Quantum Fokker-Planck Master Equation for Continuous Feedback Control

Björn Annby-Andersson,[1,*] Faraj Bakhshinezhad,[1] Debankur Bhattacharyya,[2] Guilherme De Sousa,[3] Christopher Jarzynski,[2] Peter Samuelsson,[1] and Patrick P. Potts[1,4]

[1]Physics Department and NanoLund, Lund University, Box 118, 22100 Lund, Sweden
[2]Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA
[3]Department of Physics, University of Maryland, College Park, Maryland 20742, USA
[4]Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland

(Received 8 November 2021; accepted 24 June 2022; published 25 July 2022)

Measurement and feedback control are essential features of quantum science, with applications ranging from quantum technology protocols to information-to-work conversion in quantum thermodynamics. Theoretical descriptions of feedback control are typically given in terms of stochastic equations requiring numerical solutions, or are limited to linear feedback protocols. Here we present a formalism for continuous quantum measurement and feedback, both linear and nonlinear. Our main result is a quantum Fokker-Planck master equation describing the joint dynamics of a quantum system and a detector with finite bandwidth. For fast measurements, we derive a Markovian master equation for the system alone, amenable to analytical treatment. We illustrate our formalism by investigating two basic information engines, one quantum and one classical.

DOI: 10.1103/PhysRevLett.129.050401

Introduction.—Quantum measurement and feedback control are key elements for emerging quantum technologies, enabling a wide range of applications, including quantum error correction [1], deterministic entanglement generation [2], atomic clocks [3], and quantum state stabilization [4–6]. The past two decades have also witnessed a large number of fundamental experiments on feedback control of quantum systems [7–18]. Of special interest are experiments in quantum thermodynamics [19]—by using measurement and feedback, processes that are otherwise forbidden by the second law of thermodynamics may be realized, compellingly illustrated by Maxwell's demon [20–22]. Over the past ten years, the demon has been realized in a wide range of experimental settings, both in classical [23–29] and, recently, quantum systems [30–34]. This activity has inspired further work investigating the connection between thermodynamics and information theory [35–37], and has resulted in generalizations of the second law for feedback controlled systems [38–48]. A promising platform for exploring feedback control within quantum thermodynamics is solid-state electronic systems [49], ranging from semiconductor quantum dots [50] to superconducting qubits [51]. Key features in these systems are large and fast tunability of system properties [52–54] and time resolved measurements [55,56]. Moreover, both discrete [29,57,58] and continuous [6,27] feedback protocols have been demonstrated experimentally.

The theoretical description of feedback control in quantum systems is typically based on stochastic differential equations [59–70]—powerful tools that can describe discrete as well as continuous feedback protocols. In general, these equations must be solved numerically, providing limited qualitative insight. An important exception, amenable to analytical treatment, is the Wiseman-Milburn equation [63], a Markovian master equation for continuous feedback protocols that depend linearly on the measured signal. However, optimal control often requires nonlinear protocols, for instance, bang-bang control [71,72] which has promising thermodynamic applications in solid-state architectures [27,73–75]. For such continuous, nonlinear feedback protocols, no master equation description exists, emphasizing a need for further analytical tools. We stress that the word "nonlinear" here refers to the protocol's dependence on the measured signal, not to the system's dynamics.

In this Letter, we satisfy this need by developing a general framework for continuous measurement and feedback control in quantum systems, able to provide analytical insight into nonlinear feedback protocols. Our main result, Eq. (1) below, is a quantum Fokker-Planck master equation describing the joint dynamics of a quantum system and a detector with finite bandwidth (see Fig. 1). This equation is applicable to any quantum or classical system undergoing continuous feedback control. For fast measurements, Eq. (1) reduces to a Markovian master equation for the

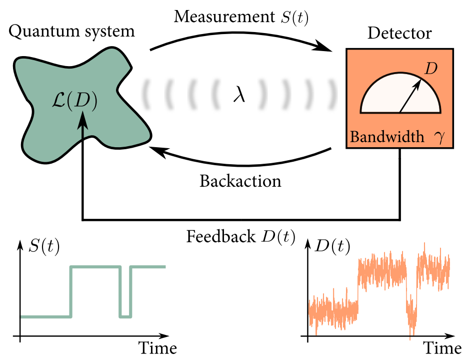

FIG. 1. Illustration of a generic measurement and feedback setup, consisting of an open quantum system and a detector with finite bandwidth γ. The detector continuously measures an arbitrary system observable. The measurement strength λ determines measurement backaction. Continuous feedback is applied using the measurement outcome D to control the Liouville superoperator $\mathcal{L}(D)$ of the system. The time traces visualize trajectories for the system state $S(t)$ and the measurement record $D(t)$.

system alone, generalizing the Wiseman-Milburn equation to nonlinear feedback protocols. The broad scope of Eq. (1) suggests that our results will impact a wide variety of topics where nonlinear, continuous feedback control can be applied, such as quantum error correction [1], entanglement generation [2], quantum state stabilization [6], Maxwell's demon [74,75], and machine learning [76].

To illustrate our formalism, we investigate two toy models, a classical and a quantum two-level system, operated via nonlinear feedback protocols. For the classical model, we also derive a fluctuation theorem, highlighting the role of continuous measurement and feedback in information thermodynamics.

Fokker-Planck master equation.—A general setup for continuous measurement and feedback is depicted in Fig. 1. We consider an open quantum system whose dynamics, in the absence of measurement and feedback, are described by a Liouville superoperator $\mathcal{L}$. A detector continuously measures a system observable $\hat{A}$. The measurement strength λ determines the magnitude of the measurement backaction, the limit $\lambda \to 0$ ($\lambda \to \infty$) corresponds to a weak, nonintrusive (strong, projective) measurement preserving (destroying) the quantum coherence of the system. Weak measurements thus reduce backaction, but increase measurement uncertainty. To provide a realistic detector description, we consider a finite bandwidth γ, acting as a low-pass frequency filter, eliminating high frequency measurement noise at the cost of introducing a time delay scaling as $1/\gamma$. Feedback control is incorporated by continuously feeding back the measurement outcome D into the system, controlling the system Liouville superoperator via $\mathcal{L}(D)$.

