based on entanglement swapping. U_3 is the interaction between the incoming signals (such as a beam splitter) that erases which-path information. Note that in both schemes, the rate of entanglement generation is bounded by U_1.

we find

$$E_1(t) = -\frac{1 - \langle\sigma_{x1}\rangle}{2} \log_2\left[\frac{1 - \langle\sigma_{x1}\rangle}{2}\right] \qquad (27)$$
$$-\frac{1 + \langle\sigma_{x1}\rangle}{2} \log_2\left[\frac{1 + \langle\sigma_{x1}\rangle}{2}\right]$$
$$\langle\sigma_{x1}\rangle = e^{-t/2}. \qquad (28)$$

This Von Neumann entropy sets an upper bound on the entanglement entropy that can be achieved by the action of a given measurement operator (as well as on the entanglement of formation if the joint evolution is pure).

The main result of this section is that by tracing out qubit 2 from Eq. (5) and substituting Eq. (11), one can show that the entanglement entropy of qubit 1 is exactly Eq. (27). For pure states, a similar result can be derived for concurrence, which has a one-to-one relation with entanglement entropy in this case [28]. Thus P_F saturates the bound Eq. (27) given by the dephasing rate of X_F, and hence no measurement-based protocol of the form Eq. (26) can generate entanglement faster than P_F. It can also be shown that P_H does not saturate this bound, nor does measurement of X_F or X_H without feedback.

This bound may be intuitively be understood by considering a remote implementation of the measurement, as depicted in Fig. 2. Eq. (26) then governs how much entanglement is generated between qubit 1 and the ancilla. An ideal entangling protocol would transfer all of this entanglement to qubit 2, so that measurement of the ancilla does not decrease the entanglement entropy. Evidently only P_F fully transfers the entanglement from the

measurement signal to qubit 2. For the other protocols, the remaining entanglement is destroyed when the signal is measured.

Note that if the effect of measurement on qubit 1 yields a dephasing operator that is not normal (*i.e.*, $[X, X^\dagger] \neq 0$), then it cannot be put in the form of Eq. (26) and our derivation does not apply. Thus our bound does not apply for quantum-demolition measurements such as a spontaneous emission process on the first qubit *i.e.*, $d\rho_1 = \mathcal{D}[\sigma]\rho_1 \, dt$, though a similar bound could be derived for such cases. An example of the latter is entanglement generation via Hong-Ou-Mandel interference.

5. CONCLUSION

We have derived equations of motion for the concurrence of a two-qubit state under non-local measurement and arbitrary local feedback control. These equations provide a simple and general method for deriving feedback protocols, allowing us to rederive a feedback protocol for deterministic entanglement generation using a half-parity measurement, and to derive a protocol for faster entanglement of initially separable pure states using a full-parity measurement. Parameterization of the state in terms of an entanglement monotone and local operations has allowed us to prove several global optimality results in both cases. In studying the full-parity protocol, our work elaborates on the connections between rapid purification and rapid entanglement generation of Ref. [14] by quantitatively linking the dynamics of the Bloch vector length to that of the concurrence. We have also derived a bound on rapid entanglement generation given by the dephasing of a single measurement, which is saturated only by the full-parity feedback protocol, indicating that it transfers entanglement with unit efficiency.

Our work suggests several general avenues for further study. Regarding the mapping between a two-qubit pure state and a one qubit mixed state, it is unclear if this connection is an isolated case or if it is indicative of a more general set of mappings between pure states and mixed states of smaller systems. If such mappings did exist, they could potentially simplify the analysis of multipartite entangled systems. Several further open questions relate to entanglement generation and feedback. Our optimality results for the min. time goals involving the half-parity measurement only apply when the target concurrence is less than $1/\sqrt{2}$. Optimal protocols for the remaining parameter regime could be found by numerical searches. Our work could also be extended to consider optimality under many other goals and cost functions, such as entanglement of formation or concurrence of the ensemble density matrix unconditioned on the measurement record.

