

Coarse-grained second-order response theory

Fenna Müller,^{1,*} Urna Basu,^{2,†} Peter Sollich,^{1,3,‡} and Matthias Krüger^{1,§}¹*Institute for Theoretical Physics, University of Göttingen, Friedrich-Hund-Platz 1, D-37077 Göttingen, Germany*²*Raman Research Institute, Bangalore 560080, India*³*Department of Mathematics, King's College London, Strand, London WC2R 2LS, United Kingdom*

(Received 8 May 2020; accepted 21 September 2020; published 23 October 2020)

While linear response theory, manifested by the fluctuation dissipation theorem, can be applied at any level of coarse-graining, nonlinear response theory is fundamentally of a microscopic nature. For perturbations of equilibrium systems, we develop an exact theoretical framework for analyzing the nonlinear (second-order) response of coarse-grained observables to time-dependent perturbations, using a path-integral formalism. The resulting expressions involve correlations of the observable with coarse-grained path weights. The time-symmetric part of these weights depends on the paths and perturbation protocol in a complex manner; in addition, the absence of Markovianity prevents slicing of the coarse-grained path integral. We show that these difficulties can be overcome and the response function can be expressed in terms of path weights corresponding to a single-step perturbation. This formalism thus leads to an extrapolation scheme where measuring linear responses of coarse-grained variables suffices to determine their second-order response. We illustrate the validity of the formalism with an exactly solvable four-state model and the near-critical Ising model.

DOI: [10.1103/PhysRevResearch.2.043123](https://doi.org/10.1103/PhysRevResearch.2.043123)

I. INTRODUCTION

Many systems of practical and scientific relevance are of an intrinsic stochastic nature with properties dominated by fluctuations, e.g., colloidal particles, protein folding networks, molecular motors, or stochastic heat engines [1]. Such systems lend themselves to descriptions by statistical physics. A variety of approaches exists for this, including the famous Jarzynski equation [2] and Crook's theorem [3], which concern the work done while driving a system far from equilibrium. In contrast, (nonlinear) response theory treats *arbitrary* observables, starting near equilibrium with the fluctuation-dissipation theorem, which relates the *linear* response to equilibrium fluctuations [4,5]. Higher orders in perturbation have also been derived, e.g., for Markov jump processes [6], using path integrals [7], or in terms of correlation functions [8–15]. Nonlinear response theory has also been applied experimentally, enabling measurement of the second-order response from an equilibrium average [16].

The above nonlinear response approaches typically rest on the assumption that all relevant degrees of freedom (d.o.f.) are known and measurable. This is not the case in many

experimental settings, and one is thus faced with the additional challenge of coarse-graining.

Taking the example of a colloidal particle in a simple solvent, the bath d.o.f. can easily be integrated out because they relax quickly on typical colloidal timescales, and they can (thus) be assumed to be in an *equilibrium* state [17,18]. Approaches such as Mori-Zwanzig projection operators formalize this idea by identifying a subset of slow d.o.f. to be relevant and integrating out the fast d.o.f. [19–21]. Indeed, fluctuation relations and response theory have been shown to hold approximately under the assumption that subsystems reach local equilibrium [22–24]. Other types of coarse-graining preserve fluctuations [25] or use other physical or computational restrictions, as, e.g., in polymer physics [26,27] or biophysics [28,29].

Adding a second colloidal particle to our example illustrates the next level of complexity: if the position of one colloidal particle is unknown, experimental estimation of potentials, entropy production, and probability distributions may be incorrect, as shown experimentally in Ref. [30]. This is the case, for example, if a driving protocol acts on the unknown degree of freedom. Such questions in relation to entropy production, work, and other thermodynamic notions in stochastic processes have been analyzed under coarse-graining, both theoretically [23,24,31,32] and experimentally [30,33]. But what about the nonlinear (second-order) response in coarse-grained systems? As detailed below, nonlinear orders remain challenging in coarse-grained systems, even if entropy productions are found correctly. Reference [34] developed second-order response theory in a system coarse-grained to a finite number of states, proposing and verifying an extrapolation scheme for the second-order response from linear contributions. Notably, this approach does not rely on a separation of timescales

*fenna.mueller@theorie.physik.uni-goettingen.de

†urna@rri.res.in

‡peter.sollich@uni-goettingen.de

§matthias.krueger@uni-goettingen.de

as demonstrated explicitly for a model system [34]. While Ref. [34] is restricted to perturbations that remain constant after an initial, instantaneous jump, in this paper we generalize this approach to include arbitrary time dependence.

Starting from microscopic response theory from path integrals, we derive a response theory for a finite number of coarse-grained states. Coarse-graining the path integrals yields coarse-grained path weights, including entropy production, but also the more difficult time-symmetric part of the corresponding weights, from which the second-order response can be found.

These formal expressions can be used in practice, e.g., via an extrapolation scheme. In this scheme, performing a linear-response experiment (or simulation) is sufficient to obtain the second-order response. We show how to measure the second-order response for any protocol from linear perturbations with one step only, thereby greatly facilitating the measurement. This concept is illustrated and verified below in an analytically solvable jump process and in simulations of the two-dimensional (2D) Ising model.

II. SYSTEM AND NONLINEAR RESPONSE THEORY

In this section, we present nonlinear (second-order) response theory, starting from the microscopic description, which is then coarse-grained to macroscopic observables.

A. Microscopic description

Consider a classical system of interacting degrees of freedom, e.g., a fluid, with phase state at time s denoted by $x_s \in \Gamma$, which is in general of high dimensionality. Assuming that the state x_s is of sufficient microscopic resolution, $(x_s)_{s \in [0,t]}$ is a Markov process. In the absence of perturbations, the system is in equilibrium at temperature $T = 1/(k_B\beta)$, with the Boltzmann constant k_B . When perturbed, the system is out of equilibrium, a situation that we aim to analyze here. We therefore start by reviewing an expansion of the system around equilibrium in terms of path integrals [7,35].

We introduce a volume form on the space of paths $p(\omega)\mathbb{D}\omega$, so that $p(\omega)$ is the probability (density) to find the path $\omega = \{x_s\}_{s \in [0,t]}$. The average of a *state observable* $O(x_t)$, which depends on the state of the system at time t , is given by

$$\langle O \rangle(t) = \int O(x_t) p(\omega) \mathbb{D}\omega, \quad (1)$$

On the right-hand side we have an integral over paths ending at time t , weighted by p .

Consider now a perturbation by a potential $v(x)$, acting on the system for times $s \geq 0$, carrying as prefactors a dimensionless perturbation strength ε and a dimensionless *protocol* $h(s)$ of order unity, so that the full perturbation to the energy of state x is given by $-\varepsilon h(s)v(x)$. The aim of response theory, as, for instance, developed in Refs. [7,35], is to express the path probability in the perturbed nonequilibrium system, $p_{\varepsilon,h}(\omega)$, in terms of the equilibrium path probability $p_{\text{eq}}(\omega)$ and orders of the perturbation strength ε . This is done in terms of a Radon-Nikodym derivative, which relates different probability measures in the Radon-Nikodym theorem [35]. Here, it

connects the probability densities [35]

$$p_{\varepsilon,h}(\omega) = e^{-a_{\varepsilon,h}(\omega)} p_{\text{eq}}(\omega), \quad (2)$$

where we have introduced an *action* a quantifying the deviation from equilibrium. It will be useful to consider the time-reversed process, described by backward paths. These are given by $\theta\omega = \pi\{x_{t-s}\}$, where π refers to the kinematic sign reversal, such as flipping the sign of velocities, and they evolve under the reversed protocol $\bar{h}(s) = h(t-s)$ [36]. Integration over the backwards path weight, $p_{\varepsilon,\bar{h}}(\theta\omega)$, yields

$$\begin{aligned} & \int O(x_t) p_{\varepsilon,\bar{h}}(\theta\omega) \mathbb{D}\omega \\ &= \int O(x_0) p_{\varepsilon,\bar{h}}(\omega) \mathbb{D}\omega = \langle O \rangle(0) = \langle O \rangle^{\text{eq}}(0) = \langle O \rangle^{\text{eq}}(t), \end{aligned} \quad (3)$$

where $\langle O \rangle^{\text{eq}}$ denotes the equilibrium average. Equation (3) uses that the system is in equilibrium at time $s = 0$: in the first term in the second line of Eq. (3) we can split the path integral into one over all paths starting at x_0 , followed by an integral over x_0 . The first integration gives the initial distribution of x_0 , and the second one then gives the average of O at $s = 0$, which is the equilibrium average. See Appendix A1 for details. Equation (3) inspires a decomposition of the action $a = d - s/2$ into its time-symmetric and -antisymmetric parts d and s , respectively. The time-antisymmetric part $s_{\varepsilon,h} = \ln(p_{\varepsilon,h}(\omega)/p_{\varepsilon,\bar{h}}(\theta\omega))$ is called the entropy production. For the potential perturbation given above and assuming the time-reversed process also starts in equilibrium, it has the form

$$s_{\varepsilon,h}(\omega) = \varepsilon\beta \left(h(t)v(x_t) - h(0)v(x_0) - \int_0^t \dot{h}_s v(x_s) ds \right), \quad (4)$$

as shown for specific examples in Ref. [7] and quite generally in Ref. [35]. The time-symmetric part $d_{\varepsilon,h} = -\frac{1}{2} \ln(p_{\varepsilon,h}(\omega)p_{\varepsilon,\bar{h}}(\theta\omega)/p_{\text{eq}}^2(\omega))$, sometimes denoted dynamical activity, depends on further details. No explicit form can be given without specifying the dynamics of the system [7].