Our main result is the following deterministic Fokker-Planck master equation (derivation outlined below),

$$\partial_t \hat{\rho}_t(D) = \mathcal{L}(D)\hat{\rho}_t(D) + \lambda \mathcal{D}[\hat{A}]\hat{\rho}_t(D) - \gamma \partial_D \mathcal{A}(D)\hat{\rho}_t(D) + \frac{\gamma^2}{8\lambda}\partial_D^2 \hat{\rho}_t(D), \tag{1}$$

describing the joint system-detector dynamics under continuous measurement and feedback control. The density operator $\hat{\rho}_t(D)$ represents the joint state of system and detector, where $\hat{\rho}_t \equiv \int dD\, \hat{\rho}_t(D)$ is the system state for an unknown measurement outcome D, and $P_t(D) \equiv \mathrm{tr}\{\hat{\rho}_t(D)\}$ defines the probability distribution of the measurement outcome D. Note that $\int dD\, P_t(D) = 1$ and $\mathrm{tr}\{\hat{\rho}_t\} = 1$; see Supplemental Material (SM) [77]. The first term on the rhs of Eq. (1) describes the feedback-controlled evolution of the system. This term allows for feedback protocols that are nonlinear in D. The second term, where $\mathcal{D}[\hat{A}]\hat{\rho} = \hat{A}\hat{\rho}\hat{A} - \frac{1}{2}\{\hat{A}^2, \hat{\rho}\}$ (note $\hat{A}^\dagger = \hat{A}$), describes how the system is dephased in the eigenbasis of $\hat{A}$ at a rate proportional to λ due to measurement backaction. The last two terms constitute a Fokker-Planck equation describing the detector time evolution. These terms define an Ornstein-Uhlenbeck process [87] with a system dependent superoperator drift coefficient $\mathcal{A}(D)\hat{\rho} \equiv \frac{1}{2}\{\hat{A} - D, \hat{\rho}\}$ and diffusion constant $\gamma/8\lambda$. This describes a noisy relaxation of the measurement outcome toward a value determined by the system state. The derivation of Eq. (1) is rather involved; see details in SM [77]. The main text instead aims to highlight its implications and applications. However, we sketch the derivation at the end of the Letter.

Equation (1) is, like most formalisms for continuous measurement and feedback, typically restricted to numerical solutions. However, when there exists a wide separation between the system and detector timescales, Eq. (1) simplifies to a Markovian master equation for the system state $\hat{\rho}_t$, allowing for analytical treatment. The detector timescale $1/\gamma$ appears in the last two terms in Eq. (1), and the system timescale $1/\Gamma$ is determined by $\mathcal{L}(D) + \lambda \mathcal{D}[\hat{A}]$. The role of λ, the measurement strength, is subtle; see below. When $\gamma \gg \Gamma$, $\hat{\rho}_t$ evolves, to first order in $1/\gamma$, according to

$$\partial_t \hat{\rho}_t = [\mathcal{L}_0 + \lambda \mathcal{D}[\hat{A}] + \gamma^{-1}\mathcal{L}_{\mathrm{corr}}]\hat{\rho}_t, \tag{2}$$

with zeroth order Liouville superoperator $\mathcal{L}_0$ and first order correction $\mathcal{L}_{\mathrm{corr}}$. $\mathcal{L}_0$ is obtained by approximating the system-detector density operator as $\hat{\rho}_t(D) = [\sum_{aa'} \pi_{aa'}(D)\mathcal{V}_{aa'}]\hat{\rho}_t$, with

$$\pi_{aa'}(D) = \sqrt{4\lambda/\pi\gamma}\, e^{-(4\lambda/\gamma)[D-(\xi_a+\xi_{a'})/2]^2}, \tag{3}$$

and superoperators $\mathcal{V}_{aa'}\hat{\rho} \equiv \langle a|\hat{\rho}|a'\rangle |a\rangle\langle a'|$, where we used the eigenvalues and eigenvectors of the measured operator $\hat{A} = \sum_a \xi_a |a\rangle\langle a|$. In this approximation, the detector is