There are also numerous possible extensions and variations to the physical systems that we have considered so far. The full-parity scheme generates entanglement very rapidly at short times, but only reaches unit concurrence asymptotically, unlike a simple CNOT gate. Combining entangling unitaries of finite strength with measurement and feedback might yield even faster entanglement generation. Parity measurements can also be useful for generating a variety of entangled states, from the three-particle GHZ and W states to entangled many-body states such as cluster states and the Surface code ground state [2]. Feedback could be used to generate such states more efficiently and perhaps optimally.

Furthermore, considerable experimental and theoretical work has focused on the task of generating non-classical states using dissipation alone [12, 40–44]. Given a fixed coupling strength to the environment, one could ask what dissipation channels are optimal for entanglement generation, or how to apply feedback if the dissipation channels are monitored [45]. The latter is particularly amenable to our techniques, since it could involve pure-state dynamics.

Because the measurements we have studied are candidates for remote entanglement generation, studying optimality in the presence of loss remains a crucial detail to be considered in future work. In this context, entanglement distillation will often be required [46], and will determine which cost functions are relevant in practice. For example, given a specific distillation protocol, one could ask to generate a distilled qubit with sufficiently high concurrence in as little time as possible, which would in turn dictate the cost function to use when optimizing the entanglement generation process. Other natural considerations include other experimental defects such as finite qubit coherence, feedback delay and finite measurement efficiency.

Appendix A: Details on the derivations of Eqs. 6 and 10

To derive Eq. (6), we use Eq. (4) and divide the quantity inside the absolute value into three terms, the initial concurrence $\mathcal{C}(t)$, the change in the real part $d\mathcal{C}_\mathbb{R}$ and the change in the imaginary part $d\mathcal{C}_\mathbb{I}$. We have chosen the global phase of ψ in Eq. (5) so that $\mathcal{C}(t)$ appears in the expression without a phase factor.

$$\mathcal{C}(t+dt) = \Big|\langle\psi(t+dt)^*|\sigma_y\otimes\sigma_y|\psi(t+dt)\rangle\Big| \quad \text{(A1)}$$

$$= |\mathcal{C}(t) + d\mathcal{C}_\mathbb{R} + id\mathcal{C}_\mathbb{I}|$$

$$d\mathcal{C}_\mathbb{R} = 2\mathcal{C}(t)\sqrt{1-\mathcal{C}(t)^2}uv\ dW$$
$$+ [(v^2-u^2)\cos(2\phi) - \mathcal{C}(t)(v^2+u^2)]\ dt$$

$$d\mathcal{C}_\mathbb{I} = \sqrt{1-\mathcal{C}(t)^2}(v^2-u^2)\sin(2\phi)\ dt$$

If $\mathcal{C}\neq 0$, we can use Ito's lemma and the standard relation $|z| = \sqrt{\mathrm{Re}[z]^2 + \mathrm{Im}[z]^2}$ to simplify the above expression.

$$d\mathcal{C} = 2\mathcal{C}\sqrt{1-\mathcal{C}^2}uv\ dW \qquad\qquad \text{(A2)}$$
$$+ [(v^2-u^2)w - \mathcal{C}(v^2+u^2)]dt \qquad \mathcal{C}>0$$

The $\mathcal{C}=0$ case must be treated separately, as $|z|$ is not twice-differentiable at the origin as required by Ito's lemma. At $\mathcal{C}=0$, ψ is independent of the value of ϕ up to a global phase. Thus the evolution of $\mathcal{C}$ here cannot be affected by the value of ϕ, and hence we can restrict the value of the latter without loss of generality. We can avoid taking the absolute value and directly assign $d\mathcal{C}|_{\mathcal{C}=0} = d\mathcal{C}_\mathbb{R}$ by enforcing $\sin(2\phi)=0$, so that $d\mathcal{C}_\mathbb{I}=0$. To keep $d\mathcal{C}$ non-negative, we choose ϕ such that $\cos(2\phi)=1$ if $(v^2-u^2)>0$ and $\cos(2\phi)=-1$ otherwise, which yields

$$d\mathcal{C} = |v^2-u^2|\ dt \qquad \mathcal{C}=0. \qquad \text{(A3)}$$

This completes the derivation of Eq. (6).

