

Enhancing qubit readout through dissipative sub-Poissonian dynamics

B. D'Anjou^{1,2} and W. A. Coish^{1,3,4}

¹*Department of Physics, McGill University, Montreal, Quebec, H3A 2T8, Canada*

²*Department of Physics, University of Konstanz, D-78464 Konstanz, Germany*

³*Quantum Information Science Program, Canadian Institute for Advanced Research, Toronto, Ontario, M5G 1Z8, Canada*

⁴*Center for Quantum Devices, Niels Bohr Institute, University of Copenhagen, 2100 Copenhagen, Denmark*

(Received 16 August 2017; published 15 November 2017)

Single-shot qubit readout typically combines high readout contrast with long-lived readout signals, leading to large signal-to-noise ratios and high readout fidelities. In recent years, it has been demonstrated that both readout contrast and readout signal lifetime, and thus the signal-to-noise ratio, can be enhanced by forcing the qubit state to transition through intermediate states. In this work, we demonstrate that the sub-Poissonian relaxation statistics introduced by intermediate states can reduce the single-shot readout error rate by orders of magnitude even when there is no increase in signal-to-noise ratio. These results hold for moderate values of the signal-to-noise ratio ($\mathcal{S} \lesssim 100$) and a small number of intermediate states ($N \lesssim 10$). The ideas presented here could have important implications for readout schemes relying on the detection of transient charge states, such as spin-to-charge conversion schemes for semiconductor spin qubits and parity-to-charge conversion schemes for topologically protected Majorana qubits.

DOI: [10.1103/PhysRevA.96.052321](https://doi.org/10.1103/PhysRevA.96.052321)

I. INTRODUCTION

High-fidelity single-shot readout of qubits is highly desirable for quantum information processing, and is crucial for achieving quantum error correction [1] and measurement-only quantum computation [2–4]. In the last decade, single-shot readout was realized experimentally for a large variety of promising qubit implementations. Prominent examples include superconducting qubits [5–10], trapped-ion and trapped-atom qubits [11–15], nitrogen-vacancy center spin and charge qubits in diamond [16–21], and semiconductor spin qubits [22–28]. In the near future, it should also be possible to perform readout of topologically protected Majorana qubits via parity-to-charge conversion [29,30].

The fidelity of these readout schemes is typically limited by spurious transitions, such as qubit relaxation, that limit the lifetime of the readout signal (i.e., the characteristic duration for which the two qubit states can be distinguished). As a result, the integration time is bounded and the signal-to-noise ratio (SNR) remains finite. Therefore, researchers have dedicated considerable effort to increasing both the readout signal lifetime and the readout contrast [9,18,20,27,28,31–33] with the aim of improving SNR and boosting readout fidelity.

One promising approach that can achieve either of these goals is to make the system transition through intermediate states. For instance, it was demonstrated that the effective readout signal lifetime for the optical readout of a nitrogen-vacancy center spin in diamond can be enhanced by forcing the system to undergo flip-flops between electron and nuclear spins before reaching the steady state [34]. Similarly, it was shown that both the readout signal lifetime and the readout contrast for singlet-triplet qubits in semiconductors can be significantly improved by mapping one of the qubit states to a long-lived metastable charge state that is easily detected by a nearby charge sensor [31,32]. These methods have since been used to bring the charge discrimination fidelities required for the readout of singlet-triplet qubits in silicon close to 99.9%, making them a serious candidate for fault-tolerant quantum computation [27,28].

In this article, we show that incoherent transitions through intermediate states can be used to significantly enhance readout fidelity even when they do not contribute to an increase in SNR. The key observation underpinning this result is that the times at which consecutive transition events occur are correlated. Consequently, the total time taken to transition through the intermediate states is not exponentially distributed as would be expected for a direct transition (see Fig. 1). Instead, the probability distribution of the total transition time becomes peaked as the number of intermediate states increases (see Fig. 2). This reflects the sub-Poissonian correlations resulting from a transition event being conditioned on the occurrence of the previous ones [35]. The randomness of the relaxation process is thus reduced, leading to an improved fidelity even in the case where the SNR is left unchanged. In particular, we show that the error probability may be reduced by orders of magnitude with the addition of a small number of intermediate states. These results identify an important resource for engineering of qubit readout that has so far been underappreciated.

The article is structured as follows. In Sec. II, we introduce the basic formalism necessary to analyze the fidelity of a single-shot qubit readout. In Sec. III, we review a minimal physical model of qubit readout that captures the effect of relaxation in the absence of intermediate states. The core of the article is Sec. IV, where we show that the addition of intermediate states in the relaxation process may enhance readout fidelity without requiring an improvement in SNR. We discuss the results in Sec. V and we conclude in Sec. VI.

II. QUBIT READOUT

A. Readout apparatus and outcomes

We consider a qubit with computational basis states $|+\rangle$ and $|-\rangle$. The qubit state is read out by making the qubit interact with some (typically macroscopic) readout apparatus. If the readout outcome \mathcal{O} registered by the apparatus is strongly dependent on the qubit state, single-shot qubit readout is

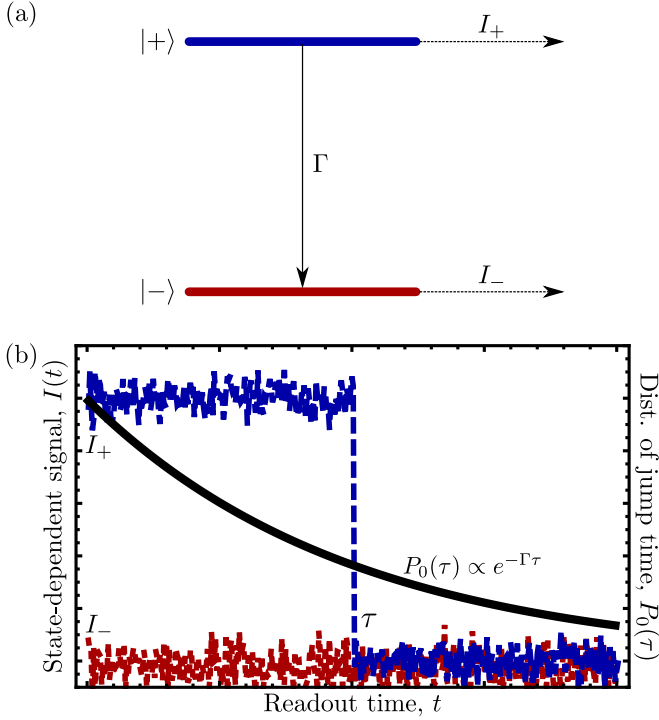


FIG. 1. Minimal model of qubit readout. (a) Level diagram of the minimal model. The detector registers a signal $I(t) = I_{\pm} + \delta I(t)$ when the qubit state is $|\pm\rangle$ at time t . Here $\delta I(t)$ is Gaussian white noise. The excited state $|+\rangle$ relaxes to the ground state with probability per unit time Γ . (b) Instances of the readout signal $I(t)$ for both qubit states. The average readout signal for $|-\rangle$ (dotted red line) is a constant I_- for all readout times $t > 0$. The average readout signal for $|+\rangle$ (dashed blue line) is initially I_+ but jumps to I_- at a random time $\tau > 0$ due to relaxation. The probability distribution $P_0(\tau)$ of the jump time τ is exponential (solid black line).