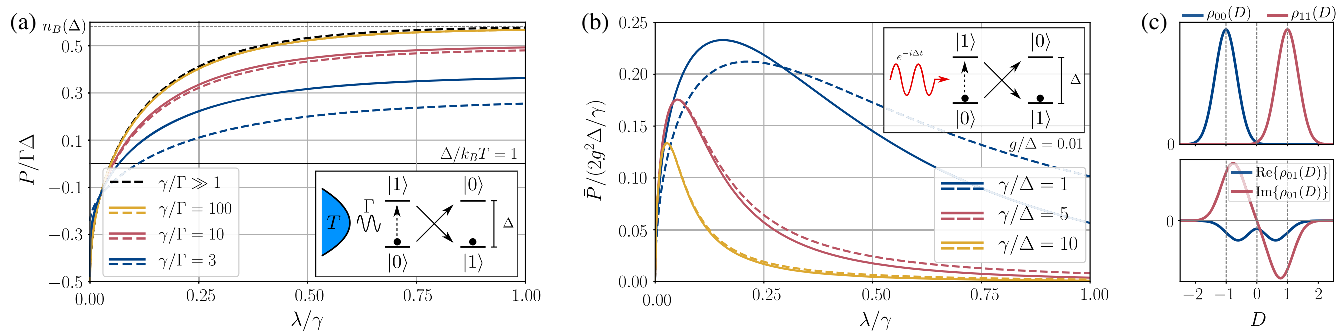

FIG. 2. Steady state power for classical (a) and quantum (b) toy models, varying the measurement strength λ. Solid lines obtained by numerically solving Eq. (1), dashed lines obtained analytically using the separation of timescales technique. The separation of timescales assumption breaks down when system and detector timescales are comparable. (a) The inset illustrates a feedback protocol of a classical two-level system coupled to a thermal reservoir. When excited (dashed arrow), the levels are flipped (solid arrows), extracting energy. For strong measurements ($\lambda \gg \gamma$), the average occupation of the bath [$n_B(\Delta)$] sets an upper limit on extracted power, see dashed grey line, and is only reached for fast detectors ($\gamma/\Gamma \gg 1$) [cf. Eq. (6)]. For weak measurements ($\lambda \ll \gamma$), feedback is applied randomly and energy is dissipated into the reservoir. (b) The inset depicts a feedback protocol for a qubit, coherently driven by an external driving field. The protocol is identical to (a). For strong measurements, the power vanishes because of the quantum Zeno effect. For weak measurements, no power can be extracted as feedback is applied randomly. (c) Visualization of $\hat{\rho}_t(D)$ for the quantum toy model, with stationary matrix elements $\rho_{ab}(D) = \langle a|\hat{\rho}_t(D)|b\rangle$. Here we use $g/\Delta = 0.01$ and $\gamma = \Delta = \lambda$. Top panel: diagonal elements of $\hat{\rho}_t(D)$. Bottom panel: real and imaginary part of $\rho_{01}(D)$.

always in a system dependent stationary distribution $\pi_{aa'}(D)$. This is justified for $\gamma \gg \Gamma$, where changes of the system occur with a rate much smaller than the inverse detector relaxation time. Inserting this approximation in Eq. (1) results in $\mathcal{L}_0 = \int dD\, \mathcal{L}(D)[\sum_{aa'} \pi_{aa'}(D)\mathcal{V}_{aa'}]$, describing the system dynamics for a detector with zero delay time. The first order correction $\gamma^{-1}\mathcal{L}_{\mathrm{corr}}$ accounts for the lag of the detector due to its finite response time γ^{-1}. As usual in linear response theory, this correction can be written in terms of time-integrated correlation functions; see SM [77]. Note that λ plays a special role in the separation of timescales since it appears in both the first and second line of Eq. (1). In general, Eq. (2) is thus only justified for $\lambda \ll \gamma$. Here we keep λ/γ arbitrary as there are scenarios where Eq. (2) also holds for strong measurements; see below.

We emphasize that Eq. (2) describes arbitrary feedback protocols, both linear and nonlinear in D. As a consistency check, we recover the Wiseman-Milburn equation [63] from Eq. (1) by employing the separation of timescales approximation to first order in $1/\gamma$, using a linear feedback Liouville superoperator $\mathcal{L}(D)\hat{\rho} = \mathcal{L}\hat{\rho} - iD[\hat{F}, \hat{\rho}]$, with feedback Hamiltonian $\hat{F}$, and taking the infinite bandwidth limit (see SM [77]). Our formalism thus generalizes the important earlier work of Ref. [63] to nonlinear feedback protocols.

In the following, we highlight the usefulness of Eq. (1) by studying protocols for power production in two toy models.

Classical toy model.—By classical system, we refer to a situation with discrete energy levels, but where the density matrix remains diagonal in the energy basis at all times. This can be achieved either by suppressing quantum coherence by environmental noise or by decoupling the diagonal and off-diagonal elements of $\hat{\rho}_t$ (see SM for details [77]). Under these conditions, $[\hat{\rho}_t(D), \hat{A}] = 0$ and the backaction term in Eq. (1) has no influence on the dynamics. To facilitate a comparison between the classical and quantum models, we use the same notation. We consider a classical two-level system, with states $|0\rangle$ and $|1\rangle$, coupled to a thermal reservoir at temperature T; see inset of Fig. 2(a). The system and reservoir exchange energy quanta with energy Δ at rate Γ. The state of the system is continuously monitored by measuring the observable $\hat{A} = \hat{\sigma}_z$, with Pauli-Z operator $\hat{\sigma}_z = |1\rangle\langle 1| - |0\rangle\langle 0|$, such that whenever the measurement outcome $D < 0$ ($D \geq 0$) for an ideal detector (low noise and delay), the system resides in $|0\rangle$ ($|1\rangle$). Feedback is incorporated by flipping the levels according to the solid arrows in Fig. 2(a) when an excitation is detected, i.e., when D changes sign, thereby extracting energy from the reservoir. The Hamiltonian is given by $\hat{H}(D) = [1 - \theta(D)]\Delta|1\rangle\langle 1| + \theta(D)\Delta|0\rangle\langle 0|$, where $\theta(D)$ is the Heaviside step function. Note that $[\hat{H}(D), \hat{A}] = 0$, ensuring that $\hat{\rho}_t(D)$ remains diagonal in the energy basis. The feedback protocol is represented by the Liouville superoperator,

$$\mathcal{L}(D) = [1 - \theta(D)]\mathcal{L}_- + \theta(D)\mathcal{L}_+, \tag{4}$$

where $\mathcal{L}_-\hat{\rho} = \Gamma n_B(\Delta)\mathcal{D}[\hat{\sigma}^\dagger]\hat{\rho} + \Gamma[n_B(\Delta) + 1]\mathcal{D}[\hat{\sigma}]\hat{\rho}$ is the protocol applied for $D < 0$, and $\mathcal{L}_+\hat{\rho} = \Gamma[n_B(\Delta) + 1]\mathcal{D}[\hat{\sigma}^\dagger]\hat{\rho} + \Gamma n_B(\Delta)\mathcal{D}[\hat{\sigma}]\hat{\rho}$ is the protocol applied for $D \geq 0$, with system ladder operator $\hat{\sigma} = |0\rangle\langle 1|$, and

Bose-Einstein distribution $n_B(x) = [\exp(x/k_B T) - 1]^{-1}$, with x denoting energy and k_B the Boltzmann constant.