There is one subtle difference between the derivations of Eq. (10) and Eq. (6), which accounts for the fact that we have allowed $\mathcal{C}$ to take on negative values in Eq. (10). In the full-parity case, $d\mathcal{C}_{\mathbb{R}/\mathbb{I}}|_{\mathcal{C}=0}$ depend on dW, so it is not possible to choose ϕ to maintain $d\mathcal{C}_\mathbb{R} \geq 0$. Even mathematically it does not make sense to choose ϕ to anticipate the value of dW, as the resulting integrand would not be an adapted process, and hence the Ito integral would not be defined [34]. To remedy this technicality, we set $\phi|_{\mathcal{C}=0}=0$ and allow $\mathcal{C}$ to take on negative values (as before, $\phi=0$ sets $d\mathcal{C}_\mathbb{I}=0$). As $\psi(\mathcal{C})=\psi(-\mathcal{C})$, one may interpret a state with $\mathcal{C}<0$ as having concurrence $|\mathcal{C}|$. This step yields the $\mathcal{C}=0$ case of Eq. (10). To make sure the equations of motion for $\mathcal{C}$ are consistent with the underlying equations of motion for ψ, we take $\mathcal{C}(t+dt) = \mathrm{sgn}(\mathcal{C}(t))|\mathcal{C}(t) + d\mathcal{C}_\mathbb{R} + id\mathcal{C}_\mathbb{I}|$ when calculating the $\mathcal{C}\neq 0$ case of Eq. (10), where $\mathrm{sgn}(x)$ is the sign function, defined to be ± 1 for $x \gtrless 0$ and 0 otherwise.

Appendix B: Feedback Hamiltonian for the full-parity protocol and its relation to that of Hill *et al.*[14]

In section 3, we derived a protocol for rapid entanglement generation using local unitary feedback and a full-parity measurement. This protocol is quite similar to that given in Ref. [14], but turns out to be distinct from it in several important respects. In this section, we first derive how to implement our protocol using Hamiltonian feedback, and then outline the similarities and differences between it and the protocol of Ref. [14].

Using the Schmidt decomposition Eq. (5), P_F specifies the state evolution $\psi(t)$ without giving the Hamiltonian feedback necessary to induce this particular evolution. We now use $\psi(t)$ to derive the required Hamiltonian. We assume as is often the case that the protocol can be implemented using proportional feedback, meaning that the controller applies the Hamiltonian $H_F\ dV$. In this case, the equation of motion for the system is the Wiseman-Milburn feedback master equation [31]

$$d\psi = \Big[\frac{-1}{2}\big(X_F - \langle X_F\rangle\big)^2 dt + \big(X_F - \langle X_F\rangle\big)dW \qquad \text{(B1)}$$
$$+ \Big[\frac{-i}{2}H_F\big(X_F + \langle X_F\rangle\big) - \frac{H_F^2}{8}\Big]dt - \frac{i}{2}H_F\ dW\Big]\psi$$

where we have set $\hbar = 1$. The first two terms match Eq. (3) and correspond to continuous measurement of X_F. The remaining terms give the evolution under feedback. We know from the derivation of P_F that H_F consists entirely of local rotations if it exists (see Sec. 3). To calculate H_F, we first express it in the Pauli basis

$$H_F = h_0 I + \sum_{i=\{1,2\}} h_{x,i}\sigma_{x,i} + h_{y,i}\sigma_{y,i} + h_{z,i}\sigma_{z,i} \quad \text{(B2)}$$