possible. The readout outcome \mathcal{O} may take various forms. For example, readout of superconducting qubits is often achieved by discriminating between two possible quadratures of a microwave tone [5–10], while readout of trapped-ion qubits and nitrogen-vacancy center spin or charge qubits in diamond is typically performed by differentiating between two fluorescence signals of distinct intensity [11–21]. Similarly, readout of semiconductor spin qubits is usually carried out by mapping the qubit states to charge states that produce different electrical currents in a nearby single-electron transistor, quantum dot, or quantum point contact [22–28]. These charge sensors could also be used to read out topologically protected qubits via a parity-to-charge conversion mechanism [29,30].

B. Average readout error rate

An ideal readout apparatus yields distinct readout outcomes when the qubit is prepared in either of the computational basis states. This means that the two probability distributions $P(\mathcal{O}|\pm)$ of the outcome conditioned on the initial state being $|\pm\rangle$ should not overlap. It can then be unambiguously decided which of the two computational basis states generated the readout outcome. Such a measurement is said to be accurate or sharp [36]. In practice, however, the probability distributions $P(\mathcal{O}|+)$ and $P(\mathcal{O}|-)$ have a finite overlap. Consequently, there

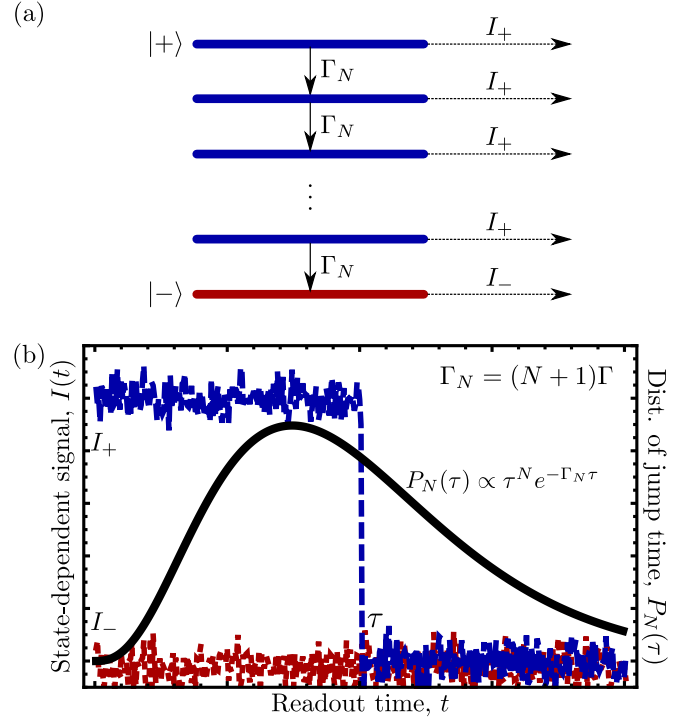


FIG. 2. Model of qubit relaxation through N intermediate states. (a) The system transitions from $|+\rangle$ to $|-\rangle$ by cascading through the intermediate states. All transitions in the sequence occur with probability per unit time $\Gamma_N = (N + 1)\Gamma$. It is assumed that all states except $|-\rangle$ yield the same average readout signal I_+ . (b) Instances of the readout signal for both qubit states. As before, the average readout signal for $|-\rangle$ (dotted red line) is a constant I_- . The average readout signal for $|+\rangle$ (dashed blue line) is initially I_+ but jumps to I_- at a time $\tau > 0$ when the system transitions from the last intermediate state to $|-\rangle$. Contrary to Fig. 1, however, the jump time τ now follows the Gamma distribution $P_N(\tau)$ (solid black line).

is a finite probability ϵ_{\pm} that the state $|\pm\rangle$ will be incorrectly identified as $|\mp\rangle$. We refer to the probabilities ϵ_{\pm} as the conditional single-shot readout error rates. The accuracy of the readout can then be quantified using the average single-shot readout error rate

$$\epsilon = \frac{\epsilon_+ + \epsilon_-}{2}. \quad (1)$$

The average single-shot readout fidelity $F = 1 - \epsilon$ is also commonly quoted as a measure of readout accuracy. Note that in Eq. (1), the conditional error rates ϵ_{\pm} are weighed equally. Thus, the quantity ϵ is the probability of error in the case where $|+\rangle$ and $|-\rangle$ occur with equal probability *a priori*. Information processing protocols are often designed to yield such equal prior probabilities to extract a maximum of information from the measurement. In the general case, Eq. (1) still provides a useful figure of merit for the accuracy of readout that has the advantage of being independent of the particular application at hand.

C. Maximum-likelihood decision

The average single-shot readout error rate ϵ , Eq. (1), is completely determined by the readout outcome distributions

$P(\mathcal{O}|\pm)$. To relate ϵ to these distributions, it is necessary to specify a decision rule mapping the readout outcome \mathcal{O} to one of the computational basis states. The decision rule that minimizes ϵ , Eq. (1), is known as a maximum-likelihood decision rule [37]. A maximum-likelihood decision compares the likelihood ratio $\Lambda = P(\mathcal{O}|+)/P(\mathcal{O}|-)$ to unity

$$\Lambda > 1 \rightarrow |+\rangle, \quad \Lambda < 1 \rightarrow |-\rangle. \quad (2)$$

The above decision rule minimizes Eq. (1) because it is a special case of a maximum *a posteriori* decision rule [37]. A maximum *a posteriori* decision maximizes the posterior probability $P(\pm|\mathcal{O})$ over the qubit states. It is thus manifestly optimal. When the two qubit states are equally likely to occur, as was assumed in Eq. (1), maximizing $P(\pm|\mathcal{O})$ leads to Eq. (2).

