Quantum mean-square predictors and thermodynamics

Marcelo Janovitch [1,2,*] and Gabriel T. Landi [2,†]

[1]*Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland*
[2]*Instituto de Física da Universidade de São Paulo, 05314-970 São Paulo, Brazil*

Thermodynamic quantities, such as heat and work, are not functions of state, but rather of the process undergone by a physical system. Assessing them can therefore be difficult, since it requires probing the system at least twice. This is even more so when these quantities are to be assessed at the stochastic level. In this paper we show how to obtain optimal estimates of thermodynamic quantities solely from indirect measurement unravelings of an auxiliary system. The method always yields the true average, and the mean-squared error of the prediction is directly proportional to how well the method estimates the variance. As an application, we study energy fluctuations in a driven system, and in an avoided crossing work protocol.

I. INTRODUCTION

At the nanoscale, thermodynamic quantities, such as heat and work, may fluctuate significantly [1–10]. Properly accounting for these fluctuations is crucial, both from a fundamental as well as from an applied perspective. For instance, in nanoscale engines both the output power [11,12] and the efficiency [13–15] may fluctuate significantly, and the consequences of this are only now starting to be explored.

A unique feature of thermodynamics, however, is that the said quantities do not depend on the state of the system, but rather on the process or transformation in question. Within a quantum setting, several approaches have been developed for tackling this problem [16–28]. But still today, the most widely used is the two-point measurement (TPM) scheme [8–10], where a projective measurement in the system energy basis is applied before and after the process. Very often, however, the "system" is actually composed of multiple parts. For instance, in the case of a heat engine operating between two baths, the full statistics of heat and work is in general only accessible by performing a TPM in both the working fluid *and* the two baths [29], which can be prohibitive.

When not all parts of the system are accessible, it becomes necessary to develop strategies to estimate the thermodynamic quantities indirectly [22,30–33]. A concrete experimental example is the calorimetric method developed in Refs. [34–36], which estimates work in a quantum system by measuring the heat flowing to an ancilla (which acts as a finite reservoir). This paradigm is also quite frequently seen in open quantum systems, e.g., in the study of full counting statistics [9,37–40], where the heat statistics is determined from measurements in the bath.

Motivated by this issue, in this paper we approach the problem from the angle of statistical inference. We consider a system S undergoing a generic open process in contact with an ancillary system A, which can have any size. It is assumed that the system can never be measured; however, one has access to a certain set of outcomes of the ancilla, corresponding to the unravelings of the open system dynamics [41,42]. Our goal is to estimate the *changes* (at the stochastic level) in some system observable G_t, between two different times (e.g., 0 and τ). More concretely, we ask *what is the best possible prediction one can make about the changes in G, given only stochastic outcomes in the ancilla?* We formulate our results using the notion of statistical predictors [43]. Being inference based, our method is thus directly applicable to experiments. Our main result is Eq. (7), which specifies the optimal mean-squared predictor as a function only of the Kraus operators determining the process. This, as we show, can then be directly applied to thermodynamic protocols, e.g., for the estimation of work. However, the result also holds for any system operator, and thus extends beyond thermodynamics. To illustrate the ideas, we consider the estimation of energy fluctuations in a driven qubit, and the determination of work in an avoided crossing protocol.

II. STATEMENT OF THE PROBLEM

We consider a system S prepared in a state ρ_S, interacting with an ancillary system A, prepared in a state ρ_A. The total Hamiltonian is taken to have the general form $H(t) = H_S(t) + H_A(t) + V(t)$, and can in principle have any kind of time dependence. The process will thus in general involve the expenditure of work, as well as the exchange of heat between S and A. The interaction lasts for a time τ, after which their joint state will be $\rho_{SA}(\tau) = U_\tau(\rho_S \otimes \rho_A)U_\tau^\dagger$, where $U_\tau = \mathcal{T} e^{-i \int_0^\tau dt\, H(t)}$ and $\mathcal{T}$ is the time-ordering operator. The reduced state of the system will then be described by the quantum channel $\rho_S(\tau) = \Lambda[\rho_S]$, where $\Lambda[\bullet] = \mathrm{tr}_A\, U_\tau(\bullet \otimes \rho_A)U_\tau^\dagger$. If one is concerned with a closed system, Λ entails the open system dynamics due to the presence of the measurement device, and, if the system is already open, Λ can include both dynamics and measurement backaction.