where I is the identity operator. We have added the physically meaningless constant h_0 fit the evolution of the global phase of $\psi(t)$, which is already implicitly specified. Substituting Eq. (B2) and $\psi(t)$ given by Eq. (12) into Eq. (B1), we solve for $h_{i,j}$. The calculation is simplified by the fact that $\langle X_F\rangle = 0$, and may be further simplified by taking advantage of the fact that $\psi(t)$ evolves deterministically, and hence the dW terms of Eq. (B1) must cancel. Physically, this means that the feedback Hamiltonian is chosen to exactly cancel the stochastic back-action from the measurement. Thus we begin by finding $h_{i,j}$ such that

$$\Big(X_F - \frac{i}{2}H_F\Big)\psi(t) = 0. \qquad \text{(B3)}$$

For $\Delta\gamma = \gamma = 0$, $u = 0$ and $v = 1$, Eq. (B3) admits the solution

$$h_{x,1} + h_{x,2} = \frac{-1}{|\mathcal{C}|},\ h_{y,i} = h_{z,i} = 0$$

$$h_0 = -\frac{1-\mathcal{C}^2}{|\mathcal{C}|}. \qquad \text{(B4)}$$

Solutions for other values of γ and $\Delta\gamma$ may be obtained by a local change of basis of H_F. Substituting the solution for $\mathcal{C}(t)$ of Eq. (11) into Eq. (B4), it can be verified that the dt terms of Eq. (B1) are also satisfied by Eq. (B4). Thus for $\mathcal{C}\neq 0$, P_F may be implemented by applying local σ_x rotations by an angle proportional to

the measurement outcome dV. Notice that from Eq. (5), this is about an axis orthogonal to that of the $\pi/2$ pulses required to prepare the initial state assuming one starts in the ground state $|11\rangle$. Also note that $h_{x,1} = 0$, $h_{x,2} = 0$ and $h_{x,1} = h_{x,2} = 1/2\mathcal{C}$ are all solutions, indicating that one may implement P_F by applying rotations on only one of the two qubits, or by applying the same rotation to both. The former is useful for remote entanglement; after the measurement signal interacts with qubit 2 and is read out by the controller, the controller can apply feedback only to qubit 2, eliminating the detrimental effect of propagation delay from qubit 1 to the controller.

For $\mathcal{C} = 0$, feedback apparently requires pulses that rotate the qubits by a finite angle. The fact that $\langle X_F \rangle = 0$ allows us to determine these pulses. It can be shown that any σ_x rotations satisfying $\theta_1 + \theta_2 = -\text{sgn}(dV)\pi/2$ induce the desired evolution. Comparing with Jacobs' rapid qubit purification protocol [16], we see that the feedback for P_F takes exactly the same form.

P_F bears many similarities to the protocol of Ref. [14]. Both involve applying Jacobs' protocol to an encoded qubit, and the concurrence as a function of time is the same. Furthermore, they can both be implemented using only single qubit rotations. P_F is defined using the pure state Schmidt decomposition, and therefore is not defined for mixed states. The protocol of Ref. [14] is defined by applying Jacobs' purification protocol to an effective qubit encoded via the mapping

$$
\begin{aligned}
\sigma'_x &\to \sigma_{x2} \\
\sigma'_y &\to \sigma_{z1}\sigma_{y2} \\
\sigma'_z &\to \sigma_{z1}\sigma_{z2}.
\end{aligned}
\tag{B5}
$$

The authors apply this protocol to a classically correlated mixture, which simultaneously purifies and entangles the qubits. To compare with P_F, we substitute Eq. (11) into Eq. (B5)

$$
\begin{aligned}
\langle \sigma'_x \rangle &= -\sqrt{1-\mathcal{C}^2}\cos(\gamma - \Delta\gamma) \\
\langle \sigma'_y \rangle &= \mathcal{C}\cos(\gamma - \Delta\gamma) \\
\langle \sigma'_z \rangle &= 0.
\end{aligned}
\tag{B6}
$$