It is often the case that the readout outcome \mathcal{O} is a real number, such as a voltage or a current. Moreover, it is common that the distributions $P(\mathcal{O}|\pm)$ are well separated and intersect at a single value $\mathcal{O} = \mathcal{O}_{\text{th}}$, given by solving $P(\mathcal{O}_{\text{th}}|+) = P(\mathcal{O}_{\text{th}}|-)$. In that case, the decision rule of Eq. (2) reduces to

$$\mathcal{O} > \mathcal{O}_{\text{th}} \rightarrow |+\rangle, \quad \mathcal{O} < \mathcal{O}_{\text{th}} \rightarrow |-\rangle. \quad (3)$$

Equation (3) is the decision rule that will be used in the remainder of this article.

III. MINIMAL READOUT MODEL

In Sec. II, we gave a general description of a single-shot qubit readout. We now review a minimal model of qubit readout [38] that incorporates the main limiting sources of error in a wide variety of implementations, namely, readout noise and qubit relaxation. Despite its simplicity, this model and its extensions [12,21,39–41] have been used as the starting point for the analysis of qubit readout in various experiments [10,12,21,23,27]. The formalism of the present section will be used as a benchmark to quantify the improvements discussed in Sec. IV.

The readout dynamics of the minimal model are illustrated in Fig. 1(a). If the qubit is in the ground state $|-\rangle$ at time $t = 0$, it remains in the ground state for all times $t > 0$. However, if the qubit is in the excited state $|+\rangle$, it relaxes to $|-\rangle$ at some random time $\tau > 0$. We assume that the relaxation is an incoherent Markov process. The time τ is then distributed according to the exponential distribution $P_0(\tau) = \Gamma e^{-\Gamma\tau}$, where Γ is the inverse relaxation time. The distribution $P_0(\tau)$ is illustrated in Fig. 1(b).

The qubit is coupled to a detector (e.g., a proximal charge sensor for readout of semiconductor spin qubits) that can distinguish between the two qubit states at any given time. More precisely, the detector registers a signal $I(t) = I_{\pm} + \delta I(t)$ if the qubit state is $|\pm\rangle$ at time t . Here I_{\pm} are state-dependent average detector signals and $\delta I(t)$ is detector noise. Without loss of generality, we assume $I_+ > I_-$ and we define the readout contrast $\Delta I = I_+ - I_-$. When the qubit transitions between different states, the transition appears as a sharp “jump” in the detector signal at time τ [42]. Particular instances of the readout signals $I(t)$ for both qubit states are illustrated in Fig. 1(b). After a transition occurs, the two qubit states can no longer be distinguished by the detector. Hence

the lifetime of the readout signal is given by the relaxation time $\langle\tau\rangle = \Gamma^{-1}$.

If the detector signal were not noisy, the qubit state could be identified with perfect accuracy by integrating the signal for short times $t \ll \Gamma^{-1}$, thus avoiding relaxation events that would cause readout errors. In practice, however, noise in the detector sets a minimum readout time that limits the fidelity. For simplicity, we assume Gaussian white noise $\delta I(t)$ with autocorrelation function $\langle\delta I(t)\delta I(t')\rangle = R^{-1}\delta(t-t')$. In this case, the two average signals I_{\pm} can be resolved in a readout time t of order $1/r$, where $r = R(\Delta I/2)^2$ is the rate of change of the (power) SNR. When $r \gg \Gamma$, the probability of a relaxation event occurring during that time is thus approximately Γ/r . Therefore, we expect the average error rate ϵ in that limit to be of order $1/S$, where

$$S = \frac{r}{\Gamma} = \frac{R\Delta I^2}{4\Gamma} \quad (4)$$

is the SNR obtained upon averaging the noise over a time Γ^{-1} . In the following, “the SNR” always refers to the quantity S .

To make the above intuition more formal, we need to specify the measurement outcomes \mathcal{O} introduced in Sec. II. A natural (although suboptimal [38]) choice that captures the tradeoff between noise averaging and relaxation is the time-averaged signal \bar{I} :

$$\mathcal{O} = \bar{I} = \frac{1}{t} \int_0^t dt' I(t'). \quad (5)$$

Here the readout time t must be chosen to optimize the aforementioned tradeoff. Analytical expressions for the probability distributions $P(\bar{I}|\pm)$ as well as the corresponding single-shot readout error rate ϵ have been obtained [38]. In particular, it can be shown that

$$\epsilon \sim \frac{1}{2S} \ln S, \quad (6)$$

as $S \rightarrow \infty$. Here “ \sim ” denotes the asymptotic equality. This result is in qualitative agreement with the heuristic argument given above, up to a numerical prefactor and logarithmic corrections.

IV. READOUT ENHANCEMENT VIA SUB-POISSONIAN DYNAMICS

In the previous section, the single-shot readout error rate was limited by random relaxation events from $|+\rangle$ to $|-\rangle$ at short readout times t . It follows that a natural strategy to improve readout is to suppress the frequency of these relaxation events. One way to achieve this would be to increase the relaxation time $\langle\tau\rangle = \Gamma^{-1}$. In this section, however, we show that these errors can instead be suppressed by using intermediate states to engineer sub-Poissonian dynamics that reduce the variance of τ , thus making the relaxation process more deterministic.

We introduce N intermediate states between the excited state $|+\rangle$ and the ground state $|-\rangle$. The dynamics are such that the system must sequentially transition through all intermediate states to go from state $|+\rangle$ to state $|-\rangle$. This process is depicted in Fig. 2(a). Such cascaded dynamics can be approximately realized in existing qubit implementations

[27,28,31,32,34,43] and were investigated in the context of transport through a chain of quantum dots [35].

We assume that all intermediate states couple to the detector as strongly as the excited state $|+\rangle$, i.e., the detector registers an average signal I_+ for all states except $|-\rangle$. Thus a jump in the readout signal occurs at the time τ when the system transitions from the last intermediate state to $|-\rangle$ [see Fig. 2(b)]. Note also that the readout contrast remains unchanged from Sec. III, $\Delta I = I_+ - I_-$. In addition, we assume that every transition along the sequence occurs with the same probability per unit time $\Gamma_N = (N+1)\Gamma$. As a result of the factor $N+1$, the lifetime of the readout signal is still $\langle \tau \rangle = \Gamma^{-1}$.