*marcelobro.pera@gmail.com
†gtlandi@gmail.com

Our main interest is in the *changes* undergone by some system observable G_t; this could be, for instance, the local system energy $H_S(t)$. The average change in G is, of course,

$$\langle \Delta G \rangle = \text{tr}\{G_\tau \Lambda[\rho_S] - G_0 \rho_S\}. \tag{1}$$

But our interest is in going beyond the average, and accounting for the fluctuations. We do this using the two-point measurement (TPM) protocol [8]. Let $G_0 = \sum_{g_0} \lambda_{g_0} |g_0\rangle\langle g_0|$ and $G_\tau = \sum_{g_\tau} \lambda_{g_\tau} |g_\tau\rangle\langle g_\tau|$ denote the eigendecompositions of G_t at $t = 0$ and $t = \tau$. For now, we assume that $[G_0, \rho_S] = 0$; the case where this does not hold is subtle, and is discussed below. The TPM protocol consists in measuring the system in the eigenbases $\{|g_0\rangle\}$ and $\{|g_\tau\rangle\}$, respectively, before and after the channel $\Lambda[\bullet]$. This leads to the distribution

$$P(\Delta G) = \sum_{g_0, g_\tau} \langle g_\tau | \Lambda[|g_0\rangle\langle g_0|] | g_\tau \rangle p(g_0) \delta(\Delta G - (g_\tau - g_0)). \tag{2}$$

where $p(g_0) = \langle g_0 | \rho_S | g_0 \rangle$. From this, higher-order statistics can be readily computed. Of particular interest is the variance $\text{var}(\Delta G) = \langle \Delta G^2 \rangle - \langle \Delta G \rangle^2$, a quantity which has seen a surge of interest, e.g., in connection with the so-called thermodynamic uncertainty relations [44–50].

III. PREDICTORS

In this paper we assume that the distribution $P(\Delta G)$ is not accessible. Instead, all one has access to is a specific unraveling of the channel Λ,

$$\Lambda[\bullet] := \sum_\gamma M_\gamma \bullet M_\gamma^\dagger, \quad \sum_\gamma M_\gamma^\dagger M_\gamma = 1, \tag{3}$$

where $M_\gamma = \langle \gamma | U_\tau | 0 \rangle$ (we take $\rho_A = |0\rangle\langle 0|$ without loss of generality [51]). The unravelings fix a positive-operator-valued measure (POVM), $E_\gamma = M_\gamma^\dagger M_\gamma$, representing a set of measurements one has access to, whose outcomes occur with probability

$$P(\gamma) = \text{tr}(M_\gamma \rho_S M_\gamma^\dagger). \tag{4}$$

Any function $\Delta\mathcal{G}(\gamma)$ of the stochastic outcomes γ can now be viewed as a predictor of ΔG, in the sense that it conveys some information about it. Our goal is to determine which function $\Delta\mathcal{G}_{\text{opt}}$ yields the best possible prediction.

We quantify the quality of the prediction in terms of the mean-squared error (MSE)

$$\text{MSE}(\Delta\mathcal{G}) = \sum_\gamma \int d\Delta G (\Delta\mathcal{G}(\gamma) - \Delta G)^2 P(\Delta G, \gamma), \tag{5}$$

where

$$P(\Delta G, \gamma) = \sum_{g_0, g_\tau} |\langle g_\tau | M_\gamma | g_0 \rangle|^2 p(g_0) \delta(\Delta G - (g_\tau - g_0)) \tag{6}$$

is the joint distribution of ΔG and γ, which would have been obtained if both S and A had been measured. The main result of this letter is as follows:

Theorem. The predictor minimizing the MSE (5) is given by

$$\Delta\mathcal{G}_{\text{opt}}(\gamma) = \frac{1}{P(\gamma)} \left\langle M_\gamma^\dagger G_\tau M_\gamma - \frac{1}{2}\{M_\gamma^\dagger M_\gamma, G_0\} \right\rangle, \tag{7}$$
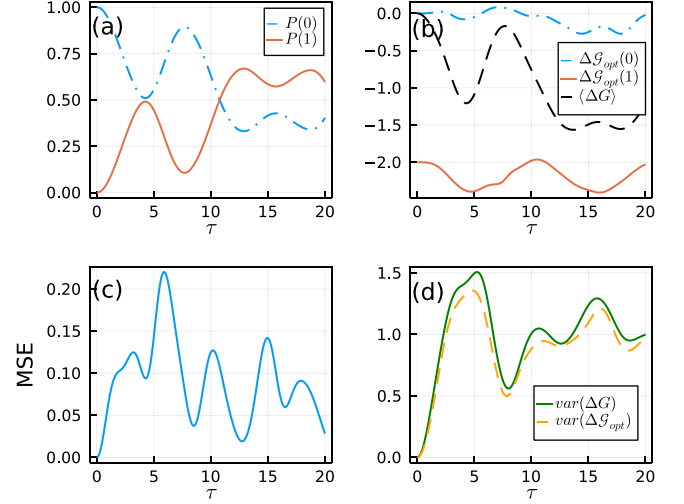

FIG. 1. Driven qubit model, as a function of the net driving time τ. (a) Probabilities $P(\gamma)$, Eq. (4); (b) optimal predictors $\Delta\mathcal{G}_{\text{opt}}(\gamma)$, Eq. (7). (c) Mean-squared error, Eq. (5). (d) True and predicted variances, $\text{var}(\Delta G)$ and $\text{var}(\Delta\mathcal{G}_{\text{opt}})$. Parameters: $\omega = \alpha = 1$, $g = \Omega = 0.3$, $s = 0.7$.

where the average is over the system's initial state ρ_S. This predictor always reproduces the true average (1),

$$\sum_\gamma P(\gamma)\Delta\mathcal{G}_{\text{opt}}(\gamma) = \langle \Delta G \rangle. \tag{8}$$

Moreover, it yields the MSE,

$$\text{MSE}(\Delta\mathcal{G}_{\text{opt}}) = \text{var}(\Delta G) - \text{var}(\Delta\mathcal{G}_{\text{opt}}), \tag{9}$$

which is thus simply the difference between the fluctuations of the true quantity and those of the predictor.

The proof of this result given in the Appendix. The fact that the optimal predictor always correctly reproduces the correct average behavior is noteworthy. In addition, it yields an MSE which directly links to the variance of ΔG. In fact, Eq. (9) implies that if ΔG does not fluctuate [$\text{var}(\Delta G) \equiv 0$], the same will also be true for $\Delta\mathcal{G}_{\text{opt}}$ (since the MSE is strictly non-negative). When applied in a thermodynamic scenario (as will be done below), this result is consistent with the no-go theorem of Ref. [52].

IV. ENERGY FLUCTUATIONS IN A DRIVEN QUBIT

To illustrate the idea, we consider a qubit system coupled to a qubit ancilla, with total Hamiltonian $H(t) = H_S(t) + \omega \sigma_z^A + g(\sigma_+^S \sigma_-^A + \sigma_-^S \sigma_+^A)$ and

$$H_S(t) = \omega \sigma_z^S + \alpha \sin(\Omega t)\sigma_x^S. \tag{10}$$

The system is thus driven by a time-dependent horizontal field $\alpha \sin(\Omega t)\sigma_x$. We look for estimates of how the system energy $[G_t \equiv H_S(t)]$ changes from time 0 up to some generic time τ.

The system starts in $\rho_S = s|0\rangle\langle 0| + (1-s)|1\rangle\langle 1|$, while the ancilla starts in $\rho_A = |0\rangle\langle 0|$. Moreover, we measure the ancillas in the computational basis, leading to two possible Kraus operators $M_0 = \langle 0|U_\tau|0\rangle$ and $M_1 = \langle 1|U_\tau|0\rangle$, where U_τ is obtained numerically. The results are shown in Fig. 1: The probabilities $P(\gamma)$ in Eq. (4) evolve as shown in Fig. 1(a), and

for each such γ, Eq. (7) predicts the results shown in Fig. 1(b). We also plot in the black-dashed lines the true average $\langle \Delta G \rangle$ [cf. Eq. (8)].