Employing the separation of timescales technique, using $\gamma \gg \Gamma$ with Eqs. (2) and (3), the system evolves, to zeroth order in $1/\gamma$, according to the feedback Liouville superoperator,

$$\mathcal{L}_0 = [(1-\eta)\mathcal{L}_- + \eta\mathcal{L}_+]\mathcal{V}_{00} + [\eta\mathcal{L}_- + (1-\eta)\mathcal{L}_+]\mathcal{V}_{11}, \qquad (5)$$

where we introduced the feedback error probability $\eta = [1 - \mathrm{erf}(2\sqrt{\lambda/\gamma})]/2$ for a single feedback event, where $\mathrm{erf}(\cdot)$ is the error function and $0 \leq \eta \leq 1/2$. Feedback is applied incorrectly when the measurement outcome does not reflect the true system state. Note that weak (strong) measurements yield high (low) detector noise and increase (decrease) the error probability.

To zeroth order in $1/\gamma$, the average power production reads

$$P = \Gamma\Delta[(1-\eta)n_B(\Delta) - \eta[n_B(\Delta) + 1]], \qquad (6)$$

where $P > 0$ corresponds to extracting energy from the bath. For strong measurements ($\eta \to 0$), feedback is consistently applied correctly and energy is only extracted from the reservoir. The maximum extraction rate $P = \Gamma\Delta n_B(\Delta)$ is limited by the coupling Γ and the average occupation $n_B(\Delta)$ of the bath. For weak measurements, feedback errors together with the asymmetry between excitation and deexcitation rates lead to a net dissipation of energy. Interestingly, the maximum dissipation rate $P = -\Gamma\Delta/2$ is independent of $n_B(\Delta)$. Equation (6) is plotted with a black, dashed line in Fig. 2, illustrating the behavior for weak and strong measurements. Additionally, we computed the power by (i) numerically solving Eq. (1) (solid colored lines) and (ii) using the separation of timescales technique to first order in $1/\gamma$ (dashed colored lines) (see SM for details [77]). As γ decreases, the extracted power decreases because the detector can no longer resolve fast changes in the system, missing opportunities to extract energy. The separation of timescales approximation gradually breaks down as γ and Γ become comparable.

Following Ref. [88], in the longtime limit, Eq. (5) implies the detailed fluctuation theorem,

$$\frac{P(-m)}{P(m)} = e^{m(\Delta/k_B T - \ln[(1-\eta)/\eta])}, \qquad (7)$$

for the number of extracted energy quanta m, where $m > 0$ ($m < 0$) corresponds to extracting (dissipating) energy from the bath. The term Δ/T is the entropy change in the bath related to the exchange of a single quantum. The information term $\ln[(1-\eta)/\eta]$ is given by the log-odds of not making an error and can be interpreted as the difference in information content between correctly and incorrectly applying feedback. Note that most information from the continuous measurement is discarded—it is only the information during a change in the system state that matters. In the error-free limit, $\eta \to 0$, the information term diverges, illustrating absolute irreversibility; i.e., all excitations are extracted. See SM for a derivation of Eq. (7) [77].

Quantum toy model.—We consider a qubit coherently driven by an external driving field; see inset of Fig. 2(b). Measurement and feedback are identical to the classical toy model, now extracting energy from the driving field. The feedback protocol is described by $\mathcal{L}_t(D)\hat{\rho} = -i[\hat{H}_t(D), \hat{\rho}]$ with Hamiltonian

$$\hat{H}_t(D) = [1-\theta(D)]\Delta|1\rangle\langle 1| + \theta(D)\Delta|0\rangle\langle 0| + g\cos(\Delta t)\hat{\sigma}_x, \qquad (8)$$

where Δ is the qubit level spacing, g the strength of the qubit-driving field coupling, and $\hat{\sigma}_x$ the Pauli-X operator.

Separating system and detector timescales to first order in $1/\gamma$ results in system Liouville superoperator (details in SM [77]),

$$[\mathcal{L}_0 + \lambda\mathcal{D}[\hat{\sigma}_z] + \gamma^{-1}\mathcal{L}_{\mathrm{corr}}]\hat{\rho} = -ig\cos(\Delta t)[\hat{\sigma}_x, \hat{\rho}] + \tilde{\lambda}\mathcal{D}[\hat{\sigma}_z]\hat{\rho} - \frac{2\Delta g}{\gamma} D_0 \cos(\Delta t)\hat{\sigma}_x, \qquad (9)$$

with effective dephasing rate $\tilde{\lambda} = \lambda + \Delta^2\ln(2)/2\gamma$, and coefficient $D_0 = 2\sqrt{\lambda/\pi\gamma}\,{}_2F_2(1/2, 1/2; 3/2, 3/2; -4\lambda/\gamma)$, where ${}_2F_2(\cdot)$ is a generalized hypergeometric function. The first term on the rhs of Eq. (9) represents the coherent drive, while the second term describes dephasing due to measurement and feedback. The third term is a source for quantum coherence, stabilizing the coherence in the longtime limit. We emphasize that the first order correction is essential to compute the power as the steady state coherence vanishes to leading order, and hence, no power can be extracted. Note that the third term, which goes beyond leading order, can lead to negativities in $\hat{\rho}_t$, which is of no concern in the separation of timescales regime where the term is small. We stress that this term is trace preserving as $\hat{\sigma}_x$ is traceless.