Because both the readout contrast and the readout signal lifetime are the same as in Sec. III, the SNR is again given by Eq. (4). It might therefore be naively expected that the readout fidelity is not significantly affected. However, the introduction of intermediate states qualitatively modifies the statistics of the jump time τ , which now follows the Gamma distribution

$$P_N(\tau) = \frac{\Gamma_N^{N+1} \tau^N}{N!} e^{-\Gamma_N \tau}. \quad (7)$$

The distribution $P_N(\tau)$ is illustrated in Fig. 2(b) for $N=3$. The distribution becomes peaked around $\tau = \Gamma^{-1}$, which strongly suppresses the probability that a jump event occurs at times smaller than Γ^{-1} . Indeed, according to Eq. (7), the probability of a jump occurring within a readout time t of order r^{-1} is now roughly $(\Gamma_N/r)^{N+1}/(N+1)!$ provided that $r/\Gamma_N = S/(N+1) \gg 1$. In this regime, we therefore expect the single-shot readout error rate to be proportional to

$$\epsilon \propto \frac{1}{(N+1)!} \left(\frac{N+1}{S} \right)^{N+1}. \quad (8)$$

For large enough S , increasing the number of intermediate states N should therefore lead to an exponential suppression of the error rate compared to the result of Sec. III for $N=0$. A direct asymptotic expansion gives the correct leading-order asymptotic contribution to the error rate

$$\epsilon \sim \frac{1}{2(N+1)!} \left\{ \frac{N+1}{S} \ln \left[2^N N! \left(\frac{S}{N+1} \right)^{N+1} \right] \right\}^{N+1} \quad (9)$$

as $S \rightarrow \infty$. Equation (9) is in qualitative agreement with the heuristic order-of-magnitude estimate, Eq. (8), again up to a numerical prefactor and logarithmic corrections.

To verify that an advantage exists in a nonasymptotic parameter regime, we perform a full analysis of the statistics of the time-averaged signal \bar{I} , Eq. (5), in Appendix A. Analytical expressions for the single-shot readout error rate ϵ are first derived in the presence of intermediate states. The error rate ϵ is then minimized numerically with respect to the readout time t and decision threshold \bar{I}_{th} . We plot ϵ on a logarithmic scale as a function of SNR for N ranging from 0 to 4 in Fig. 3. For moderate values of the SNR ($S \lesssim 100$), the addition of a single intermediate state leads to a near order-of-magnitude reduction in ϵ .

Our analysis of the time-averaged signal \bar{I} , Eq. (5), has enabled us to analytically and efficiently calculate the single-shot readout error rate ϵ . However, as mentioned in Sec. III, the time-averaged signal is a suboptimal statistic for qubit readout. This is because time-averaging the signal discards

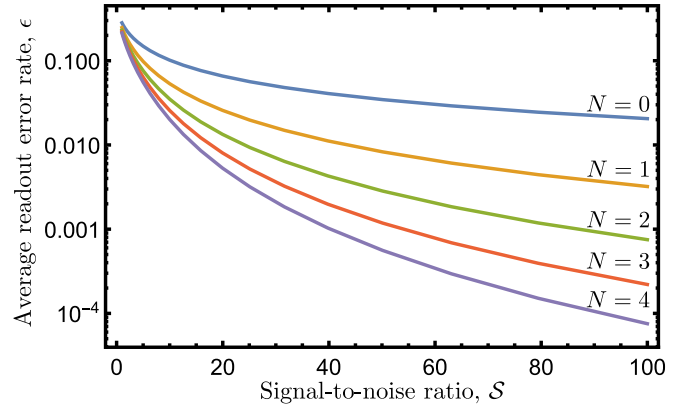


FIG. 3. Average single-shot readout error rate Eq. (1) on a logarithmic scale as a function of the SNR for different numbers N of intermediate states. The reported values of ϵ are calculated as described in the main text and as detailed in Appendix A.

single-shot information about the statistics of the jump time [38]. In contrast, optimal readout is achieved by choosing the readout outcome to be the full state-dependent signal $\mathcal{O} = I(t)$. Because time averaging is suboptimal, it is prudent to verify that the advantage discussed above is intrinsic to the system dynamics and not merely an artifact of the processing method. We therefore implemented the optimal processing method [40,41] for the readout dynamics of Fig. 2(a). The relevant formalism is outlined in Appendix B. We use this method to obtain a Monte Carlo estimate of the minimum theoretically achievable error rate. The results are illustrated in Fig. 5 of Appendix B for $S=20$ and several values of N . While the optimal method reduces the error rate compared to the time-averaging method, the readout enhancement due to sub-Poissonian dynamics remains the same.

Our results were obtained under the assumption that the detector registers the same contrast ΔI for all states except $|-\rangle$. Moreover, it was assumed that all transitions occur at the same rate $\Gamma_N = (N+1)\Gamma$. In practice, however, there will necessarily be variations in the contrast and in the transition rate at each step of the cascade. Therefore, it is important to verify that the readout enhancement persists in spite of small asymmetries in these parameters. Using the optimal method described in Appendix B, we show numerically that the enhanced error rate is weakly dependent on either an asymmetry in contrast or an asymmetry in transition rates. The results are shown in Fig. 6 of Appendix C for $N=1$ and $S=20$. Thus, no fine-tuning is required for the readout to benefit from the sub-Poissonian dynamics.

V. DISCUSSION

The results presented here show that intermediate states can be useful even in situations where the SNR is unchanged. This means that the SNR is, in general, not sufficient to fully characterize the readout. We thus identify sub-Poissonian qubit dynamics as an *additional resource* for qubit readout engineering. Depending on the experimental constraints, either or both of these resources can be used to optimize readout fidelity.

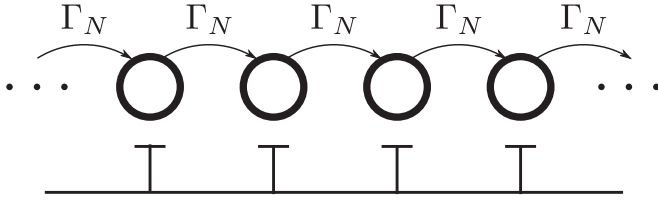


FIG. 4. Sub-Poissonian charge dynamics in a chain of quantum dots [35]. Metallic gates are used to maintain strong capacitive coupling between distant dots while the charge hops from one quantum dot to the next (circles) with probability per unit time Γ_N . The resulting statistics for the total transition time τ are given by Eq. (7).