We can estimate the quality of our predictions by plotting the MSE (5) [Fig. 1(c)]. The predictions are better at some values of τ and worse at others. This can also be inferred by comparing the actual variance of ΔG with the predicted one $\text{var}(\Delta \mathcal{G}_{\text{opt}})$, as shown in Fig. 1(d). This result makes it particularly clear that very good predictions are possible, despite the complexity of the problem. Of course, this will depend on the model and the parameters.

V. INTERPRETATION, PARTICULAR CASES, AND GENERALIZATIONS

The optimal predictor is a function $f(\gamma)$, which provides the best possible guess for the changes in the system observable, from G_0 to G_τ, given only the specific unraveling in the ancilla. Computing it thus requires a *model*. In classical statistics, this is usually associated with a transition probability $P(y|x)$. In our case, this is determined by the the Kraus operators $\{M_\gamma\}$, as well as the system's initial state. In the way the problem has been formulated, the operator G can represent any system property. Hence, the results extend well beyond thermodynamics. The unique thermodynamic flavor appears in the fact that we are not estimating the value of G itself, but rather the *change* in G due to a certain process. In fact, if we artificially adjust $G_0 \equiv 0$, Eq. (7) reduces to $\Delta \mathcal{G}_{\text{opt}}(\gamma) = \text{tr}(G_\tau \rho_{\tau|\gamma})$, where $\rho_{\tau|\gamma} = M_\gamma \rho M_\gamma^\dagger / P(\gamma)$ is the conditional state, given γ. In this case the predictor is therefore only estimating the value of G_τ itself. Similarly, if $G_\tau = 0$, Eq. (7) will provide a retrodiction [53] of the initial value G_0 before the open dynamics. We also note that the choice of unraveling in the ancilla is arbitrary and model dependent. Hence, some unravelings will be more precise than others. This can be quantified using the MSE (9).

Equation (7) can be readily extended to more complex types of unravelings. For instance, continuous measurements will be characterized by a string of indices $\gamma = (\gamma_1, \ldots, \gamma_N)$. Equation (7) still applies. But now the Kraus operators will have the form $M_\gamma = M_{\gamma_N} \cdots M_{\gamma_1}$. Predictions can also be made when only parts of the outcomes are known. For instance, if there are two outcomes γ_1, γ_2, but only the latter is known, then the optimal predictor will be simply $\Delta \mathcal{G}(\gamma_2) = \sum_{\gamma_1} P(\gamma_1) \Delta \mathcal{G}_{\text{opt}}(\gamma_1, \gamma_2)$.

VI. THERMODYNAMIC APPLICATIONS

To make a connection with thermodynamics, we consider the problem of determining the work associated with an avoided crossing protocol [34]. We consider a qubit with $\rho_S = s|0\rangle\langle 0| + (1-s)|1\rangle\langle 1|$ and time-independent Hamiltonian $H_S = \omega \sigma_+ \sigma_-$. Initially the qubit is isolated and undergoes a unitary pulse described by $U_w = \sigma_x$. The work associated with this pulse can be either $\pm\omega$. In order to estimate this, we then couple the system to an ancilla, *after* the protocol. For simplicity, we take the ancilla at zero temperature, $\rho_A = |0\rangle\langle 0|$, and take $H_A = \omega|1\rangle\langle 1|$. Moreover, the interaction is assumed to be a swap unitary, which will hence transfer any excitations

TABLE I. Trajectories, probabilities, and predicted work [Eq. (11)] for the avoided crossing model. Here, $\Delta \mathcal{G}_{\text{opt}}$ is the predictor for the system energy, while $\mathcal{W}_{\text{opt}}$ refers to the work, which also includes the heat, as in Eq. (11).