Expanding in terms of the perturbation strength ε and subtracting the path integral over backwards paths, as considered in Eq. (3), yields

$$\langle O \rangle = \langle O \rangle^{\text{eq}} + \varepsilon \langle s'_h O \rangle^{\text{eq}} - \varepsilon^2 \langle s'_h d'_h O \rangle^{\text{eq}} + O(\varepsilon^3). \quad (5)$$

We introduced the notation $f' = \frac{df}{d\varepsilon}|_{\varepsilon=0}$ so that s' is immediately found from Eq. (4). The derivative d' of the time-symmetric component is given in terms of the derivative of p [35],

$$d'_h(\omega) = -\frac{1}{2p_{\text{eq}}(\omega)} [p'_h(\omega) + p'_{\bar{h}}(\theta\omega)]. \quad (6)$$

Examples for different dynamics may be found in Refs. [7,35].

We finally introduce a notation for the n th-order response of the nonequilibrium average $\langle O \rangle$ [37],

$$\langle O \rangle^{(n)} := \frac{1}{n!} \frac{d^n}{d\varepsilon^n} \langle O \rangle|_{\varepsilon=0}, \quad (7)$$

which we analyze up to $n = 2$ in this paper.

It is important to note that Eq. (5) relies on the unperturbed system being in equilibrium, via detailed balance. For nonequilibrium cases, already the linear response involves

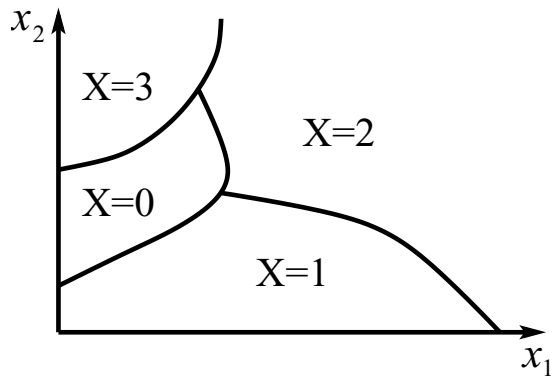


FIG. 1. Illustration of coarse-graining: A continuous microscopic phase space Γ (here spanned by x_1 and x_2) is coarse-grained into states $X = 0, 1, 2, 3$ that make up the coarse-grained space Γ' .

d'_h [38], and the second order involves d''_h . This would lead to differences in the coarse-grained relations derived below.

B. Coarse-grained description

We now turn to a coarse-grained version of the stochastic process introduced above. This is motivated by the fact that experimental resolution is always limited, so that in general only coarse-grained observables can be monitored. Furthermore, developing nonequilibrium thermodynamics for macroscopic variables is an important goal of statistical physics. The coarse-graining as performed here allows for a practical extrapolation scheme, as detailed below.

We thus consider a countable number of coarse-grained, discrete (stochastic) states $X_s \in \Gamma'$, with a function φ uniquely mapping Γ to Γ' , i.e., $X_s = \varphi(x_s)$.

Figure 1 illustrates this mapping of a microscopic continuous state space Γ to $\Gamma' = \{0, 1, 2, 3\}$ consisting of four coarse-grained states. Note that this approach does not rely on a separation of timescales of slow and fast variables and thus remains valid even when the coarse-grained process is not Markovian. The spirit of this coarse-graining is hence distinct from the idea that underlies typical approximations based on local equilibrium.

A crucial *physical* assumption or requirement is that the perturbation potential v acts on the coarse level as well, in other words, v is a function of coarse-grained states. In the introductory example of colloids in a solvent, it means that the perturbation acts only on those colloids whose positions are being monitored. To emphasize this, we introduce a potential $V(X)$ acting on coarse-grained states, so that $v(x) = V(X)$ for all x satisfying $\varphi(x) = X$. Importantly, under this assumption, the entropy production is a functional of coarse-grained paths $\Omega = \{X_s\}_{s \in [0,t]}$ (see Ref. [23] for a statement in a similar spirit). We define the coarse-grained entropy production;

$$\begin{aligned} S'_h(\Omega) &:= \frac{1}{P_{\text{eq}}(\Omega)} \int_{\Omega} s'_h(\omega) p_{\text{eq}}(\omega) \mathbb{D}\omega \\ &= \beta \left(h(t)V(X_t) - h(0)V(X_0) - \int_0^t \dot{h}(s)V(X_s) ds \right). \end{aligned} \quad (8)$$

The index \int_{Ω} indicates that the integral runs over all micropaths belonging to the coarse-grained path Ω , and $P_{\text{eq}}(\Omega)$ is the equilibrium weight for path Ω . Equation (8) follows directly from Eq. (4), by noting that $v(x) = V(X)$ for x with $\varphi(x) = X$. The *linear* response of a state observable $O(X_t)$ is thus

$$\langle O \rangle^{(1)} = \langle S'_h O \rangle^{\text{eq}}. \quad (9)$$

Comparing with Eq. (5), we see that the formalism of linear response is the same when applied to a microscopic or any coarse-grained description, so that the process of coarse-graining is easily performed for the linear response.

To obtain the second-order response, we coarse-grain the second-order response Eq. (5), making use of the coarse-grained entropy production given in Eq. (8),

$$\begin{aligned} \langle O \rangle^{(2)}[h](t) &= - \int s'_h(\omega) d'_h(\omega) O(X_t) p_{\text{eq}}(\omega) \mathbb{D}\omega \\ &= - \int S'_h(\Omega) \left[\int_{\Omega} d'_h(\omega) p_{\text{eq}}(\omega) \mathbb{D}\omega \right] O(X_t) \mathbb{D}\Omega \\ &= - \int S'_h(\Omega) D'_h(\Omega) O(X_t) P_{\text{eq}}(\Omega) \mathbb{D}\Omega. \end{aligned} \quad (10)$$

Here we have identified the coarse-grained D' as the average of d' over micropaths belonging to Ω ,

$$D'_h(\Omega) P_{\text{eq}}(\Omega) := \int_{\Omega} d'_h(\omega) p_{\text{eq}}(\omega) \mathbb{D}\omega, \quad (11)$$

similarly to the coarse-grained entropy production of Eq. (8). In Eq. (10) we split the integration over microscopic paths by first integrating over paths belonging to a given coarse-grained path Ω and subsequently integrating over the latter. We then used that the entropy production takes the same value for all microscopic paths ω belonging to the same Ω , so that s'_h can be taken out of the integral over $\mathbb{D}\omega$ in the second line of Eq. (10), thereby turning into S'_h . This is why the macroscopic parts $S'_h(\Omega)$ and $D'_h(\Omega)$ of the action factorize, with the consequence that Eq. (10) takes a form similar to Eq. (5). The coarse-grained path integral appearing here can be written as

$$\int \mathbb{D}\Omega = \sum_{X_1 \in \Gamma'} \cdots \sum_{X_N \in \Gamma'}, \quad (12)$$

by discretizing time into N lattice points and making use of the discrete nature of X . Other ways of representing $\int \mathbb{D}\Omega$ can be found in Appendix A 2. $D'_h(\Omega)$ in Eq. (11) can be written, using (6), as

$$\begin{aligned} D'_h(\Omega) P_{\text{eq}}(\Omega) &= \int_{\Omega} -\frac{1}{2} [p'_h(\omega) + p'_h(\theta\omega)] \mathbb{D}\omega \\ &= -\frac{1}{2} [P'_h(\Omega) + P'_h(\theta\Omega)], \end{aligned} \quad (13)$$

where we have introduced (derivatives of the) nonequilibrium weight $P'_h(\Omega)$. One important difference between Eqs. (13) and (6) lies in the Markov property of x , which is absent for X : while $p_{\varepsilon,h}(\omega)$ can be cut into (temporal) pieces according to the Chapman-Kolmogorov-Equation, this is not possible for $P_{\varepsilon,h}(\Omega)$.