In Sec. IV, for instance, we fix the readout signal lifetime $\langle \tau \rangle = \Gamma^{-1}$ by increasing the transition rates proportionally to N , $\Gamma_N = (N + 1)\Gamma$. This is done to demonstrate that sub-Poissonian statistics of the jump time τ may reduce the error rate without increasing SNR. In general, however, the addition of intermediate states may increase $\langle \tau \rangle$ (and thus the SNR) as well as reduce the randomness in τ . As can be seen in Fig. 3, such a combined increase in SNR and N would lead to an even greater suppression of errors. In particular, the enhancement of the readout of the NV -center spin in diamond reported in Ref. [34] should already benefit from this combination of effects even though high-fidelity single-shot readout has not been achieved.

Readout of semiconductor spin qubits [22–28] and recent proposals for readout of topologically protected qubits [29,30] frequently rely on the detection of a transient charge state (or a transient flow of charge) in quantum dots. In these schemes, charge relaxation often limits the readout signal lifetime. Stabilizing the charge state is therefore crucial to improving the readout fidelity. Engineering sub-Poissonian charge dynamics could be a promising avenue for achieving this goal. For instance, the relaxation process could be made generically sub-Poissonian by forcing the charge to tunnel through a chain of quantum dots in the Coulomb blockade regime [35]. In principle, such chains can be made long enough to include several dots by using gates to enforce long-range capacitive coupling between the dots [44–47]. Such a long-range coupling prevents double occupancy of the chain, leading to highly sub-Poissonian statistics of the total transition time (see Fig. 4).

Although we restricted our attention to a cascade of incoherent transitions, the engineering of sub-Poissonian relaxation statistics could, in principle, improve readout for more general dynamics including, e.g., coherent dynamics described by a Lindblad quantum master equation. The only requirement is that the distribution $P(\tau)$ becomes narrow enough to suppress the probability of a jump occurring during the readout. More generally, it should only be necessary to make τ more deterministic by reducing the (differential) entropy of $P(\tau)$.

VI. CONCLUSION

In summary, we show that sub-Poissonian dynamics is a useful resource for qubit readout engineering. In particular, we

demonstrate that forcing a qubit to relax through a sequence of intermediate states can reduce the single-shot readout error rate by orders of magnitude. This enhancement is entirely due to the induced sub-Poissonian statistics of the relaxation process and thus holds even when the SNR is unchanged. We also verify that the enhancement is significant for moderate values of the SNR ($S \lesssim 100$) and small numbers of intermediate states ($N \lesssim 10$). The example given here makes it clear that the single-shot readout fidelity should be optimized accounting for the full readout statistics, rather than simply quoting the SNR.

A particularly promising application of our results is the enhanced detection of short-lived charge states using charge sensors. As discussed in Sec. V, charge relaxation dynamics can be made generically sub-Poissonian with the help of quantum dot chains. This could have important implications for recent proposals to read out topologically protected Majorana qubits via parity-to-charge conversion.

ACKNOWLEDGMENTS

We acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Institute for Advanced Research (CIFAR), the Institut Transdisciplinaire d'Information Quantique (INTRIQ), and Nordea Fonden.

APPENDIX A: ANALYTICAL EXPRESSIONS FOR THE SINGLE-SHOT READOUT ERROR RATE

In this Appendix, we briefly outline the derivation of the analytical expressions used to plot Fig. 3. The formalism presented here may also be used as the starting point for the derivation of asymptotic results such as Eqs. (6) and (9).

The probability distribution of the time-averaged signal \bar{I} , Eq. (5), depends on the qubit state. When the qubit state is $|-\rangle$, the average detector signal is I_- for all readout times $t > 0$. Assuming Gaussian white noise of the form $\langle \delta I(t) \delta I(t') \rangle = R^{-1} \delta(t - t')$, \bar{I} is a Gaussian random variable with mean I_- and variance $1/Rt$:

$$P(\bar{I}|-) = \mathcal{N}(\bar{I}; I_-, 1/\sqrt{Rt}), \quad (\text{A1})$$

where

$$\mathcal{N}(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (\text{A2})$$

is the normal distribution of mean μ and variance σ^2 . When the qubit state is $|+\rangle$, the presence of relaxation leads to a non-Gaussian distribution $P(\bar{I}|+)$. However, the distribution $P(\bar{I} +, \tau)$ conditioned on a particular value of the jump time τ is Gaussian. Using Bayes' rule, we have

$$P(\bar{I}|+) = \int_0^\infty d\tau P(\bar{I} +, \tau) P_N(\tau), \quad (\text{A3})$$

where

$$P(\bar{I} +, \tau) = \begin{cases} \mathcal{N}(\bar{I}; I_- + (\tau/t)\Delta I, 1/\sqrt{Rt}) & \text{if } \tau < t, \\ \mathcal{N}(\bar{I}; I_+, 1/\sqrt{Rt}) & \text{if } \tau \geq t. \end{cases} \quad (\text{A4})$$

According to the decision rule outlined in Sec. II, the conditional error rates $\epsilon_\pm^{(N)}$ for a given value of N are given by

the probabilities that \bar{I} is above (below) the decision threshold \bar{I}_{th} when the qubit state is $|-\rangle$ ($|+\rangle$):

$$\begin{aligned}\epsilon_+^{(N)} &= \int_{-\infty}^{\bar{I}_{\text{th}}} d\bar{I} P(\bar{I}|+), \\ \epsilon_-^{(N)} &= \int_{\bar{I}_{\text{th}}}^{\infty} d\bar{I} P(\bar{I}|-). \end{aligned} \quad (\text{A5})$$

The above integrals were performed in Ref. [38] for the case $N = 0$. The resulting analytical expressions for the conditional error rates $\epsilon_{\pm}^{(0)}$ are

$$\begin{aligned}\epsilon_+^{(0)} &= \frac{1}{2} e^{\rho \frac{\gamma}{r} (\frac{\gamma}{4r} - \nu)} \\ &\times \left\{ \text{erf} \left[\sqrt{2\rho} \left(\frac{\gamma}{4r} - \nu \right) \right] - \text{erf} \left[\sqrt{2\rho} \left(\frac{\gamma}{4r} + 1 - \nu \right) \right] \right\} \\ &+ \frac{1}{2} \text{erfc}(-\sqrt{2\rho\nu^2}), \\ \epsilon_-^{(0)} &= \frac{1}{2} \text{erfc}(\sqrt{2\rho\nu^2}). \end{aligned} \quad (\text{A6})$$

Here $r = R(\Delta I/2)^2$, $\rho = rt$, $\nu = (\bar{I}_{\text{th}} - I_-)/\Delta I$, and γ is the probability per unit time to transition from one state to the next [set to $\gamma = \Gamma_N = (N+1)\Gamma$ in the main text].