M_γ	$P[\gamma]$	$\Delta \mathcal{G}_{\text{opt}}(\gamma)$	$\mathcal{W}_{\text{opt}}(\gamma)$	$\mathcal{W}_{\text{opt}}(\gamma)$ (coherent)
M_0	s	$-\omega$	$-\omega$	$-\omega + \frac{\omega}{2} \frac{\sin^2\theta}{1+(2s-1)\cos\theta}$
M_1	$1-s$	0	ω	$\omega - \frac{\omega}{2} \frac{\sin^2\theta}{1-(2s-1)\cos\theta}$

from S to A. Overall, this will thus be a two-step process, with a net unitary $U = U_{\text{swap}}(U_w \otimes I_A)$.

After the process, we measure the ancilla in the computational basis, which leads to two Kraus operators $M_\gamma = \langle \gamma | U_{\text{swap}} | 0 \rangle U_w$, with $\gamma = 0, 1$. We use the predictor (7), with $G_0 = G_\tau = H_S$. To obtain the actual work performed in the system, we must also include the heat transferred to the bath. In more complicated scenarios, such as strong coupling, one might also need to include the system-ancilla interaction. Our formalism currently cannot account for this. In the present case, however, the swap incurs no additional energy cost since the system and ancilla are resonant. Hence, the work will simply be given by $W = Q + \Delta H_S$. Consequently, the optimal predictor will be

$$\mathcal{W}_{\text{opt}}(\gamma) = Q(\gamma) + \Delta \mathcal{G}_{\text{opt}}(\gamma), \tag{11}$$

where, in this case, $Q(1) = \omega$ and $Q(0) = 0$. There are thus two possibilities (Table I). If M_1 is detected, the ancilla must have absorbed an excitation, which means that the work protocol must have excited the system from $|0\rangle \to |1\rangle$. The work associated to this is thus $+\omega$. Conversely, M_0 means that no excitation was detected in the ancilla, so that the state of the system before the swap must have been $|0\rangle$. Hence, the work performed must have been $-\omega$.

VII. INITIALLY COHERENT SYSTEMS

The optimal predictor (7) assumes that $[G_0, \rho_S] = 0$. When this is not the case, the problem becomes more delicate [52], and several approaches have been put forward for handling it [16–28]. Within our framework, the problem remains essentially unaltered, in the sense that the quality of any prediction is still given by the MSE (5). The subtle part is in how to define the joint distribution $P(\Delta G, \gamma)$, as this would generally be susceptible to the backaction from the first measurement. A way of constructing $P(\Delta G, \gamma)$, which does not suffer from this problem, is through the concept of quantum Bayesian networks (QBNs) [26,27]. Let $\rho_S = \sum_\alpha p_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|$, with the bases $\{|\psi_\alpha\rangle\}$ and $\{|g_0\rangle\}$ being generally incompatible. We then consider [26,27]

$$P(\Delta G, \gamma) = \sum_{g_0, g_\tau, \alpha} |\langle g_\tau | M_\gamma | \psi_\alpha \rangle|^2 p_{g_0|\alpha} p_\alpha \delta(\Delta G - (g_\tau - g_0)), \tag{12}$$

where $p_{g_0|\alpha} = |\langle g_0 | \psi_\alpha \rangle|^2$ is the conditional probability of observing $|g_0\rangle$ given $|\psi_\alpha\rangle$. The QBN (i) is always non-negative, (ii) reproduces the correct average (1), and (iii) reduces to the TPM (6) when $[G_0, \rho_S] = 0$. From an operational perspective, it was also recently shown that QBNs can be directly accessed
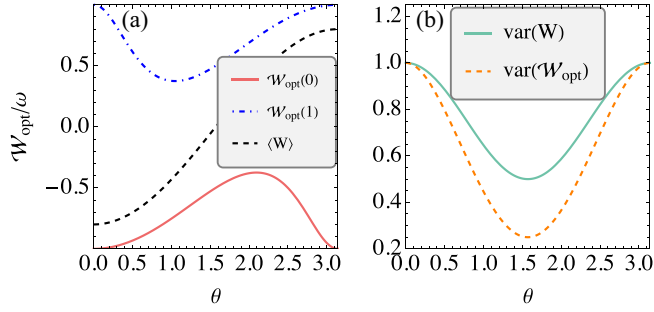
FIG. 2. (a) Work predictions in the avoided crossing model (last column of Table I), as a function of the system initial coherence angle θ, with $s = 0.9$. (b) True and predicted variances. Their difference yields the MSE (9), which quantifies the quality of the prediction.

in an experiment, provided one uses two identical copies of the system, together with measurement postprocessing [54].