P_F does indeed maintain $\langle \sigma'_z \rangle = 0$ like Jacobs' rapid purification protocol, but it's action on this effective qubit is not to purify it, but rather to rotate it about the σ'_z axis. Its purity remains constant, determined by the initial preparation of the qubits. When $\tilde{\Gamma} = 2$, it is natural to consider $\tilde{r}$ of Eq. (25) to represent two effective qubits encoded as

$$
\begin{aligned}
\langle \tilde{\sigma}_x^1 \rangle &\to -\langle \sigma_{x1}\sigma_{z2} \rangle = \mathcal{C}\sin(\gamma_1)\sin(2\phi) \\
\langle \tilde{\sigma}_y^1 \rangle &\to \langle \sigma_{y1}\sigma_{z2} \rangle = \mathcal{C}\cos(\gamma_1)\sin(2\phi) \\
\langle \tilde{\sigma}_z^1 \rangle &\to -\langle \sigma_{z1}\sigma_{z2} \rangle = \mathcal{C}\cos(2\phi)
\end{aligned}
\tag{B7}
$$

$$
\begin{aligned}
\langle \tilde{\sigma}_x^2 \rangle &\to -\langle \sigma_{z1}\sigma_{x2} \rangle = \mathcal{C}\sin(\gamma_2)\sin(2\phi) \\
\langle \tilde{\sigma}_y^2 \rangle &\to \langle \sigma_{z1}\sigma_{y2} \rangle = \mathcal{C}\cos(\gamma_2)\sin(2\phi) \\
\langle \tilde{\sigma}_z^2 \rangle &\to -\langle \sigma_{z1}\sigma_{z2} \rangle = \mathcal{C}\cos(2\phi).
\end{aligned}
\tag{B8}
$$

where we have set $\theta = \pi/2$ and $\Delta\theta = 0$, although other allowed values yield similar results. Note that the operators on the left-hand side of each equality together obey the Pauli commutation relations $[\sigma_a, \sigma_b] = 2i\epsilon_{abc}\sigma_c$ and therefore represent valid encoded qubits (although they are not independent, as they share the same z component i.e., $\langle \tilde{\sigma}_z^1 \rangle = \langle \tilde{\sigma}_z^2 \rangle$). The right-hand sides are functions of local unitary operations on the physical qubits and coincide with a Bloch sphere representation in spherical coordinates when we make the mapping $\{\mathcal{C}, 2\phi, \gamma_1\}$ or $\{\mathcal{C}, 2\phi, \gamma_2\} \to \{\tilde{r}, \tilde{\theta}, \tilde{\phi}\}$.

To conclude, we show that even though the protocols involve different initial states, the performance as measured by the concurrence is equivalent. This equivalence is somewhat non-trivial since the expression for the concurrence of mixed states is a non-trivial generalization of that for pure states. Our analysis also uncovers further explanation for the connection between rapid purification and rapid entanglement generation.

We begin again with the Wiseman-Milburn equation for unit efficiency measurement, this time written as a master equation

$$
\begin{aligned}
d\rho =& \mathcal{D}[X_F]\rho\,dt + \mathcal{H}[X_F]\rho\,dW - i[H_F, \rho]\frac{dW}{2} \\
& -i[H_f, X_F\rho + \rho X_F]\frac{dt}{2} + \mathcal{D}[H_F]\rho\frac{dt}{4}.
\end{aligned}
\tag{B9}
$$

We follow the prescription given in [14], starting with the mixed state $\rho_0 = (|00\rangle\langle00| + |11\rangle\langle11|)/2$. This state lies within a decoherence-free subspace of X_F, so Hadamard rotations are first applied to both qubits so that the measurement acts non-trivially on the state. Jacobs' rapid purification protocol is then applied on an encoded qubit by applying local σ_x rotations on qubit 1. For this, we set $H_F = P\sigma_{x,1}/2$ where P is to be determined. Like Jacobs' protocol, the optimal feedback rotation is then determined by requiring $\langle X_F \rangle = 0$ after each application of feedback, which keeps the state maximally unbiased with respect to the measurement operator. The resulting evolution preserves spin flip symmetry $(\rho \to \sigma_y \otimes \sigma_y \rho^* \sigma_y \otimes \sigma_y)$, exchange symmetry, and is deterministic, which makes the dynamics easy to find. The state evolution starting in ρ_0 may be fully parameterized by one variable as