To obtain the error rates for $N > 0$, we notice that

$$P_N(\tau) = \frac{(-1)^N \gamma^{N+1}}{N!} \frac{d^N}{d\gamma^N} \left(\frac{P_0(\tau)}{\gamma} \right), \quad (\text{A7})$$

where $P_0(\tau) = \gamma e^{-\gamma\tau}$ is the exponential distribution. It then follows from Eqs. (A3) and (A5) that analytical expressions for $\epsilon_{\pm}^{(N)}$ when $N > 0$ can be obtained by taking derivatives of Eq. (A6) with respect to the transition rate γ :

$$\begin{aligned}\epsilon_+^{(N)} &= \frac{(-1)^N \gamma^{N+1}}{N!} \frac{d^N}{d\gamma^N} \left(\frac{\epsilon_+^{(0)}}{\gamma} \right), \\ \epsilon_-^{(N)} &= \epsilon_-^{(0)}. \end{aligned} \quad (\text{A8})$$

These expressions are used to minimize the average single-shot readout error rate $\epsilon^{(N)} = [\epsilon_+^{(N)} + \epsilon_-^{(N)}]/2$ with respect to both ρ and ν . Setting $\gamma = (N+1)\Gamma$ and $r/\Gamma = S$, this yields the results presented in Fig. 3.

APPENDIX B: COMPARISON TO OPTIMAL PROCESSING

The optimal processing method calculates the likelihood ratio Λ from the full signal $\mathcal{O} = I(t)$. This is achieved by solving a set of linear classical Itô stochastic differential equations for the unconditioned system state ℓ [40,41]. These equations, also known as filtering equations, have the form

$$d\ell = [\mathcal{L} + I(t)\mathcal{I}R]\ell dt \quad (\text{It}\bar{\text{o}}). \quad (\text{B1})$$

Here ℓ is a vector with components ℓ_i such that the probability of finding the system in state $|i\rangle$ conditioned on past observations is $\ell_i / \sum_j \ell_j$. The matrix \mathcal{L} is the rate matrix that describes the transitions illustrated in Fig. 2, while the matrix $\mathcal{I} = \text{diag}(I_0, I_1, \dots, I_{N+1})$ encodes the average signals I_i registered by the detector for each state $|i\rangle$ along the cascade.

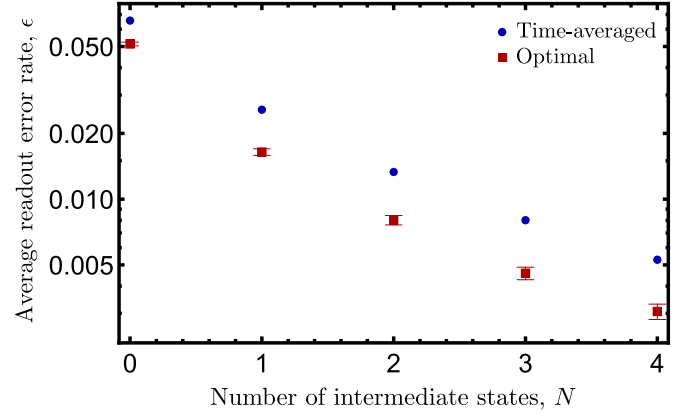


FIG. 5. Average error rate ϵ as a function of the number of intermediate states N using the time-averaged method Eq. (5) (blue dots) and using the optimal method Eq. (B3) (red squares). The SNR is $S = 20$. In the optimal case, each error rate is obtained by simulating 5×10^4 signals $I(t)$ for each state $|\pm\rangle$. The error bars for the optimal method are given by the standard deviation of a binomial process with frequency ϵ .

The likelihood ratio is given by

$$\Lambda = \frac{\sum_i \ell_i^{(+)}}{\sum_i \ell_i^{(-)}}, \quad (\text{B2})$$

where $\ell^{(\pm)}$ is the solution of Eq. (B1) assuming an initial state $|\pm\rangle$. In practice, Eq. (B1) requires a very small time step to solve accurately. We therefore use an alternative set of ordinary differential equations [38,39,48]

$$\frac{d\ell}{dt} = \left[\mathcal{L} + \left(I(t) - \frac{1}{2}\mathcal{I} \right) \mathcal{I}R \right] \ell. \quad (\text{B3})$$

It can be shown that Eq. (B3) is equivalent to Eq. (B1) provided that it is solved with a higher-order numerical method such as the fourth-order Runge-Kutta method [49]. The use of a higher-order method enables an accurate computation of Λ with a reasonable number of time steps. Note that Eq. (B3) may not be solved with the Euler method since its first-order increment differs from that of Eq. (B1).

For both states $|\pm\rangle$, a large number of signals $I(t)$ are simulated. For a given set of parameters, the measurement time is chosen to be long enough to achieve the minimum error rate. For each signal, Eq. (B3) is solved using a fourth-order Runge-Kutta method. The likelihood ratio Λ is then calculated using Eq. (B2). If $\Lambda < 1$ ($\Lambda > 1$) when the state is $|+\rangle$ ($|-\rangle$), an error is recorded. This procedure gives a Monte Carlo estimate of the conditional single-shot readout error rates ϵ_{\pm} . These values are then used to calculate $\epsilon = (\epsilon_+ + \epsilon_-)/2$. The results are shown in Fig. 5 for $S = 20$ and several values of N .