The main challenge that remains is the determination of $D'(\Omega)$. How $D'(\Omega)$ appears in practice will be analyzed in Sec. III by decomposing the time dependence of the protocol into a number of discrete steps. Section IV verifies and illustrates these findings via analytical solutions of a four-state model, and Sec. V will apply the extrapolation scheme to the Ising model.

III. FROM STEPWISE PERTURBATION TO THE SECOND-ORDER SUSCEPTIBILITY

Equation (10) describes the second-order response $\langle O \rangle^{(2)}$ in terms of a path integral with linear contributions at most. However, evaluating the path integral holds the challenge of finding $D'(\Omega)$ in Eqs. (11) or (13). In this section, we demonstrate that, starting with protocols of a finite number of discrete steps, $D'(\Omega)$ turns into a tensor of finite order. We discuss the simplifications arising if the coarse-grained process is itself Markovian in Appendix A 4.

A. A single-step perturbation

The case of a perturbation with a single step in time, i.e.,

$$h(s) = \Theta(s), \quad (14)$$

was considered in Ref. [34], and for the sake of completeness we summarize the derivation here. The entropy production reads in this case

$$S'(\Omega) = \beta[V(j) - V(i)] = S'_{ij}. \quad (15)$$

We have introduced the abbreviations $i = X_0$ and $j = X_t$ for the states at times 0 and t . The above form of S' reduces the path integral in Eq. (10) into a sum of terms [34]:

$$\begin{aligned} \langle O \rangle^{(2)}(t) &= - \int S'(\Omega) D'(\Omega) O(X_t) P_{\text{eq}}(\Omega) \mathbb{D}\Omega \\ &= - \sum_{ij \in \Gamma'} S'_{ij} \int_{ij} D'(\Omega) O(X_t) P_{\text{eq}}(\Omega) \mathbb{D}\Omega \\ &= - \sum_{ij} S'_{ij} D'_{ij} P_{ij}^{\text{eq}} O(j). \end{aligned} \quad (16)$$

Here, \int_{ij} is a path integral with fixed start and end states i and j , which yields the joint probability

$$P_{ij} = \int_{ij} P(\Omega) \mathbb{D}\Omega. \quad (17)$$

According to Eq. (13), the time-symmetric component is given by

$$D'_{ij} P_{ij}^{\text{eq}} = \int_{ij} D'(\Omega) P_{\text{eq}}(\Omega) \mathbb{D}\Omega \quad (18)$$

$$= -\frac{1}{2}(P'_{ij} + P'_{ji}). \quad (19)$$

For the step perturbation of Eq. (14) the protocol equals its reverse, and the protocol reversal appearing in Eq. (13) is obsolete. We note that, for a single step, $D'(\Omega)$ turns into a matrix D'_{ij} , which is related to the linear response of the coarse-grained probability P_{ij} . The latter can be measured

easily, giving rise to the extrapolation scheme introduced in Ref. [34], which we discuss further in Sec. V below.

B. A two-step perturbation

We add one more step to the protocol at time $0 \leq \tau \leq t$, introducing the corresponding state $k = X_\tau$. Denoting the step sizes by Δh_0 and Δh_1 , the protocol is then

$$h(s) = \Delta h_0 \Theta(s) + \Delta h_1 \Theta(s - \tau). \quad (20)$$

Recalling the definition of S'_{ij} and $S'(\Omega)$ in Eq. (15), Eq. (8) yields the entropy production for two steps,

$$S'(\Omega) = \Delta h_1 S'_{kj} + \Delta h_0 S'_{ij} =: S'_{ikj}. \quad (21)$$

Similarly to Eq. (16), the path integral turns into sums over states at times 0, τ , and t ,

$$\begin{aligned} \langle O \rangle^{(2)}(t) &= - \int S'(\Omega) D'(\Omega) O(X_t) P_{\text{eq}}(\Omega) \mathbb{D}\Omega \\ &= - \sum_{i,k,j \in \Gamma'} S'_{ikj} \int_{ikj} D'(\Omega) O(X_t) P_{\text{eq}}(\Omega) \mathbb{D}\Omega \\ &= - \sum_{ikj} S'_{ikj} D'_{ikj} P_{ikj}^{\text{eq}} O(j). \end{aligned} \quad (22)$$

Consistent with the notation above, the path integral $\int_{ikj} \mathbb{D}\Omega$ is restricted to paths passing through the states i, k, j and yields the probability of being in state i at time $s = 0$, in k at time τ , and in state j at the time of the measurement t :

$$P_{ikj} = \int_{ikj} P(\Omega) \mathbb{D}\Omega. \quad (23)$$

Applying this notation, we can identify the time-symmetric contribution by integrating in Eq. (13)

$$D'_{ikj} P_{ikj}^{\text{eq}} = -\frac{1}{2}(P'_{ikj} + \overline{P'_{ikj}}), \quad (24)$$

where we introduced the probability $\overline{P_{ikj}}$ under time and protocol reversal. More specifically, $\overline{P_{ikj}}$ is the probability to measure j at time $s = 0$, k at time $s = t - \tau$, and i at time $s = t$ under the backwards protocol $\overline{h}(s) = (\Delta h_0 + \Delta h_1) \Theta(s) - \Delta h_1 \Theta(s - (t - \tau))$. To arrive at \overline{P} we have swapped the two operations of time reversal and integration over coarse-grained paths.

By construction, D' must be linear in the protocol h , so that it can be decomposed into

$$D'_{ikj} = \Delta h_0 D'_{ikj}[\Theta_0] + \Delta h_1 D'_{ikj}[\Theta_\tau]. \quad (25)$$

The quantities $D'_{ikj}[\Theta_s]$ appearing on the right give the value of D'_{ikj} for a single perturbation step at time s , e.g., $D'_{ikj}[\Theta_0]$ is extracted for a perturbation consisting of a step at time $s = 0$. We have thus obtained a tensor D' with three indices, which as before is connected to coarse-grained probabilities P [39].

C. Second-order susceptibility for any protocol

In this section, we consider the response to a general protocol h by deriving a formula for the second-order susceptibility in terms of “one-step probabilities.” Given that the unperturbed equilibrium systems is invariant under time

translations, we can express the second-order response for a protocol h in terms of the second-order susceptibility χ [40],

$$\langle O \rangle^{(2)}(t) = \int_0^t \int_0^t \dot{h}(t_1) \dot{h}(t_2) \chi(t - t_1, t - t_2) dt_1 dt_2. \quad (26)$$

The $\chi(t, t - \tau)$ defined in this way can be determined from the second-order response to a protocol with one and two steps, since the time derivatives of such a protocol are δ distributions at the times where the field jumps. We may thus find χ from the relations given in the previous subsections. Comparing the definition of the second-order susceptibility, Eq. (26), to the response formula given in terms of indices (22), and using the linearity of D' in Eq. (25) yields

$$\chi(t, t - \tau) = -\frac{1}{2} \sum_{ikj} (S'_{kj} D'_{ikj}[\Theta_0] + S'_{ij} D'_{ikj}[\Theta_\tau]) P_{ikj}^{\text{eq}} O(j), \quad (27)$$

which is valid for $0 \leq \tau \leq t$. The case of arbitrary times $\chi(t_1, t_2)$ is obtained by inserting the respective arguments $\tau = t_1 - t_2$ and $t = t_1$.

Equation (27) is an intermediate result: it gives the second-order response function χ for any arguments in terms of sums over indices of the tensors S'_{ij} and D'_{ikj} obtained in Eqs. (A4), (A5), and (15). Note that as (27) relates to different times, it comes from the cross-terms in the product of (21) and (25). This is why, e.g., the first term contains a factor S'_{kj} relating to the state k at time τ , combined with $D'_{ikj}[\Theta_0]$ for a perturbation at time zero.