Plugging (12) in (5) and repeating the same procedure in the Appendix yields the optimal predictor

$$\Delta \mathcal{G}_{\text{opt}}(\gamma) = \frac{1}{P(\gamma)} \left\langle M_\gamma^\dagger G_\tau M_\gamma - \frac{1}{2} \{ M_\gamma^\dagger M_\gamma, \mathbb{D}(G_0) \} \right\rangle, \quad (13)$$

where $\mathbb{D}(\bullet) = \sum_\alpha |\psi_\alpha\rangle\langle\psi_\alpha| \bullet |\psi_\alpha\rangle\langle\psi_\alpha|$ is the full dephasing operator in the basis of ρ_S. Compared to (7), the only difference is that G_0 is now replaced by $\mathbb{D}(G_0)$. This therefore clearly reduces to (7) when $[G_0, \rho_S] = 0$. Equation (13) continues to yield the correct average, as in Eq. (8). And the MSE is still given by Eq. (9).

As an application, we revisit the avoided crossing model in Table I and include the effects of initial coherence, by assuming that the system is prepared in $\rho_S^{\text{coh}} = e^{-i\theta\sigma_y/2} \rho_S e^{i\theta\sigma_y/2}$, where $\rho_S = s|0\rangle\langle 0| + (1 - s)|1\rangle\langle 1|$. The results are shown in the last column of Table I, as well as in Fig. 2(a). In the latter, we also plot the average work $\langle W \rangle$ for comparison. To assess the quality of the prediction, we plot in Fig. 2(b) the true and predicted variances, var(W) and var($\mathcal{W}_{\text{opt}}$). They are equal in the incoherent case and grow with increasing θ, being maximal when $\theta = \pi/2$. The presence of quantum coherences therefore generally degrades the quality of the prediction, which is intuitive, although this may not be true for other models.

VIII. DISCUSSION

We have put forth a framework for finding the optimal function to predict changes of a system property indirectly, solely from measurements in an auxiliary system. The method is applicable to both coherent and incoherent systems, and is summarized by an explicit expression for the predictor solely in terms of the Kraus operators determining the open system unraveling. The predictor can always capture the correct average, and the quality of the estimation is directly related to how well it captures the variance. These results are timely, due to the growing progress in the manipulation of coherent quantum devices and their potential thermodynamic applications. To illustrate that, we analyzed the problem of estimating energy fluctuations in driven qubit systems, and in the reconstruction of work in an avoided crossing model.

APPENDIX: PROOF OF THE MAIN THEOREM

We prove here the theorem in Eq. (7) of the main text, namely, that the predictor $\Delta \mathcal{G}_{\text{opt}}(\gamma)$ which minimizes the mean-squared error (5) is that given by Eq. (7). Define

$$f(\gamma) = \frac{1}{P(\gamma)} \int d\Delta G (\Delta G) P(\Delta G, \gamma). \quad (A1)$$

We consider the MSE in Eq. (5) for a generic predictor $\Delta \mathcal{G}$. Adding and subtracting $f(\gamma)$ leads to

$$\text{MSE}(\Delta \mathcal{G}) = \sum_\gamma \int d\Delta G (\Delta G - f(\gamma))^2 P(\Delta G, \gamma)$$

$$\times \sum_\gamma \int d\Delta G (\Delta \mathcal{G}(\gamma) - f(\gamma))^2 P(\Delta G, \gamma)$$

$$\times \sum_\gamma \int d\Delta G (\Delta \mathcal{G}(\gamma) - f(\gamma))[f(\gamma)$$

$$- \Delta G] P(\Delta G, \gamma). \quad (A2)$$

Due to (A1), however, the last term vanishes. Moreover, in the second term we can marginalize over ΔG and write the result as an average over the original distribution $P(\gamma)$:

$$\text{MSE}(\Delta \mathcal{G}) = \sum_\gamma \int d\Delta G (\Delta G - f(\gamma))^2 P(\Delta G, \gamma)$$

$$+ \sum_\gamma P(\gamma)[\Delta \mathcal{G}(\gamma) - f(\gamma)]^2. \quad (A3)$$