APPENDIX C: ASYMMETRICAL PARAMETERS

In this Appendix, we discuss the effect of an asymmetry in contrast or in transition rates. A convenient way to parametrize

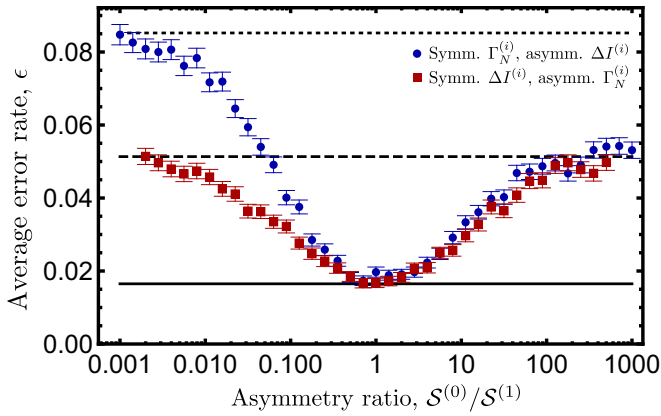


FIG. 6. Average error rate ϵ as a function of the asymmetry ratio $S^{(0)}/S^{(1)}$ in the case of $N = 1$ intermediate state. The asymmetry in SNR arises from an asymmetry in contrast $\Delta I^{(i)}$ (blue dots) or from an asymmetry in transition rates $\Gamma_N^{(i)}$ (red squares). The total SNR is $S = S^{(0)} + S^{(1)} = 20$. Each error rate is obtained by simulating 10^4 signals $I(t)$ for each state $|\pm\rangle$ and then applying the optimal processing method, Eqs. (B2) and (B3). The error bars are given by the standard deviation of a binomial process with frequency ϵ . The horizontal lines are the error rates obtained from simulations of the symmetric case (solid) and completely asymmetric cases (dashed and dotted) with 10^5 samples for each state. The completely asymmetric cases correspond to the readout dynamics discussed in Refs. [38,39].

the asymmetry is to define partial SNRs, $S^{(i)} = R\Delta I^{(i)2}/4\Gamma_N^{(i)}$, for $i = 0, 1, \dots, N$. Here, $\Delta I^{(i)}$ and $\Gamma_N^{(i)}$ are the i^{th} contrast and transition rate in the cascade, respectively. The total SNR is then defined as $S = \sum_i S^{(i)}$. For a fair comparison of the asymmetric case with the symmetric case, the total SNR must remain fixed as the $S^{(i)}$ are varied. When the $S^{(i)}$ differ significantly from each other, the advantage due to sub-Poissonian dynamics is expected to disappear since only a minority of intermediate states are responsible for most of the SNR. To demonstrate the resilience of the sub-Poissonian enhancement to an asymmetry in parameters, we use the optimal method of Appendix B to calculate the error rate as a function of the asymmetry ratio $S^{(0)}/S^{(1)}$ for the case $N = 1$ and $S = 20$. We introduce an asymmetry

- (1) by choosing symmetric $\Gamma_N^{(i)}$ and varying the $\Delta I^{(i)}$, and
- (2) by choosing symmetric $\Delta I^{(i)}$ and varying the $\Gamma_N^{(i)}$.

The results are shown in Fig. 6. In either case, the symmetry point is a minimum. Moreover, the error rate does not change significantly provided that the asymmetry ratio $S^{(0)}/S^{(1)}$ differs from unity by less than an order of magnitude. For large asymmetries, we recover the readout dynamics discussed in Refs. [38,39], both of which lead to higher error rates than in the symmetric case.

- [1] D. P. DiVincenzo, in *Scalable Quantum Computers: Paving the Way to Realization*, edited by S. L. Braunstein, H.-K. Lo, and P. Kok, (Wiley-VCH, Berlin, 2001), Chap. 1, pp. 1–13.
- [2] R. Raussendorf and H. J. Briegel, *Phys. Rev. Lett.* **86**, 5188 (2001).
- [3] P. Bonderson, M. Freedman, and C. Nayak, *Phys. Rev. Lett.* **101**, 010501 (2008).
- [4] T. Karzig, C. Knapp, R. M. Lutchyn, P. Bonderson, M. B. Hastings, C. Nayak, J. Alicea, K. Flensberg, S. Plugge, Y. Oreg, C. M. Marcus, and M. H. Freedman, *Phys. Rev. B* **95**, 235305 (2017).
- [5] A. Lupașcu, S. Saito, T. Picot, P. C. de Groot, C. J. P. M. Harmans, and J. E. Mooij, *AIP Conf. Proc.* **1074**, 26 (2008).
- [6] F. Mallet, F. R. Ong, A. Palacios-Laloy, F. Nguyen, P. Bertet, D. Vion, and D. Esteve, *Nat. Phys.* **5**, 791 (2009).
- [7] Y. Liu, S. J. Srinivasan, D. Hover, S. Zhu, R. McDermott, and A. A. Houck, *New J. Phys.* **16**, 113008 (2014).
- [8] D. Hover, S. Zhu, T. Thorbeck, G. Ribeill, D. Sank, J. Kelly, R. Barends, J. M. Martinis, and R. McDermott, *App. Phys. Lett.* **104**, 152601 (2014).
- [9] E. Jeffrey, D. Sank, J. Y. Mutus, T. C. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Megrant, P. J. J. O'Malley, C. Neill, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and J. M. Martinis, *Phys. Rev. Lett.* **112**, 190504 (2014).
- [10] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potočnik, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, and A. Wallraff, *Phys. Rev. Appl.* **7**, 054020 (2017).
- [11] D. B. Hume, T. Rosenband, and D. J. Wineland, *Phys. Rev. Lett.* **99**, 120502 (2007).
- [12] A. H. Myerson, D. J. Szwer, S. C. Webster, D. T. C. Allcock, M. J. Curtis, G. Imreh, J. A. Sherman, D. N. Stacey, A. M. Steane, and D. M. Lucas, *Phys. Rev. Lett.* **100**, 200502 (2008).
- [13] R. Gehr, J. Volz, G. Dubois, T. Steinmetz, Y. Colombe, B. L. Lev, R. Long, J. Estève, and J. Reichel, *Phys. Rev. Lett.* **104**, 203602 (2010).
- [14] R. Noek, G. Vrijsen, D. Gaultney, E. Mount, T. Kim, P. Maunz, and J. Kim, *Opt. Lett.* **38**, 4735 (2013).
- [15] T. P. Harty, D. T. C. Allcock, C. J. Ballance, L. Guidoni, H. A. Janacek, N. M. Linke, D. N. Stacey, and D. M. Lucas, *Phys. Rev. Lett.* **113**, 220501 (2014).
- [16] L. Jiang, J. S. Hodges, J. R. Maze, P. Maurer, J. M. Taylor, D. G. Cory, P. R. Hemmer, R. L. Walsworth, A. Yacoby, A. S. Zibrov, and M. D. Lukin, *Science* **326**, 267 (2009).
- [17] P. Neumann, J. Beck, M. Steiner, F. Rempp, H. Fedder, P. R. Hemmer, J. Wrachtrup, and F. Jelezko, *Science* **329**, 542 (2010).
- [18] L. Robledo, L. Childress, H. Bernien, B. Hensen, P. F. Alkemade, and R. Hanson, *Nature* **477**, 574 (2011).
- [19] A. Dréau, P. Spinicelli, J. R. Maze, J.-F. Roch, and V. Jacques, *Phys. Rev. Lett.* **110**, 060502 (2013).
- [20] B. J. Shields, Q. P. Unterreithmeier, N. P. de Leon, H. Park, and M. D. Lukin, *Phys. Rev. Lett.* **114**, 136402 (2015).
- [21] B. D'Anjou, L. Kurek, L. Childress, and W. A. Coish, *Phys. Rev. X* **6**, 011017 (2016).
- [22] J. M. Elzerman, R. Hanson, L. H. W. Van Beveren, B. Witkamp, L. M. K. Vandersypen, and L. P. Kouwenhoven, *Nature* **430**, 431 (2004).
- [23] C. Barthel, D. J. Reilly, C. M. Marcus, M. P. Hanson, and A. C. Gossard, *Phys. Rev. Lett.* **103**, 160503 (2009).