As a final simplification, we note that the entropy production in Eq. (27) carries only two indices so that we can sum over the remaining index. This sum eliminates one index from the expressions $D'_{ikj} P_{ikj}^{\text{eq}}$, which always occur jointly [compare Eqs. (A4) and (A5)]. One term is given by summing over the center index k in the probabilities $\sum_k P'_{ikj}(\tau, t) = P'_{ij}(t)$. We add notation to include time and protocol, as these are varied below. For example, $P_{ikj}[\Theta_{s_1}](\tau, t)$ denotes the probability of measuring state i at time $s = 0$, state k at time τ , and state j at time t under a perturbation switched on at time s_1 . With this notation in mind, the summation over k yields

$$\begin{aligned} & \sum_k D'_{ikj}[\Theta_\tau](\tau, t) P_{ikj}^{\text{eq}}(\tau, t) \\ &= -\frac{1}{2} [P'_{ij}[\Theta_\tau](t) + P'_{ji}[\Theta_0](t) - P'_{ji}[\Theta_{t-\tau}](t)] \\ &= D'_{ij}[\Theta_\tau](t) P_{ij}^{\text{eq}}(t), \end{aligned} \quad (28)$$

where we used the linearity of $P'_{ji}[\Theta_0 - \Theta_{t-\tau}](t) = \overline{P'_{ij}[\Theta_\tau](t)}$ in the backwards protocol, as given in Eq. (A5). The second term involves summation over the first index,

$$\begin{aligned} & \sum_i D'_{ikj}[\Theta_0](\tau, t) P_{ikj}^{\text{eq}}(\tau, t) \\ &= -\frac{1}{2} (P'_{kj}[\Theta_{-\tau}](t - \tau) + P'_{jk}[\Theta_0](t - \tau)) \\ &= D'_{kj}[\Theta_{-\tau}](t - \tau) P_{kj}^{\text{eq}}(t - \tau). \end{aligned} \quad (29)$$

Notably, the perturbation in Eq. (29) starts at negative times $-\tau < 0$ and probabilities cover a time interval of $t - \tau$ between measurements due to integrating out the first state.

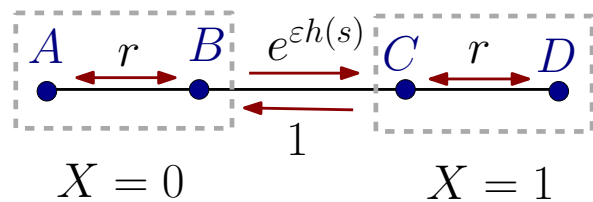


FIG. 2. Illustration of coarse-graining the Markov four-state model into a non-Markovian two-state model. Arrows denote the transitions with the given rates.

After renaming indices, we finally obtain for the second-order susceptibility

$$\begin{aligned} \chi(t, t - \tau) &= -\frac{1}{2} \sum_{ij} S'_{ij} O(j) (D'_{ij}[\Theta_{-\tau}](t - \tau) P_{ij}^{\text{eq}}(t - \tau) \\ &\quad + D'_{ij}[\Theta_\tau](t) P_{ij}^{\text{eq}}(t)). \end{aligned} \quad (30)$$

This expression is symmetrical under exchange of its arguments $t_1 = t$ and $t_2 = t - \tau$, as expected from the definition of the susceptibility, Eq. (26). For the single-step perturbation the second order is given by the response function for equal-time arguments $\langle O \rangle^{(2)}[\Theta_0] = \chi(t, t)$, i.e., by setting $\tau = 0$. The two terms in Eq. (22) then become the same and we recover Eq. (16) as it should be.

IV. ILLUSTRATION AND VERIFICATION: THE FOUR-STATE MODEL

In this section, we use a simple example system, namely a (driven) four-state model that can be solved analytically, to verify and illustrate the concepts introduced in Sec. III.

A. Model and coarse-graining

The second-order response can be expressed in terms of the entropy production and a time-symmetric component, as in Eqs. (10) and (30). As a proof of concept, we consider a Markov jump process with four states $\Gamma = \{A, B, C, D\}$; see Fig. 2, which is then coarse-grained to a two-state one. Such Markov jump processes may be used to describe a variety of systems; see, e.g., Ref. [35].

The conditional probabilities $p_{\alpha\delta}(s, t)$ for occupying state δ at time t if occupying α at an earlier time s are described by the Master equation (an equivalent equation holds for occupation densities)

$$\frac{\partial}{\partial t} p_{\alpha\delta}(s, t) = \sum_{\gamma \in \Gamma} p_{\alpha\gamma}(s, t) q_{\gamma\delta}(t) \quad (31)$$

with the rate $q_{\alpha\delta}(t)$ for the transition from the state α to δ and setting $q_{\alpha\alpha} = -\sum_{\delta \neq \alpha} q_{\alpha\delta}$ to ensure probability conservation; cf. Ref. [35].

More explicitly, we use time-independent rates $q_{AB} = q_{BA} = q_{CD} = q_{DC} = r$ for the side links, and the center links have rates $q_{BC}(t) = e^{\epsilon h(t)}$ and $q_{CB} = 1$ (with all other rates being 0). Only the rate of the center link $q_{BC}(t)$ is time-dependent, and this will be used to drive the system. The rate matrix $q(s)$ is given explicitly in Appendix B, Eq. (B1). We use $\beta = 1$ in this section.

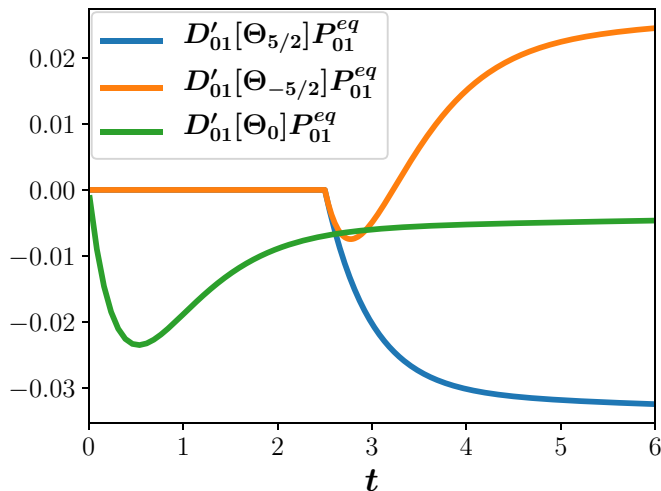


FIG. 3. The time-symmetric components D'_{01} contributing to the second-order susceptibility $\chi(t, t - \tau)$ and $\chi(t, t)$ in the four-state model, Eq. (32), with rate $r = 0.1$ and $\tau = 5/2$. Time arguments are chosen as they appear in Eq. (32). Due to causality, the time-symmetric component vanishes for negative time arguments, $D'_{01}[\Theta_\tau](t) = D'_{01}[\Theta_{-\tau}](t - \tau) = 0$ for $t < \tau$.

This system is *coarse-grained* into two states $X = 0$ and 1 by assigning $\varphi(A) = \varphi(B) = 0$ and $\varphi(C) = \varphi(D) = 1$. The two coarse-grained states are connected by the center link BC and the associated rates q_{BC} and q_{CB} of the underlying Markov process, which yields a non-Markovian two-state process. Notably, in the limit $r \gg 1$ the resulting two-state process is Markovian, while it is strongly non-Markovian in the opposite

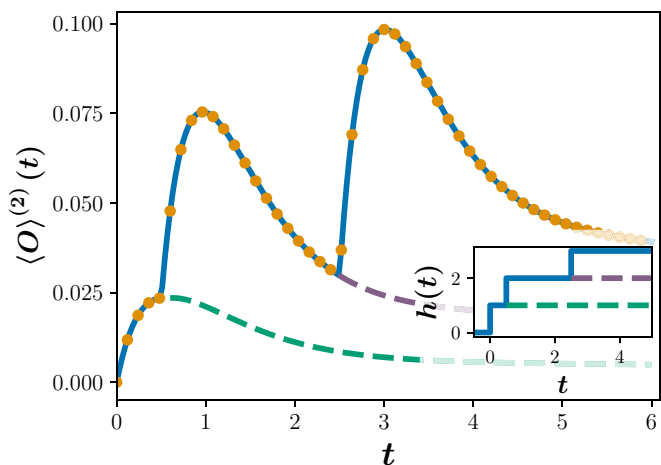


FIG. 4. The second-order response $\langle O \rangle^{(2)}$ of the coarse-grained four-state model, for a driving protocol containing one, two, and three instantaneous steps (protocols shown in the inset), as a function of time t . Lines show the analytical solution (see Appendix B 1), and data points are found from Eq. (27) as well as analytically. The “internal” rate is set to $r = 0.1$, which corresponds to a strongly non-Markovian case, as also reflected in the curves: The response shows two distinct timescales of relaxation. States, rates, and time are naturally dimensionless. As $t \rightarrow \infty$, the curves approach zero, as, for symmetry reasons, the system has no static second-order response.