Since the last term is always non-negative, comparing this with Eq. (5) leads to

$$\text{MSE}(\Delta \mathcal{G}) \geqslant \text{MSE}(f), \quad (A4)$$

for any other predictor $\Delta \mathcal{G}$. Hence, the optimal predictor is exactly $\Delta \mathcal{G}_{\text{opt}}(\gamma) \equiv f(\gamma)$.

For the optimal predictor the MSE (A3) reduces to

$$\text{MSE}(\Delta \mathcal{G}) = \sum_\gamma \int d\Delta G (\Delta G - f(\gamma))^2 P(\Delta G, \gamma). \quad (A5)$$

Expanding and using Eq. (A1) leads to Eq. (9) of the main text.

Next, we obtain the explicit form in Eq. (7), in which the optimal predictor is cast solely in terms of the Kraus operators. Inserting Eq. (6) into Eq. (A1) leads to

$$\Delta \mathcal{G}_{\text{opt}}(\gamma) = \frac{1}{P(\gamma)} \sum_{g_0, g_\tau} |\langle g_\tau | M_\gamma | g_0 \rangle|^2 p(g_0)(g_\tau - g_0), \quad (A6)$$

which can also be written as

$$\Delta\mathcal{G}_{\text{opt}}(\gamma) = \frac{1}{P(\gamma)} \sum_{g_\tau} g_\tau \langle g_\tau | M_\gamma \rho_S M_\gamma^\dagger | g_\tau \rangle - \langle g_\tau | M_\gamma \rho_S G_0 M_\gamma^\dagger | g_\tau \rangle. \tag{A7}$$

Writing the remaining sum over g_τ as a trace yields precisely to Eq. (7). The coherent case, where $P(\Delta G, \gamma)$ is given by Eq. (12), is treated similarly.

[1] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993); D. J. Evans and D. J. Searles, Phys. Rev. E **50**, 1645 (1994).

[2] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995).

[3] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).

[4] C. Jarzynski, Phys. Rev. E **56**, 5018 (1997).

[5] G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998).

[6] J. Kurchan, J. Phys. A: Math. Gen. **31**, 3719 (1998).

[7] C. Jarzynski and D. K. Wójcik, Phys. Rev. Lett. **92**, 230602 (2004).

[8] P. Talkner, E. Lutz, and P. Hänggi, Phys. Rev. E **75**, 050102(R) (2007).

[9] M. Esposito, U. Harbola, and S. Mukamel, Rev. Mod. Phys. **81**, 1665 (2009).

[10] M. Campisi, P. Hänggi, and P. Talkner, Rev. Mod. Phys. **83**, 771 (2011).

[11] P. Pietzonka and U. Seifert, Phys. Rev. Lett. **120**, 190602 (2018).

[12] T. Denzler and E. Lutz, Phys. Rev. Research **3**, L032041 (2021).

[13] G. Verley, M. Esposito, T. Willaert, and C. Van den Broeck, Nat. Commun. **5**, 4721 (2014).

[14] T. Denzler and E. Lutz, Phys. Rev. Research **2**, 032062(R) (2020).

[15] T. Denzler, J. F. G. Santos, E. Lutz, and R. Serra, arXiv:2104.13427.

[16] A. E. Allahverdyan and T. M. Nieuwenhuizen, Phys. Rev. E **71**, 066102 (2005).

[17] A. Engel and R. Nolte, Europhys. Lett. **79**, 10003 (2007).

[18] A. E. Allahverdyan, Phys. Rev. E **90**, 032137 (2014).

[19] P. P. Hofer and A. A. Clerk, Phys. Rev. Lett. **116**, 013603 (2016).