$$
\rho(t) = \begin{bmatrix} 1/4 & \alpha(t) & \alpha(t) & 1/4 \\ -\alpha(t) & 1/4 & 1/4 & -\alpha(t) \\ -\alpha(t) & 1/4 & 1/4 & -\alpha(t) \\ 1/4 & \alpha(t) & \alpha(t) & 1/4 \end{bmatrix}
\tag{B10}
$$

with $\alpha(t)$ purely imaginary. Note that from Eq. (B5), one can show that $\alpha(t)$ is proportional to the Bloch vector length of the effective qubit considered in Ref. [14]. This evolution follows from $P = i/(2\alpha)$, which may be determined by requiring $\langle X_F \rangle = 0$ or by searching for a deterministic solution. Using Eq. (B9), one finds the solution $\alpha(t) = -i\sqrt{1 - e^{-t}}/4$

Because $\rho(t)$ is spin flip symmetric, its concurrence is directly related to its eigenvalues. Thus purification and entanglement generation are nearly equivalent tasks. If the eigenvalues of ρ are $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ written in decreasing order, then the concurrence is $\max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}$ and the purity is $\sum_i \lambda_i^2$. Both functions are large when the magnitude of one eigenvalue dominates, and thus protocols for entanglement and purification should attempt to maximize the one eigenvalue of ρ while suppressing the others. Using the former equation, we find $\mathcal{C}(t) = \sqrt{1 - e^{-t}}$, which coincides with the solution Eq. (11) with $\mathcal{C}(0) = 0$.

[1] A. Broadbent, J. Fitzsimons, and E. Kashefi, in *Foundations of Computer Science, 2009. FOCS'09. 50th Annual IEEE Symposium on* (IEEE, 2009), pp. 517–526.

[2] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Phys. Rev. A **86**, 032324 (2012).

[3] M. H. Devoret and R. J. Schoelkopf, Science **339**, 1169 (2013).

[4] N. Nickerson, Y. Li, and S. Benjamin, Nat. Commun. **4**, 1756 (2013).

[5] F. Motzoi, K. B. Whaley, and M. Sarovar, Phys. Rev. A **92**, 032308 (2015).

[6] N. Roch, M. E. Schwartz, F. Motzoi, C. Macklin, R. Vijay, A. W. Eddins, A. N. Korotkov, K. B. Whaley, M. Sarovar, and I. Siddiqi, Phys. Rev. Lett. **112**, 170501 (2014).

[7] J. Hofmann, M. Krug, N. Ortegel, L. Gérard, M. Weber, W. Rosenfeld, and H. Weinfurter, Science **337**, 72 (2012).

[8] H. Bernien, B. Hensen, W. Pfaff, G. Koolstra, M. Blok, L. Robledo, T. H. Taminiau, M. Markham, D. Twitchen, L. Childress, et al., Nature **497**, 86 (2013).

[9] O. Černotík and K. Hammerer, arXiv preprint arXiv:1512.00768 (2015).

[10] A. Roy, L. Jiang, A. D. Stone, and M. Devoret, Phys. Rev. Lett. **115**, 150503 (2015).

[11] M. Silveri, E. Zalys-Geller, M. Hatridge, Z. Leghtas, M. H. Devoret, and S. Girvin, Phys. Rev. A **93**, 062310 (2016).

[12] F. Motzoi, E. Halperin, X. Wang, K. B. Whaley, and S. Schirmer, Phys. Rev. A **94**, 032313 (2016).

[13] M. Sarovar, H.-S. Goan, T. P. Spiller, and G. J. Milburn, Phys. Rev. A **72**, 062327 (2005).