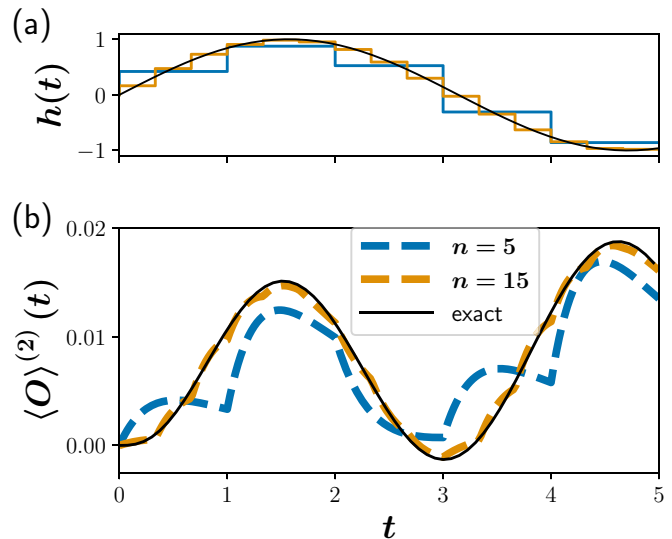


FIG. 5. The second-order response $\langle O \rangle^{(2)}$ of the coarse-grained four-state model as a function of time t , for a sinusoidal protocol of different temporal resolution, as shown in the upper part: The sine (black) is approximated by step functions with n steps of height $\Delta h_k = [\sin((k+1)\delta t) - \sin((k-1)\delta t)]/2$ and time increments $\delta t = t_{\max}/n$, here for $t_{\max} = 5$ and for $n = 5$ (blue) and $n = 15$ steps (orange). The lower panel shows the corresponding second-order response in the same colors as the corresponding protocols. Rates and time are given in dimensionless units.

limit, $r \ll 1$. Choosing $r = 0.1$ as in Figs. 4 and 5 results in the system being in the latter regime.

The associated potential is $v(A) = v(B) = 0$ and $v(C) = v(D) = 1$, hence $V(0) = 0$ and $V(1) = 1$. The rates thus fulfill $\frac{q_{BC}(t)}{q_{CB}} = e^{\epsilon h(t)[v(C) - v(B)]}$, which is called the microscopic reversibility condition [3] or local detailed balance [35]. This is a sufficient condition to have an entropy production of the form given by Eq. (4) [35]. For single-step perturbations, the Master equation can be solved analytically. For more complex protocols, the solution is formally given by a time-ordered exponential, which may be expanded in orders of ϵ using a Dyson-expansion. This allows us to illustrate our approach of computing the second-order susceptibility from linear quantities analytically.

B. One, two, and three steps

We compute the second-order susceptibility $\chi(t, t - \tau)$ from linear contributions S'_{ij} and $D'_{ij}P_{ij}^{eq}$ for perturbations switched on at different times $\pm\tau$, according to Eq. (30). Equation (30) for the average in the coarse-grained two-state system with $O(j) = j$ reduces to

$$\chi(t, t - \tau) = -\frac{1}{2} \left(D'_{01}[\Theta_{-\tau}](t - \tau) P_{01}^{eq}(t - \tau) + D'_{01}[\Theta_\tau](t) P_{01}^{eq}(t) \right). \quad (32)$$

The entropy production that contributes is $S'_{01} = 1$ and the relevant time-symmetric components D'_{01} for different perturbations are shown in Fig. 3. The explicit form of the second-order susceptibility in the four-state model is given in Appendix B, Eq. (B2). Employing this function together with

Eq. (26) enables prediction of the second-order response for arbitrary protocols $\langle O \rangle^{(2)}$.

Here, we demonstrate this by means of a protocol $h = \Theta_0 + \Theta_{1/2} + \Theta_{5/2}$ with three steps. As shown in Fig. 4, the response formula for the second order $\langle O \rangle^{(2)}$ coincides with the explicit solution. This is an example of employing the second-order susceptibility from linear contributions to correctly predict the second-order response under a protocol with several steps.

C. Continuous protocol: Exact and discretized

As noted above, Eq. (26) readily describes any protocol, which we further illustrate by using a protocol of a sinusoidal oscillation of the form

$$h(s) = \sin(s)\Theta_0(s). \quad (33)$$

The resulting response is shown in Fig. 5. In addition to the response corresponding to the protocol of Eq. (33), we show the response to the discretized versions of the protocol in the upper part of Fig. 5. This illustrates the possibility of an additional coarse-graining along the time axis of a protocol, which is one natural way of implementing Eq. (30) in practice (see also Sec. V below). How fine a discretization is needed? As seen in the graph, discretizing with an increment of unity (resulting in $n = 5$ steps in the given time range) yields pronounced deviations from the exact result. On the other hand, discretizing with an increment of $1/3$, resulting in $n = 15$ steps, produces a more precise approximation. This can be understood from the curves in Fig. 4, where the (shortest) relaxation time, or the response time, is of the order of unity. This analysis suggests that the time increment should be small compared to that response time to accurately resolve the perturbation protocol.

V. EXTRAPOLATION: ISING MODEL

In this section, we illustrate the validity of the time-dependent coarse-grained response theory for an interacting system with many degrees of freedom, using the example of a near-critical Ising model. Let us consider a 2D lattice of size $L \times L$ with periodic boundaries; each lattice site i contains a spin $\eta_i = \pm 1$, which interacts with its nearest-neighbor spins. Let the coarse-grained variable X correspond to a single site, say site k , so that we have $X = \frac{1}{2}(1 + \eta_k)$. In other words, all spins except for spin k will be coarse-grained away, and play the role of a complex (non-Markovian) bath. This scenario may mimic the experimental situation in which a system is perturbed and monitored at a local position in space. We thus introduce a magnetic field, which acts on the spin k , i.e., a potential $V(X) = \eta_k = 2X - 1$. The Hamiltonian describing the system at any time s is (setting the spin coupling to unity)

$$\mathcal{H} = - \sum_{\langle ij \rangle} \eta_i \eta_j - [g + \varepsilon h(s)] \eta_k. \quad (34)$$

The explicit time dependence assigned to the magnetic field via the protocol $h(s)$ gives rise to a perturbation of the system from its equilibrium state.

In the absence of the magnetic field, i.e., with $g = h(s) = 0$, the system shows a paramagnetic to ferromagnetic

transition at temperature $T_c = 2.269$ in the limit of thermodynamically large size L (having set the Boltzmann constant to unity). Here we consider a system of size $L = 16$ at a slightly supercritical temperature $T = 2.45$. This finite-sized system shows a nonzero magnetization at this temperature, which randomly flips its sign on a slow timescale. We thus expect the resulting bath for the tagged spin η_k to be highly non-Markovian.

For the sake of simplicity, we take the two-step protocol introduced in Sec. III B [see Eq. (20)] with $\Delta h_0 = \Delta h_1 = 1$ and consider the response of the observable $O(X) = X$. We also use a time-independent offset magnetic field of strength $g = 2.0$, which renders the equilibrium system nonsymmetric, yielding a finite second-order response.

Unlike the four-state model, the susceptibility and response function cannot be calculated analytically here, and we take recourse to Monte Carlo simulations. To be specific, we use Glauber dynamics, where a randomly selected spin flips with rate $\min\{1, e^{-\beta\Delta H}\}$, ΔH being the change in energy due to the proposed flip and $\beta = T^{-1}$ is the inverse temperature of the system. One Monte Carlo step consists of L^2 attempted flips, which defines the unit of time.

To demonstrate the validity of the response formalism, we compare the response $\langle O \rangle^{(2)}(t)$ predicted by Eq. (26) with $\langle O \rangle_{\text{per}}^{(2)}(t)$, obtained from directly applying a larger perturbation. The latter is extracted accurately from

$$\langle O \rangle_{\text{per}}^{(2)}(t) = \frac{1}{2\varepsilon^2} [\langle X \rangle^\varepsilon(t) + \langle X \rangle^{-\varepsilon}(t) - 2\langle X \rangle^{\text{eq}}], \quad (35)$$

where $\langle \cdot \rangle^\varepsilon$ denotes the expectation value in the presence of the perturbation protocol of Eq. (20) with strength ε , and $\langle X \rangle^{\text{eq}}$ is the expected value in equilibrium. We use measurements with strengths $\pm\varepsilon$ to avoid errors of $O(\varepsilon^3)$ [34].