The average power of the system is given by $P(t) = \mathrm{tr}\{[\partial_t \hat{H}(D)]\hat{\rho}_t\}$, where power is extracted [dissipated] when $P(t) > 0$ [$P(t) < 0$]. Over one driving period $\tau = 2\pi/\Delta$, the time averaged power reads

$$\bar{P} = \frac{2g^2\Delta}{\gamma} D_0 \frac{\Delta^2}{\Delta^2 + 4\tilde{\lambda}^2}. \qquad (10)$$

For strong measurements $\lambda \gg \gamma$, the power vanishes because of the quantum Zeno effect. For weak measurements $\lambda \ll \gamma$, large detector noise leads to completely random feedback, and the power goes to zero because of the symmetric driving. This is highlighted in Fig. 2(b), where we plot Eq. (10) as dashed lines. The solid lines were computed numerically by solving the full Eq. (1). The corresponding steady state matrix elements of $\hat{\rho}_t(D)$ are visualized in Fig. 2(c) (details in SM [77]). Similar to the classical toy model, the separation of timescales assumption breaks down when system and detector timescales are comparable.

Outline derivation main result.—To outline the main steps in the derivation of Eq. (1), we start by describing the continuous measurement. For a single instantaneous measurement, the system state $\hat{\rho}_t$ transforms as

$$\hat{\rho}_t(z) = \hat{K}(z)\hat{\rho}_t\hat{K}^\dagger(z), \tag{11}$$

where $\hat{K}(z)$ is the measurement operator for obtaining outcome z, obeying the completeness relation $\int dz\, \hat{K}^\dagger(z)\hat{K}(z) = \mathbb{1}$, $\mathrm{tr}\{\hat{\rho}_t(z)\}$ is the probability of obtaining z, and $\int dz\, \hat{\rho}_t(z)$ is the system state for an unknown measurement outcome. Stressing that temporal coarse graining results in Gaussian noise for any measurement operator [89], we consider Gaussian measurement operators [89,90],

$$\hat{K}(z) = \left(\frac{2\lambda\delta t}{\pi}\right)^{1/4} e^{-\lambda\delta t(z-\hat{A})^2}, \tag{12}$$

where δt is the time between measurements. A weak continuous measurement is obtained by repeatedly measuring the system, taking the limit $\lambda\delta t \to 0$ for a fixed measurement strength λ. In this limit, the sequence of outcomes becomes a continuous signal $z(t)$.

The detector bandwidth γ is introduced through a low-pass frequency filter [1,12,91–95],

$$D(t) = \int_{-\infty}^{t} ds\, \gamma e^{-\gamma(t-s)} z(s), \tag{13}$$

such that the measurement outcome $D(t)$ is a smoothened version of the signal $z(t)$. The filter reduces the high frequency measurement noise and introduces a detector delay. This provides a realistic detector model, but the filter is also necessary for nonlinear feedback protocols because higher orders of $z(t)$ are ill defined due to its white noise spectrum which includes diverging frequencies [1,12,93].

Feedback is incorporated by controlling the system time evolution in between measurements, i.e., making the Liouville superoperator $\mathcal{L}(D)$ dependent on the frequency filtered measurement outcome D. Combining time evolution due to measurements and due to the Liouvillian, we find Eq. (1) in the continuous limit $\delta t \to 0$. The derivation can be carried out either in the framework of stochastic calculus following the methods outlined in Refs. [68,89] or under the rules of conventional calculus. See details in SM [77].

Conclusions.—We have derived a Fokker-Planck master equation for continuous feedback control, describing the joint system-detector dynamics for detectors with finite bandwidth. By separating system and detector timescales, we obtain a Markovian master equation for the system alone, opening a new avenue for analytical modeling of nonlinear feedback protocols. The Markovian description further implies fluctuation theorems, providing insight into the connection between thermodynamics and information theory. With two simple toy models, we highlighted the usefulness of our formalism, showing that it can be applied to a large variety of systems in both the classical and quantum regimes. Future endeavors include extensions of the formalism to include non-Markovian effects and state-estimation feedback [61,96].

We thank Mark T. Mitchison for fruitful discussions. This research was supported by FQXi Grant No. FQXi-IAF19-07 from the Foundational Questions Institute Fund, a donor advised fund of Silicon Valley Community Foundation. P. S. and B. A.-A. were supported by the Swedish Research Council, Grant No. 2018-03921. P. P. P. acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 796700, from the Swedish Research Council (Starting Grant No. 2020-03362), and from the Swiss National Science Foundation (Eccellenza Professorial Fellowship PCEFP2_194268).

*bjorn.annby-andersson@teorfys.lu.se

[1] M. Sarovar, C. Ahn, K. Jacobs, and G. J. Milburn, Practical scheme for error control using feedback, Phys. Rev. A **69**, 052324 (2004).

[2] D. Risté, M. Dukalski, C. A. Watson, G. De Lange, M. J. Tiggelman, Y. M. Blanter, K. W. Lehnert, R. N. Schouten, and L. DiCarlo, Deterministic entanglement of superconducting qubits by parity measurement and feedback, Nature (London) **502**, 350 (2013).

[3] A. D. Ludlow, M. M. Boyd, J. Ye, E. Peik, and P. O. Schmidt, Optical atomic clocks, Rev. Mod. Phys. **87**, 637 (2015).

[4] W. P. Smith, J. E. Reiner, L. A. Orozco, S. Kuhr, and H. M. Wiseman, Capture and Release of a Conditional State of a Cavity QED System by Quantum Feedback, Phys. Rev. Lett. **89**, 133601 (2002).

[5] C. Sayrin, I. Dotsenko, X. Zhou, B. Peaudecerf, T. Rybarczyk, S. Gleyzes, P. Rouchon, M. Mirrahimi, H. Amini, M. Brune, J.-M. Raimond, and S. Haroche, Real-time quantum feedback prepares and stabilizes photon number states, Nature (London) **477**, 73 (2011).

[6] R. Vijay, C. Macklin, D. H. Slichter, S. J. Weber, K. W. Murch, R. Naik, A. N. Korotkov, and I. Siddiqi, Stabilizing Rabi oscillations in a superconducting qubit using quantum feedback, Nature (London) **490**, 77 (2012).