- [24] A. Morello, J. J. Pla, F. A. Zwanenburg, K. W. Chan, K. Y. Tan, H. Huebl, M. Möttönen, C. D. Nugroho, C. Yang, J. A. van Donkelaar, A. D. C. Alves, D. N. Jamieson, C. C. Escott, L. C. L. Hollenberg, R. G. Clark, and A. S. Dzurak, *Nature* **467**, 687 (2010).
- [25] J. J. Pla, K. Y. Tan, J. P. Dehollain, W. H. Lim, J. J. Morton, F. A. Zwanenburg, D. N. Jamieson, A. S. Dzurak, and A. Morello, *Nature* **496**, 334 (2013).
- [26] K. Eng, T. D. Ladd, A. Smith, M. G. Borselli, A. A. Kiselev, B. H. Fong, K. S. Holabird, T. M. Hazard, B. Huang, P. W. Deelman, I. Milosavljevic, A. E. Schmitz, R. S. Ross, M. F. Gyure, and A. T. Hunter, *Science Advances* **1**, e150021 (2015).
- [27] P. Harvey-Collard, B. D'Anjou, M. Rudolph, N. T. Jacobson, J. Dominguez, G. A. T. Eyck, J. R. Wendt, T. Pluym, M. P. Lilly, W. A. Coish, M. Pioro-Ladrière, and M. S. Carroll, [arXiv:1703.02651](https://arxiv.org/abs/1703.02651).
- [28] T. Nakajima, M. R. Delbecq, T. Otsuka, P. Stano, S. Amaha, J. Yoneda, A. Noiri, K. Kawasaki, K. Takeda, G. Allison, A. Ludwig, A. D. Wieck, D. Loss, and S. Tarucha, *Phys. Rev. Lett.* **119**, 017701 (2017).
- [29] D. Aasen, M. Hell, R. V. Mishmash, A. Higginbotham, J. Danon, M. Leijnse, T. S. Jespersen, J. A. Folk, C. M. Marcus, K. Flensberg, and J. Alicea, *Phys. Rev. X* **6**, 031016 (2016).
- [30] S. Plugge, A. Rasmussen, R. Egger, and K. Flensberg, *New J. Phys.* **19**, 012001 (2017).
- [31] S. A. Studenikin, J. Thorgrimson, G. C. Aers, A. Kam, P. Zawadzki, Z. R. Wasilewski, A. Bogan, and A. S. Sachrajda, *App. Phys. Lett.* **101**, 233101 (2012).
- [32] J. D. Mason, S. A. Studenikin, A. Kam, Z. R. Wasilewski, A. S. Sachrajda, and J. B. Kycia, *Phys. Rev. B* **92**, 125434 (2015).
- [33] N. Earnest, S. Chakram, Y. Lu, N. Irons, R. K. Naik, N. Leung, J. Lawrence, J. Koch, and D. I. Schuster, [arXiv:1707.00656](https://arxiv.org/abs/1707.00656).
- [34] M. Steiner, P. Neumann, J. Beck, F. Jelezko, and J. Wrachtrup, *Phys. Rev. B* **81**, 035205 (2010).
- [35] A. N. Korotkov and K. K. Likharev, *Phys. Rev. B* **61**, 15975 (2000).
- [36] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, England, 2010).
- [37] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. II: Detection Theory*, Vol. II, (Prentice Hall, Princeton, NJ, 1998).
- [38] J. Gambetta, W. A. Braff, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, *Phys. Rev. A* **76**, 012325 (2007).
- [39] B. D'Anjou and W. A. Coish, *Phys. Rev. A* **89**, 012313 (2014).
- [40] S. Ng and M. Tsang, *Phys. Rev. A* **90**, 022325 (2014).
- [41] B. D'Anjou, Optimization of Real-World Qubit Measurements, Ph.D. thesis, McGill University, 2017, available at <http://digitool.library.mcgill.ca/R/>. See Chap. 1.3. for an introduction to the optimal filtering equations.
- [42] For simplicity, we assume that the detector response time is much shorter than the relaxation time Γ^{-1} .
- [43] S. Gasparinetti, M. Pechal, J.-C. Besse, M. Mondal, C. Eichler, and A. Wallraff, *Phys. Rev. Lett.* **119**, 140504 (2017).
- [44] I. H. Chan, R. M. Westervelt, K. D. Maranowski, and A. C. Gossard, *App. Phys. Lett.* **80**, 1818 (2002).
- [45] Y. Hu, H. O. H. Churchill, D. J. Reilly, J. Xiang, M. Lieber, Charles, and C. M. Marcus, *Nat. Nanotech.* **2**, 622 (2007).
- [46] K. Flensberg and C. M. Marcus, *Phys. Rev. B* **81**, 195418 (2010).
- [47] L. Trifunovic, O. Dial, M. Trif, J. R. Wootton, R. Abebe, A. Yacoby, and D. Loss, *Phys. Rev. X* **2**, 011006 (2012).
- [48] J. Wilkie, *Phys. Rev. E* **70**, 017701 (2004).
- [49] H. Press, A. Teukolsky, T. Vetterling, and P. Flannery, *Numerical Recipes in C++: The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 2002).