On the other hand, the response theory predicts the second-order susceptibility via Eq. (26). For the protocol of Eq. (20) with $\Delta h_0 = \Delta h_1 = 1$ it reduces to

$$\langle O \rangle^{(2)}(t) = \chi(t, t) + 2\chi(t, t - \tau) + \chi(t - \tau, t - \tau), \quad (36)$$

where $\chi(t_1, t_2)$ is given by Eq. (30). As mentioned before, for $O(X) = X$, the sum reduces to only one term, namely $i = 0, j = 1$. Moreover, in this case, $S'_{01} = \beta[V(1) - V(0)] = 2\beta$, and we only need to measure the linear parts of D_{ij} under single-step perturbations at times 0 and $\pm\tau$.

Using the Monte Carlo simulations and applying a (small) perturbation of strength $\varepsilon = \pm 0.05$, we measure the *linear responses* of the relevant path probabilities $P_{ij}[h](t)$. The corresponding matrices D' are then computed using Eq. (B4) in Appendix B 2. Figure 6 shows plots of the resulting D'_{01} , evaluated for the three different protocols as needed, in analogy to Fig. 3. Qualitative differences to Fig. 3 result from the fact that here, a finite second-order response remains in the long-time limit. The presence of a slow timescale is visible in the slow relaxation of the curves in Fig. 6.

For the particular quantity $P_{ji}[\theta_{t-\tau}]$, the numerical fluctuations are substantial and we therefore obtain the derivative by fitting the $P_{ji}^{\pm\varepsilon}[\theta_{t-\tau}]$ to a compressed exponential form and taking the difference of these fitted functions; see Appendix B 2 for more details. The dark blue curve shows the $D'_{01}[\theta_\tau]$ obtained using this fit; the light gray curve shows the original data.

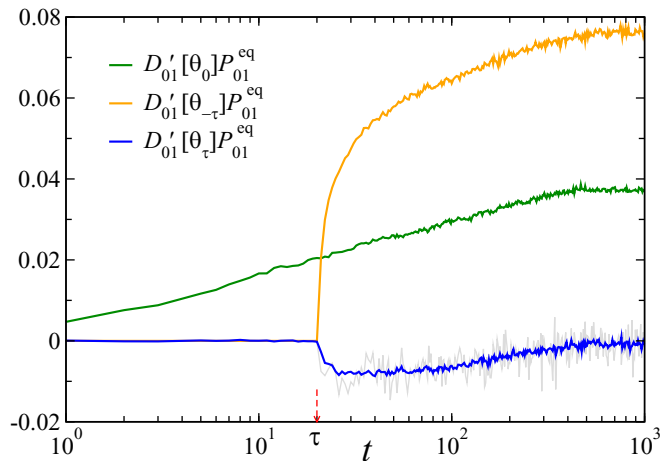


FIG. 6. The relevant time-symmetric components D'_{01} contributing to the second-order response for the Ising model measured from numerical simulations. Here we have considered a fixed $\tau = 20$, and the linear response is calculated using $\varepsilon = 0.05$. The blue curve is obtained using the best fit for $P_{10}^{\varepsilon}[\theta_{-\tau}]$, while the gray curve shows the original data (see the main text and Appendix B 2 for details).

The second-order response is obtained using Eq. (36) along with Eq. (30). Figure 7 compares the susceptibility $\langle O \rangle_{\text{per}}^{(2)}(t)$ that we measure directly (symbols) with the predicted response $\langle O \rangle^{(2)}(t)$ (solid lines) for two different values of τ . At late times $t \rightarrow \infty$, the susceptibility reaches a stationary value that is independent of τ and is simply the equilibrium second-order response for a perturbation $\varepsilon(\Delta h_0 + \Delta h_1)V(X) = 2\varepsilon(2X - 1)$. This can be calculated by a series expansion of the Boltzmann weight and turns out to be $8\beta^2 \langle X \rangle^{\text{eq}}(1 - 2\langle X \rangle^{\text{eq}})(1 - \langle X \rangle^{\text{eq}})$ as shown in detail in Appendix B 3. This value is indicated by a black dashed line in the figure.

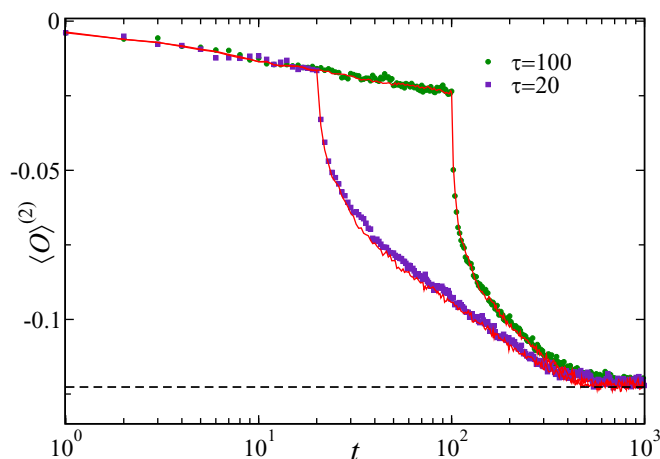


FIG. 7. The second-order response of a single spin in the Ising model. The solid red lines show the response $\langle O \rangle^{(2)}$ predicted from Eq. (36) while the symbols show the susceptibility $\langle O \rangle_{\text{per}}^{(2)}$ measured directly from simulations. The dashed line indicates the *static* response, found by expanding the Boltzmann weight; see the main text.

It is worth reemphasizing here that the procedure used in this section generalizes the extrapolation scheme introduced in Ref. [34] to arbitrary time-dependent perturbations: the second-order response, which is relevant for a comparatively stronger perturbation, can be predicted exactly by measuring path probabilities close to equilibrium, i.e., within the linear response regime.

VI. CONCLUSIONS

We have developed a second-order response theory for coarse-grained observables, which is valid for arbitrary time-dependent perturbation protocols and thus provides a significant extension of Ref. [34]. One application of this theory is an extrapolation scheme that uses measurements within the linear regime to predict the second-order response. The relevant linear measurements only need to be performed for simple perturbation protocols consisting of a single switch-on event, and from these the second-order response for arbitrary protocols follows.

The necessary spatial resolution, i.e., the degree of coarse-graining possible in this approach, is set by the perturbation. Returning to the introductory example of two colloidal particles, if the perturbation acts only on one of the two colloids, the other one can be coarse-grained, i.e., its position does not have to be monitored. An important difference to approaches that use fast and slow variables is thus that, in the present scheme, the coarse-grained variables are allowed to be of non-Markovian type.

The scheme can be applied to any *time* dependence of protocol. As is the case for spatial resolution, it is the protocol that sets the (experimental) time resolution required to apply our method. However, as found in explicit examples, a temporal resolution that is fine compared to the reaction time of the coarse-grained variables is also sufficient.

Technically, our approach relies on being able to resolve fully the entropy production. This is ensured by the assumption that the perturbing potential depends only on macrovariables, and it implies that the entropy production and the time-symmetric part of the action decouple when coarse-graining. This work is thus naturally in agreement with (macroscopic and stochastic) thermodynamics and with the known fluctuation relations. Its new contribution lies in the description of the nonthermodynamic, symmetric part of the action.

Future work will consider higher orders of perturbation, as well as simultaneous perturbations by multiple potentials with different time dependences, and will investigate possibilities of combining this scheme with approaches that rely on separation of fast and slow timescales. It may also be insightful to combine our approach with estimates of the entropy production for cases in which the potential acts on partly inaccessible d.o.f. [41]. Another important extension to be addressed is perturbations of nonequilibrium systems, where already the linear order is difficult as regards coarse-graining.

Finally, one could explore the question of a Gallavotti-Cohen symmetry for coarse-grained descriptions [3,42]: from the definition of $s_{\varepsilon,h}(\omega)$ and Eq. (8) we see that $\langle e^{-s_{\varepsilon,h}(\omega)} \rangle = \langle e^{-\varepsilon S_h^{\varepsilon}(\Omega)} \rangle = 1$. Because the entropy production $\varepsilon S_h^{\varepsilon}(\Omega)$ depends only on the coarse-grained paths, this implies that in

our setting the coarse-grained dynamics does itself obey the Gallavotti-Cohen symmetry. It will then be interesting to see whether statements about nonlinear response can be deduced from this, by extending existing results for the linear response regime [3,42].

ACKNOWLEDGMENTS

U.B. acknowledges support from Science and Engineering Research Board (SERB), India under Ramanujan Fellowship (Grant No. SB/S2/RJN-077/2018). M.K. acknowledges support from DFG Grant No. KR 3844/3-1.