[7] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi, Adaptive Homodyne Measurement of Optical Phase, Phys. Rev. Lett. **89**, 133602 (2002).

[8] B. D'Urso, B. Odom, and G. Gabrielse, Feedback Cooling of a One-Electron Oscillator, Phys. Rev. Lett. **90**, 043001 (2003).

[9] P. Bushev, D. Rotter, A. Wilson, F. Dubin, C. Becher, J. Eschner, R. Blatt, V. Steixner, P. Rabl, and P. Zoller, Feedback Cooling of a Single Trapped Ion, Phys. Rev. Lett. **96**, 043003 (2006).

[10] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Entanglement-free Heisenberg-limited phase estimation, Nature (London) **450**, 393 (2007).

[11] G. G. Gillett, R. B. Dalton, B. P. Lanyon, M. P. Almeida, M. Barbieri, G. J. Pryde, J. L. O'Brien, K. J. Resch, S. D. Bartlett, and A. G. White, Experimental Feedback Control of Quantum Systems Using Weak Measurements, Phys. Rev. Lett. **104**, 080503 (2010).

[12] T. A. Wheatley, D. W. Berry, H. Yonezawa, D. Nakane, H. Arao, D. T. Pope, T. C. Ralph, H. M. Wiseman, A. Furusawa, and E. H. Huntington, Adaptive Optical Phase Estimation Using Time-Symmetric Quantum Smoothing, Phys. Rev. Lett. **104**, 093601 (2010).

[13] G.-Y. Xiang, B. L. Higgins, D. W. Berry, H. M. Wiseman, and G. J. Pryde, Entanglement-enhanced measurement of a completely unknown optical phase, Nat. Photonics **5**, 43 (2011).

[14] X. Zhou, I. Dotsenko, B. Peaudecerf, T. Rybarczyk, C. Sayrin, S. Gleyzes, J. M. Raimond, M. Brune, and S. Haroche, Field Locked to a Fock State by Quantum Feedback with Single Photon Corrections, Phys. Rev. Lett. **108**, 243602 (2012).

[15] R. Okamoto, M. Iefuji, S. Oyama, K. Yamagata, H. Imai, A. Fujiwara, and S. Takeuchi, Experimental Demonstration of Adaptive Quantum State Estimation, Phys. Rev. Lett. **109**, 130404 (2012).

[16] D. Ristè, C. C. Bultink, K. W. Lehnert, and L. DiCarlo, Feedback Control of a Solid-State Qubit Using High-Fidelity Projective Measurement, Phys. Rev. Lett. **109**, 240502 (2012).

[17] H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arao, K. Ohki, K. Tsumura, D. W. Berry, T. C. Ralph, H. M. Wiseman, E. H. Huntington, and A. Furusawa, Quantum-enhanced optical-phase tracking, Science **337**, 1514 (2012).

[18] Z. K. Minev, S. O. Mundhada, S. Shankar, P. Reinhold, R. Gutirrez-Juregui, R. J. Schoelkopf, M. Mirrahimi, H. J. Carmichael, and M. H. Devoret, To catch and reverse a quantum jump mid-flight, Nature (London) **570**, 200 (2019).

[19] S. Vinjanampathy and J. Anders, Quantum thermodynamics, Contemp. Phys. **57**, 545 (2016).

[20] J. C. Maxwell, *Theory of Heat* (Longmans, Green, and Co., London, 1871).

[21] *Maxwell's Demon 2 Entropy, Classical and Quantum Information, Computing*, edited by H. S. Leff and A. F. Rex (CRC Press, Boca Raton, FL, 2002).

[22] K. Maruyama, F. Nori, and V. Vedral, Colloquium: The physics of Maxwell's demon and information, Rev. Mod. Phys. **81**, 1 (2009).

[23] V. Serreli, C. F. Lee, E. R. Kay, and D. A. Leigh, A molecular information ratchet, Nature (London) **445**, 523 (2007).

[24] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality, Nat. Phys. **6**, 988 (2010).

[25] J. V. Koski, V. F. Maisi, J. P. Pekola, and D. V. Averin, Experimental realization of a Szilard engine with a single electron, Proc. Natl. Acad. Sci. U.S.A. **111**, 13786 (2014).

[26] J. V. Koski, V. F. Maisi, T. Sagawa, and J. P. Pekola, Experimental Observation of the Role of Mutual Information in the Nonequilibrium Dynamics of a Maxwell Demon, Phys. Rev. Lett. **113**, 030601 (2014).

[27] K. Chida, S. Desai, K. Nishiguchi, and A. Fujiwara, Power generator driven by Maxwell's demon, Nat. Commun. **8**, 15310 (2017).

[28] A. Kumar, T.-Y. Wu, F. Giraldo, and D. S. Weiss, Sorting ultracold atoms in a three-dimensional optical lattice in a realization of Maxwell's demon, Nature (London) **561**, 83 (2018).

[29] D. Barker, M. Scandi, S. Lehmann, C. Thelander, K. A. Dick, M. Perarnau-Llobet, and V. F. Maisi, Experimental Verification of the Work Fluctuation-Dissipation Relation for Information-to-Work Conversion, Phys. Rev. Lett. **128**, 040602 (2022).

[30] M. D. Vidrighin, O. Dahlsten, M. Barbieri, M. S. Kim, V. Vedral, and I. A. Walmsley, Photonic Maxwell's Demon, Phys. Rev. Lett. **116**, 050401 (2016).

[31] N. Cottet, S. Jezouin, L. Bretheau, P. Campagne-Ibarcq, Q. Ficheux, J. Anders, A. Auffèves, R. Azouit, P. Rouchon, and B. Huard, Observing a quantum Maxwell demon at work, Proc. Natl. Acad. Sci. U.S.A. **114**, 7561 (2017).