[14] C. Hill and J. Ralph, Phys. Rev. A **77**, 014305 (2008).

[15] L. Martin, F. Motzoi, H. Li, M. Sarovar, and K. B. Whaley, Phys. Rev. A **92**, 062321 (2015).

[16] K. Jacobs, Phys. Rev. A **67** (2003).

[17] H. M. Wiseman and J. Ralph, New J. Phys. **8**, 90 (2006).

[18] H. M. Wiseman and L. Bouten, Quantum Information Processing **7**, 71 (2008).

[19] C. Teo, J. Combes, and H. M. Wiseman, New J. Phys. **16**, 105010 (2014).

[20] H. Li, A. Shabani, M. Sarovar, and B. K. Whaley, Phys. Rev. A **87**, 032334 (2013).

[21] J. Combes and K. Jacobs, Phys. Rev. Lett. **96**, 010504 (2006).

[22] A. Shabani and K. Jacobs, Phys. Rev. Lett. **101**, 230403 (2008).

[23] J. Combes, H. M. Wiseman, K. Jacobs, and A. J. OConnor, Phys. Rev. A **82**, 022307 (2010).

[24] S. Sridharan, M. Yanagisawa, and J. Combes, arXiv preprint arXiv:1211.5617 (2012).

[25] A. Serafini and S. Mancini, Phys. Rev. Lett. **104**, 220501 (2010).

[26] S. Mancini and H. M. Wiseman, Phys. Rev. A **75**, 012330 (2007).

[27] M. G. Genoni, S. Mancini, and A. Serafini, Phys. Rev. A **87**, 042333 (2013).

[28] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998), ISSN 0031-9007.

[29] D. Ristè, M. Dukalski, C. A. Watson, G. de Lange, M. J. Tiggelman, Y. M. Blanter, K. W. Lehnert, R. N. Schouten, and L. DiCarlo, Nature **502**, 350 (2013).

[30] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).

[31] H. M. Wiseman and G. J. Milburn, *Quantum measurement and control* (Cambridge University Press, 2009).

[32] K. Jacobs and D. A. Steck, Contemp. Phys. **47**, 279 (2006).

[33] K. Jacobs, *Quantum measurement theory and its applications* (Cambridge University Press, 2014).

[34] B. Øksendal, *Stochastic Differential Equations* (Spinger-Verlag, 2003).

[35] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, 2010).

[36] K. Jacobs and P. Knight, Phys. Rev. A **57**, 2301 (1998).

[37] K. Jacobs and A. Shabani, Contemp. Phys. **49**, 435 (2008).

[38] F. Verstraete and H. Verschelde, Phys. Rev. A **66**, 022307 (2002).

[39] A. A. Clerk, S. M. Girvin, F. Marquardt, and R. J. Schoelkopf, Rev. Mod. Phys. **82**, 1155 (2010).

[40] S. Shankar, M. Hatridge, Z. Leghtas, K. Sliwa, A. Narla, U. Vool, S. M. Girvin, L. Frunzio, M. Mirrahimi, and M. H. Devoret, Nature **504**, 419 (2013).

[41] M. E. Kimchi-Schwartz, L. Martin, E. Flurin, C. Aron, M. Kulkarni, H. E. Tureci, and I. Siddiqi, Phys. Rev. Lett. **116**, 240503 (2016).

[42] F. Ticozzi and L. Viola, Automatica **45**, 2002 (2009).

[43] F. Ticozzi and L. Viola, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **370**, 5259 (2012).

[44] P. D. Johnson, F. Ticozzi, and L. Viola, arXiv preprint arXiv:1506.07756 (2015).

[45] Y. Liu, S. Shankar, N. Ofek, M. Hatridge, A. Narla, K. Sliwa, L. Frunzio, R. J. Schoelkopf, and M. H. Devoret, Phys. Rev. X **6**, 011022 (2016).

[46] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76** (1996).