APPENDIX A: RELATIONS AND DEFINITIONS

1. Microscopic response formalism

The results presented here are based on the known microscopic response formalism presented in Sec. II. In this Appendix, we aim to describe in detail how time-dependent perturbations can be treated, as these details may not be provided in other literature. The key is the reversal of paths and protocol in the integration. Mathematically, the path weight considered here is a joint one. This is in contrast to some common literature on the subject of path integrals, which considers the probability $p(\omega|x_i)$ to find a path ω given that the system starts in a fixed state x_i . This probability is related to the full path probability by the probability $\rho_0(x_i)$ of the initial state at time $s = 0$, namely $p(\omega) = p(\omega|x_i)\rho_0(x_i)$. In fact, one may think of this path integral as three integrals: one Lebesgue integral over the starting point x_0 , one over the end point x_t , and a third real path integral over possible paths connecting these two points. It is then immediately clear what happens when integrating over reversed paths in the average of a state observable as given by Eq. (1): We can integrate out paths starting in a fixed state $x_0 = x_i$ at which the observable $O(x_0)$ is then evaluated. Since every path starting in x_i has to go somewhere, the integral over all these paths reduces to the probability density $\rho_0(x_i)$ of x_i ,

$$\int_{x_0=x_i} p_{\varepsilon,g}(\omega) \mathbb{D}\omega = \rho_0(x_i). \quad (\text{A1})$$

Note that the same integral, taken over the *conditional* path weight, would yield one. These arguments hold for *arbitrary protocols*, and in particular for the reversed protocol $\bar{h}(s) = h(t-s)$.

2. Coarse-grained path integral

Let us now consider the *coarse-grained path integral* as used in Eq. (10) in more detail. Instead of expressing the coarse-grained path integral by summing over possible states, see Eq. (12), it can also be written by integration over possible paths as follows.

For processes with discrete state space, paths are given by a sequence of states and jump times $\Omega = (X_0, t_0; X_1, t_1, \dots, X_n, t_n)$, i.e., $X_s = X_k$ for all $s \in [t_k, t_{k+1})$. We always have $t_0 = 0$ and we set $t_{n+1} = t$ for notational simplicity. A path integral is then given by summing over

possible states and integrating over jump times t_i ,

$$\int f(\Omega) \mathbb{D}\Omega = \sum_{n=0}^{\infty} \sum_{X_0 \in \Gamma'} \sum_{X_1 \in \Gamma' \setminus \{X_0\}} \cdots \sum_{X_n \in \Gamma' \setminus \{X_{n-1}\}} \times \int_0^t \int_{t_1}^{t_1} \cdots \int_{t_n}^{t_n} f(\Omega) \prod_{i=0}^{n-1} dt_{i+1}. \quad (\text{A2})$$

Analogously one may write down the exact same equation for Markov jump processes, as done in more detail in Ref. [35].

3. Derivation of Eq. (30)

For clarity, we provide a few more steps for the derivation of the second-order susceptibility in terms of sums, Eqs. (27) and (30). We insert the linear decomposition Eq. (25) of the dynamical activity D'_{ikj} into the second-order response $\langle O \rangle^{(2)}$ expressed as a sum over states, Eq. (22). For a protocol with two steps, as defined in Sec. III B, separating the different combinations of Δh_0 and Δh_1 yields

$$\begin{aligned} \langle O \rangle^{(2)} = & - \sum_{ikj} (\Delta h_0^2 S'_{ij} D'_{ikj}[\Theta_0] + \Delta h_1^2 S'_{kj} D'_{ikj}[\Theta_\tau] \\ & + \Delta h_0 \Delta h_1 S'_{kj} D'_{ikj}[\Theta_0] \\ & + \Delta h_0 \Delta h_1 S'_{ij} D'_{ikj}[\Theta_\tau]) P_{ikj}^{\text{eq}} O(j). \end{aligned} \quad (\text{A3})$$

Adding the explicit time arguments, the two versions of D' appearing in Eqs. (25) and (A3) read

$$\begin{aligned} D'_{ikj}[\Theta_0](\tau, t) P_{ikj}^{\text{eq}}(\tau, t) \\ = & -\frac{1}{2} (P'_{ikj}[\Theta_0](\tau, t) + P'_{jki}[\Theta_0](t - \tau, t)) \quad \text{and} \quad (\text{A4}) \\ D'_{ikj}[\Theta_\tau](\tau, t) P_{ikj}^{\text{eq}}(\tau, t) \\ = & -\frac{1}{2} (P'_{ikj}[\Theta_\tau](\tau, t) + P'_{jki}[\Theta_0](t - \tau, t) \\ & - P'_{jki}[\Theta_{t-\tau}](t - \tau, t)). \end{aligned} \quad (\text{A5})$$

In Eq. (A5) we have expanded the probability under the backward protocol $\bar{\Theta}_\tau = \Theta_0 - \Theta_{t-\tau}$ by thinking of it as a two-step perturbation and using the general linearity property

$$P'_{ikj}[h^{(2)}] = \Delta h_0 P'_{ikj}[\Theta_0](\tau, t) + \Delta h_1 P'_{ikj}[\Theta_\tau](\tau, t). \quad (\text{A6})$$

The above equations are the basis for integrating out one index, specifically the initial state i in Eq. (A4) and the state k at time τ in Eq. (A5). Summing over possible states in a joint probability yields $\sum_j \mathcal{P}(X_{t_1} = i, X_{t_2} = j, X_{t_3} = k) = \mathcal{P}(X_{t_1} = i, X_{t_3} = k)$, thus yielding probabilities P_{ij} for different protocols; see Eqs. (28) and (29), respectively.

4. Markov case

The results derived in Secs. II B and III do not rely on a Markovian property of the coarse-grained variables. There might be practical cases, however, where the degrees of freedom under consideration *are* in fact Markovian, for example if a local equilibrium approximation is justified for the d.o.f. that are integrated out. In that case, $(X_s)_{s \in [0, t]}$ is a Markov process, and hence follows the formulas of the microscopic response formalism in Sec. II A (see Refs. [7,13] for specifics). Notably,

the linear contribution of the time-symmetric part is given as a superposition of instantaneous values, denoted $\tilde{d}(x_s)$,

$$d'_h(\omega) = \int_0^t h(s)\tilde{d}(x_s)ds \quad (\text{A7})$$

as explained in Ref. [13]. We can thus decompose

$$D'_{ikj}[\Delta h_0 \Theta_0 + \Delta h_1 \Theta_\tau] = \Delta h_0 D'_{ik}(\tau) + \Delta h_1 D'_{kj}(t - \tau) \quad (\text{A8})$$

and for the probability

$$P_{ikj} = P_{ik}(\tau)p_{kj}(\tau, t) \quad (\text{A9})$$

with the conditional probability p_{kj} introduced before Eq. (31). Equation (A9) is in contrast to the case of non-Markovian processes where states at different times couple due to memory effects. With only the quantity $D'_{ik}[\Theta_0]$ appearing in Eq. (A4) (evaluated at different times), the second-order response in the Markov case thus takes the complexity of the single-step protocol described in Sec. III A. This simplifies the extrapolation scheme introduced in Sec. V, as only $P'_{ij}[\Theta_0]$ and P_{ij}^{eq} need to be measured in order to find the second-order response for any protocol.

APPENDIX B: MODELS AND MEASUREMENT

1. The four-state model

The rate matrix for the four-state model analyzed in Sec. IV is given by

$$q(s) = \begin{pmatrix} -r & r & 0 & 0 \\ r & -e^{\varepsilon h(s)} - r & e^{\varepsilon h(s)} & 0 \\ 0 & 1 & -1 - r & r \\ 0 & 0 & r & -r \end{pmatrix} \quad (\text{B1})$$

so that the row sums are 0 and as explained r is a dimensionless parameter. For $r \ll 1$ this system exhibits much slower rates within the macrostates than in the transition connecting the two coarse-grained states, which is of order 1. Still our extrapolation technique succeeds (Fig. 4). This illustrates that our method does not rely on separation of timescales as also demonstrated in Ref. [34]. For the average of the coarse-grained observable $O(X) = X$ the second-order susceptibility is computed from Eq. (32).