[32] Y. Masuyama, K. Funo, Y. Murashita, A. Noguchi, S. Kono, Y. Tabuchi, R. Yamazaki, M. Ueda, and Y. Nakamura, Information-to-work conversion by Maxwell's demon in a superconducting circuit quantum electrodynamical system, Nat. Commun. **9**, 1291 (2018).

[33] M. Naghiloo, J. J. Alonso, A. Romito, E. Lutz, and K. W. Murch, Information Gain and Loss for a Quantum Maxwell's Demon, Phys. Rev. Lett. **121**, 030604 (2018).

[34] M. Ribezzi-Crivellari and F. Ritort, Large work extraction and the Landauer limit in a continuous Maxwell demon, Nat. Phys. **15**, 660 (2019).

[35] T. Sagawa, Thermodynamics of information processing in small systems, Prog. Theor. Phys. **127**, 1 (2012).

[36] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, Thermodynamics of information, Nat. Phys. **11**, 131 (2015).

[37] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, The role of quantum information in thermodynamics—a topical review, J. Phys. A **49**, 143001 (2016).

[38] T. Sagawa and M. Ueda, Second Law of Thermodynamics with Discrete Quantum Feedback Control, Phys. Rev. Lett. **100**, 080403 (2008).

[39] T. Sagawa and M. Ueda, Generalized Jarzynski Equality Under Nonequilibrium Feedback Control, Phys. Rev. Lett. **104**, 090602 (2010).

[40] M. Ponmurugan, Generalized detailed fluctuation theorem under nonequilibrium feedback control, Phys. Rev. E **82**, 031129 (2010).

[41] J. M. Horowitz and S. Vaikuntanathan, Nonequilibrium detailed fluctuation theorem for repeated discrete feedback, Phys. Rev. E **82**, 061120 (2010).
[42] Y. Morikuni and H. Tasaki, Quantum Jarzynski-Sagawa-Ueda relations, J. Stat. Phys. **143**, 1 (2011).
[43] T. Sagawa and M. Ueda, Fluctuation Theorem with Information Exchange: Role of Correlations in Stochastic Thermodynamics, Phys. Rev. Lett. **109**, 180602 (2012).
[44] T. Sagawa and M. Ueda, Nonequilibrium thermodynamics of feedback control, Phys. Rev. E **85**, 021104 (2012).
[45] D. Abreu and U. Seifert, Thermodynamics of Genuine Nonequilibrium States Under Feedback Control, Phys. Rev. Lett. **108**, 030601 (2012).
[46] K. Funo, Y. Watanabe, and M. Ueda, Integral quantum fluctuation theorems under measurement and feedback control, Phys. Rev. E **88**, 052121 (2013).
[47] C. W. Wächtler, P. Strasberg, and T. Brandes, Stochastic thermodynamics based on incomplete information: Generalized Jarzynski equality with measurement errors with or without feedback, New J. Phys. **18**, 113042 (2016).
[48] P. P. Potts and P. Samuelsson, Detailed Fluctuation Relation for Arbitrary Measurement and Feedback Schemes, Phys. Rev. Lett. **121**, 210603 (2018).
[49] J. P. Pekola, Towards quantum thermodynamics in electronic circuits, Nat. Phys. **11**, 118 (2015).
[50] W. G. van der Wiel, S. De Franceschi, J. M. Elzerman, T. Fujisawa, S. Tarucha, and L. P. Kouwenhoven, Electron transport through double quantum dots, Rev. Mod. Phys. **75**, 1 (2002).
[51] M. Kjaergaard, M. E. Schwartz, J. Braumller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, Superconducting qubits: Current state of play, Annu. Rev. Condens. Matter Phys. **11**, 369 (2020).
[52] C. Fasth, A. Fuhrer, M. T. Björk, and L. Samuelson, Tunable double quantum dots in InAs nanowires defined by local gate electrodes, Nano Lett. **5**, 1487 (2005).
[53] K. W. Murch, R. Vijay, and I. Siddiqi, Weak measurement and feedback in superconducting quantum circuits, in *Superconducting Devices in Quantum Optics*, edited by R. H. Hadfield and G. Johansson (Springer International Publishing, Cham, 2016), pp. 163–185.
[54] D. Barker, S. Lehmann, L. Namazi, M. Nilsson, C. Thelander, K. A. Dick, and V. F. Maisi, Individually addressable double quantum dots formed with nanowire polytypes and identified by epitaxial markers, Appl. Phys. Lett. **114** (2019).
[55] B. Küng, C. Rössler, M. Beck, M. Marthaler, D. S. Golubev, Y. Utsumi, T. Ihn, and K. Ensslin, Irreversibility on the Level of Single-Electron Tunneling, Phys. Rev. X **2**, 011001 (2012).
[56] A. Hofmann, V. F. Maisi, J. Basset, C. Reichl, W. Wegscheider, T. Ihn, K. Ensslin, and C. Jarzynski, Heat dissipation and fluctuations in a driven quantum dot, Phys. Status Solidi B **254**, 1600546 (2017).
[57] D. Ristè, C. C. Bultink, K. W. Lehnert, and L. DiCarlo, Feedback Control of a Solid-State Qubit Using High-Fidelity Projective Measurement, Phys. Rev. Lett. **109**, 240502 (2012).
[58] P. Campagne-Ibarcq, E. Flurin, N. Roch, D. Darson, P. Morfin, M. Mirrahimi, M. H. Devoret, F. Mallet, and B. Huard, Persistent Control of a Superconducting Qubit by Stroboscopic Measurement Feedback, Phys. Rev. X **3**, 021008 (2013).
[59] V. P. Belavkin, On the theory of controlling observable quantum systems, Autom. Remote Control (Engl. Transl.) **44**, 178 (1983).
[60] V. P. Belavkin, Non-demolition measurement and control in quantum dynamical systems, in *Information Complexity and Control in Quantum Physics*, edited by A. Blaquiere, S. Diner, and G. Lochak (Springer, Vienna, 1987), pp. 311–329.
[61] V. P. Belavkin, Quantum stochastic calculus and quantum nonlinear filtering, J. Multivariate Anal. **42**, 171 (1992).
[62] V. P. Belavkin, Quantum continual measurements and a posteriori collapse on CCR, Commun. Math. Phys. **146**, 611 (1992).
[63] H. M. Wiseman and G. J. Milburn, Quantum Theory of Optical Feedback via Homodyne Detection, Phys. Rev. Lett. **70**, 548 (1993).
[64] H. M. Wiseman, Quantum theory of continuous feedback, Phys. Rev. A **49**, 2133 (1994).
[65] M. Yanagisawa and H. Kimura, A control problem for Gaussian states, in *Learning, Control and Hybrid Systems* (Springer, New York, 1999), pp. 294–313.
[66] A. C. Doherty and K. Jacobs, Feedback control of quantum systems using continuous state estimation, Phys. Rev. A **60**, 2700 (1999).
[67] A. N. Korotkov, Selective quantum evolution of a qubit state due to continuous measurement, Phys. Rev. B **63**, 115403 (2001).
[68] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge Univeristy Press, Cambridge, England, 2010).
[69] K. Jacobs, *Quantum Measurement Theory and its Applications* (Cambridge University Press, Cambridge, England, 2014).
[70] J. Zhang, Y.-X. Liu, R.-B. Wu, K. Jacobs, and F. Nori, Quantum feedback: Theory, experiments, and applications, Phys. Rep. **679**, 1 (2017).
[71] D. E. Kirk, *Optimal Control Theory—An Introduction* (Dover Publications, Inc., New York, 2004).
[72] V. Cavina, A. Mari, A. Carlini, and V. Giovannetti, Optimal thermodynamic control in open quantum systems, Phys. Rev. A **98**, 012139 (2018).
[73] G. Schaller, C. Emary, G. Kiesslich, and T. Brandes, Probing the power of an electronic Maxwell's demon: Single-electron transistor monitored by a quantum point contact, Phys. Rev. B **84**, 085418 (2011).
[74] D. V. Averin, M. Möttönen, and J. P. Pekola, Maxwell's demon based on a single-electron pump, Phys. Rev. B **84**, 245448 (2011).
[75] B. Annby-Andersson, P. Samuelsson, V. F. Maisi, and P. P. Potts, Maxwell's demon in a double quantum dot with continuous charge detection, Phys. Rev. B **101**, 165404 (2020).
[76] R. Porotti, A. Essig, B. Huard, and F. Marquardt, Deep reinforcement learning for quantum state preparation with weak nonlinear measurements, Quantum **6**, 747 (2022).