[20] P. Solinas and S. Gasparinetti, Phys. Rev. A **94**, 052103 (2016).

[21] P. P. Hofer, J. B. Brask, M. Perarnau-Llobet, and N. Brunner, Phys. Rev. Lett. **119**, 090603 (2017).

[22] H. J. D. Miller and J. Anders, New J. Phys. **19**, 062001 (2017).

[23] P. Solinas, H. J. D. Miller, and J. Anders, Phys. Rev. A **96**, 052115 (2017).

[24] N. Yunger Halpern, B. Swingle, and J. Dressel, Phys. Rev. A **97**, 042105 (2018).

[25] A. Levy and M. Lostaglio, PRX Quantum **1**, 010309 (2020).

[26] J. J. Park, S. W. Kim, and V. Vedral, arXiv:1705.01750.

[27] K. Micadei, J. P. S. Peterson, A. M. Souza, R. S. Sarthour, I. S. Oliveira, G. T. Landi, R. M. Serra, and E. Lutz, arXiv:2012.06294.

[28] S. Gherardini, A. Belenchia, M. Paternostro, and A. Trombettoni, arXiv:2006.06208.

[29] P. Talkner, M. Campisi, and P. Hänggi, J. Stat. Mech., (2009) P02025.

[30] M. Mehboudi, A. Sanpera, and J. M. R. Parrondo, Quantum **2**, 66 (2018).

[31] G. Manzano, F. Plastina, and R. Zambrini, Phys. Rev. Lett. **121**, 120602 (2018).

[32] S. Deffner, J. P. Paz, and W. H. Zurek, Phys. Rev. E **94**, 010103(R) (2016).

[33] A. Sone, Q. Zhuang, and P. Cappellaro, Phys. Rev. A **98**, 012115 (2018).

[34] J. P. Pekola, P. Solinas, A. Shnirman, and D. V. Averin, New J. Phys. **15**, 115006 (2013).

[35] S. Suomela, R. Sampaio, and T. Ala-Nissila, Phys. Rev. E, **94**, 032138 (2016).

[36] B. Karimi, F. Brange, P. Samuelsson, and J. P. Pekola, Nat. Commun. **11**, 367 (2020).

[37] L. Levitov and G. Lesovik, JETP Lett. **58**, 230 (1993).

[38] B. K. Agarwalla and D. Segal, Phys. Rev. B **98**, 155438 (2018).

[39] S. Saryal, O. Sadekar, and B. K. Agarwalla, Phys. Rev. E **103**, 022141 (2021).

[40] G. Guarnieri, S. Campbell, J. Goold, S. Pigeon, B. Vacchini, and M. Paternostro, New J. Phys. **19**, 103038 (2017).

[41] H. J. Carmichael, Phys. Rev. X **5**, 031028 (2015).

[42] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, New York, 2009).

[43] P. J. Bickel and K. A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics, Volume I* (Prentice-Hall, Hoboken, NJ, 1977).

[44] A. C. Barato and U. Seifert, Phys. Rev. Lett. **114**, 158101 (2015).

[45] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Phys. Rev. Lett. **116**, 120601 (2016).

[46] P. Pietzonka, A. C. Barato, and U. Seifert, Phys. Rev. E **93**, 052145 (2016).

[47] A. M. Timpanaro, G. Guarnieri, J. Goold, and G. T. Landi, Phys. Rev. Lett. **123**, 090604 (2019).

[48] G. Guarnieri, G. T. Landi, S. R. Clark, and J. Goold, Phys. Rev. Research **1**, 033021 (2019).

[49] Y. Hasegawa, Phys. Rev. Lett. **126**, 010602 (2021).

[50] Y. Hasegawa, Phys. Rev. Lett. **125**, 050601 (2020).

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

[52] M. Perarnau-Llobet, E. Bäumer, K. V. Hovhannisyan, M. Huber, and A. Acin, Phys. Rev. Lett. **118**, 070601 (2017).

[53] S. M. Barnett, J. Jeffers, and D. T. Pegg, Symmetry **13**, 586 (2021).

[54] K. Micadei, G. T. Landi, and E. Lutz, arXiv:2103.14570.