[77] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.129.050401 for detailed derivations of the results in this Letter, which includes Refs. [79–87].

[78] V. P. Belavkin, Nondemolition measurements, nonlinear filtering and dynamic programming of quantum stochastic processes, in *Modeling and Control of Systems*, edited by A. Blaquiére (Springer, Berlin, 1989), p. 245.

[79] S. Nakajima, On quantum theory of transport phenomena: Steady diffusion, Prog. Theor. Phys. **20**, 948 (1958).

[80] R. Zwanzig, Ensemble method in the theory of irreversibility, J. Chem. Phys. **33**, 1338 (1960).

[81] D. Mandal and C. Jarzynski, Analysis of slow transitions between nonequilibrium steady states, J. Stat. Mech. (2016) 063204.

[82] M. Scandi and M. Perarnau-Llobet, Thermodynamic length in open quantum systems, Quantum **3**, 197 (2019).

[83] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (CRC Press, Boca Raton, FL, 2018).

[84] J. A. Gyamfi, Fundamentals of quantum mechanics in Liouville space, Eur. J. Phys. **41**, 063002 (2020).

[85] D. Manzano, A short introduction to the Lindblad master equation, AIP Adv. **10**, 025106 (2020).

[86] G. Schaller, *Open Quantum Systems Far from Equilibrium* (Springer, New York, 2014), Vol. 881.

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

[88] M. Esposito and G. Schaller, Stochastic thermodynamics for Maxwell demon feedbacks, Europhys. Lett. **99**, 30003 (2012).

[89] K. Jacobs and D. A. Steck, A straightforward introduction to continuous quantum measurement, Contemp. Phys. **47**, 279 (2006).

[90] A. Bednorz, W. Belzig, and A. Nitzan, Nonclassical time correlation functions in continuous quantum measurement, New J. Phys. **14** (2012).

[91] P. Warszawski and H. M. Wiseman, Quantum trajectories for realistic photodetection: I. General formalism, J. Opt. B **5**, 1 (2003).

[92] P. Warszawski and H. M. Wiseman, Quantum trajectories for realistic photodetection: II. Application and analysis, J. Opt. B **5**, 15 (2003).

[93] M. Sarovar, H.-S. Goan, T. P. Spiller, and G. J. Milburn, High-fidelity measurement and quantum feedback control in circuit QED, Phys. Rev. A **72**, 062327 (2005).

[94] Z. Liu, L. Kuang, K. Hu, L. Xu, S. Wei, L. Guo, and X.-Q. Li, Deterministic creation and stabilization of entanglement in circuit QED by homodyne-mediated feedback control, Phys. Rev. A **82**, 032335 (2010).

[95] W. Feng, P. Wang, X. Ding, L. Xu, and X.-Q. Li, Generating and stabilizing the Greenberger-Horne-Zeilinger state in circuit QED: Joint measurement, Zeno effect, and feedback, Phys. Rev. A **83**, 042313 (2011).

[96] M. Yanagisawa, Non-Gaussian State Generation from Linear Elements via Feedback, Phys. Rev. Lett. **103**, 203601 (2009).