Using the abbreviations $\gamma = \sqrt{1 + r^2}$ and $\gamma_\pm = 1 + r \pm \gamma$ the exact second-order susceptibility reads

$$\begin{aligned} \chi(t_1, t_2) = & \frac{1}{64\gamma^3} (r(e^{2t_2\gamma} - 1)(e^{-t_1\gamma_+} + 2e^{-t_2\gamma_+} + e^{-2t_2\gamma - t_1\gamma_-}) \\ & + 2t_2\gamma[(1 + \gamma)(\gamma - r)(e^{-t_1\gamma_+} + e^{-t_2\gamma_+}) \\ & + (\gamma - 1)(\gamma + r)(e^{-t_2\gamma_-} + e^{-t_1\gamma_-})]g). \end{aligned} \quad (\text{B2})$$

2. Measurement in the Ising model

To compute the first derivatives of the path weight $P_{ij}[h](t)$ accurately, we use two measurements, namely with perturbation strengths $\pm\varepsilon$. Expanding $P_{ij}^{\pm\varepsilon}[h](t)$ in a Taylor series around equilibrium, i.e., around $\varepsilon = 0$, we get

$$P_{ij}^{\pm\varepsilon}[h](t) = P_{ij}^{\text{eq}}(t) \pm \varepsilon P'_{ij}[h](t) + \frac{\varepsilon^2}{2} P''_{ij}[h](t) + O(\varepsilon^3).$$

TABLE I. Numerical values of the fitting parameters used for $P_{10}^{\pm\varepsilon}[\theta_{t-\tau}](t)$.

	ε	a	b	c
$\tau = 20$	0.05	0.0238	0.0296	0.7752
	-0.05	0.0253	0.0327	0.7575
$\tau = 100$	0.05	0.0096	0.0068	1.0227
	-0.05	0.0095	0.0045	1.0968

The first derivative of the path probabilities can then be extracted from

$$P'_{ij}[h](t) = \frac{1}{2\varepsilon} [P_{ij}^\varepsilon[h](t) - P_{ij}^{-\varepsilon}[h](t)] + O(\varepsilon^2), \quad (\text{B3})$$

and similarly for $P_{ij}[\bar{h}](t)$. Note that the error here is one order smaller than if the derivative was computed only from $P_{ij}^\varepsilon[h](t)$ and $P_{ij}^{\text{eq}}(t)$. Using the above equation and its counterpart for $P_{ij}[\bar{h}](t)$ we can also extract D'_{ij} ,

$$\begin{aligned} D'_{ij}[h]P_{ij}^{\text{eq}} = & -\frac{1}{4\varepsilon} (P_{ij}^\varepsilon[h] + P_{ji}^\varepsilon[\bar{h}] - P_{ij}^{-\varepsilon}[h] - P_{ji}^{-\varepsilon}[\bar{h}]) \\ & + O(\varepsilon^2). \end{aligned} \quad (\text{B4})$$

As mentioned in the main text, for the particular case of $P_{ij}[\theta_{t-\tau}](t)$, instead of calculating the derivative directly from the numerically measured path probabilities, we use a functional fit. We first fit $P_{10}^\varepsilon[\theta_{t-\tau}](t) - P_{10}^\varepsilon[\theta_{t-\tau}](\tau)$ to a functional form $a(1 - \exp[-b(t - \tau)^c])$ (remember that the path probability difference is zero for $t < \tau$ in this case) with a, b, c as fitting parameters. The derivative is then calculated using Eq. (B3) along with these fitted functions. For the sake of completeness, we provide the values of the fitting parameters in Table I.

3. Static response in the Ising model

The long-time limiting value of the second-order response in the Ising model can be computed from the equilibrium Boltzmann distribution. Under the perturbation protocol (20), in the long-time limit, the system reaches an equilibrium state characterized by configuration weights

$$P(\{\eta_i\}) = \frac{1}{Z_\varepsilon} e^{-\beta[\mathcal{H}_0 - \varepsilon(\Delta h_0 + \Delta h_1)V(X)]}, \quad (\text{B5})$$

where Z_ε is the equilibrium partition function and \mathcal{H}_0 is the Hamiltonian in the absence of the perturbation. The second-order response of any observable O can be calculated by expanding the above weight around $\varepsilon = 0$, multiplying by O , and summing over all possible configurations. This straightforward exercise leads to a formal expression,

$$\begin{aligned} \langle O \rangle^{(2)} = & \frac{\beta^2}{2} (\Delta h_0 + \Delta h_1)^2 [\langle OV^2 \rangle - \langle O \rangle \langle V^2 \rangle] \\ & + 2\langle O \rangle \langle V \rangle^2 - 2\langle OV \rangle \langle V \rangle. \end{aligned}$$

For the case $O(X) = X$ and $V(X) = 2X - 1$ with $\Delta h_0 = \Delta h_1 = 1$ the above expression simplifies to

$$\langle O \rangle^{(2)} = 8\beta^2 \langle X \rangle^{\text{eq}} (1 - 2\langle X \rangle^{\text{eq}}) (1 - \langle X \rangle^{\text{eq}}), \quad (\text{B6})$$

where we have used the fact that $X^2 = X$.

- [1] U. Seifert, *Rep. Prog. Phys.* **75**, 126001 (2012).
- [2] C. Jarzynski, *Phys. Rev. Lett.* **78**, 2690 (1997).
- [3] G. E. Crooks, *Phys. Rev. E* **60**, 2721 (1999).
- [4] H. B. Callen and T. A. Welton, *Phys. Rev.* **83**, 34 (1951).
- [5] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II: Nonequilibrium Statistical Mechanics* (Springer Science & Business Media, 2012), Vol. 31.
- [6] G. Diezemann, *Phys. Rev. E* **85**, 051502 (2012).
- [7] U. Basu, M. Krüger, A. Lazarescu, and C. Maes, *Phys. Chem. Chem. Phys.* **17**, 6653 (2015).
- [8] T. Yamada and K. Kawasaki, *Prog. Theor. Phys.* **38**, 1031 (1967).
- [9] D. J. Evans and G. P. Morriss, *Mol. Phys.* **64**, 521 (1988).
- [10] D. Andrieux and P. Gaspard, *J. Stat. Mech. Theor. Exp.* (2007) P02006.
- [11] J.-P. Bouchaud and G. Biroli, *Phys. Rev. B* **72**, 064204 (2005).
- [12] E. Lippiello, F. Corberi, A. Sarracino, and M. Zannetti, *Phys. Rev. B* **77**, 212201 (2008).
- [13] M. Colangeli, C. Maes, and B. Wynants, *J. Phys. A* **44**, 095001 (2011).
- [14] V. Lucarini and M. Colangeli, *J. Stat. Mech.: Theor. Exp.* (2012) P05013.
- [15] R. Kubo and K. Tomita, *J. Phys. Soc. Jpn.* **9**, 888 (1954).
- [16] L. Helden, U. Basu, M. Krüger, and C. Bechinger, *Europhys. Lett.* **116**, 60003 (2017).
- [17] U. Seifert, *Eur. Phys. J. B* **64**, 423 (2008).
- [18] A. Altland and B. D. Simons, *Condensed Matter Field Theory* (Cambridge University Press, Cambridge, 2010).
- [19] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, Oxford, 2001).
- [20] M. te Vrugt and R. Wittkowski, *Phys. Rev. E* **99**, 062118 (2019).
- [21] H. Mori, *Prog. Theor. Phys.* **33**, 423 (1965).
- [22] B. Bravi and P. Sollich, *Phys. Biol.* **14**, 045010 (2017).
- [23] S. Rahav and C. Jarzynski, *J. Stat. Mech.: Theor. Exp.* (2007) P09012.
- [24] M. Esposito, *Phys. Rev. E* **85**, 041125 (2012).
- [25] B. Altaner and J. Vollmer, *Phys. Rev. Lett.* **108**, 228101 (2012).
- [26] T. Vettorel, G. Besold, and K. Kremer, *Soft Matter* **6**, 2282 (2010).
- [27] E. J. Sambriski and M. G. Guenza, *Phys. Rev. E* **76**, 051801 (2007).
- [28] M. G. Saunders and G. A. Voth, *Ann. Rev. Biophys.* **42**, 73 (2013).
- [29] G. S. Ayton and G. A. Voth, *J. Phys. Chem. B* **113**, 4413 (2009).
- [30] J. Mehl, B. Lander, C. Bechinger, V. Blickle, and U. Seifert, *Phys. Rev. Lett.* **108**, 220601 (2012).
- [31] M. Kahlen and J. Ehrlich, *J. Stat. Mech.: Theor. Exp.* (2018) 063204.
- [32] R. García-García, S. Lahiri, and D. Lacoste, *Phys. Rev. E* **93**, 032103 (2016).
- [33] M. Ribezzi-Crivellari and F. Ritort, *Proc. Natl. Acad. Sci. (USA)* **111**, E3386 (2014).
- [34] U. Basu, L. Helden, and M. Krüger, *Phys. Rev. Lett.* **120**, 180604 (2018).
- [35] B. Wynants, Ph.D. thesis, KU Leuven, 2010.
- [36] We consider fields h that are even under time reversal since they are scalar. If h is itself odd, its sign also has to be flipped in time reversal to obtain the proper entropy production.
- [37] This differs from previous notations in Refs. [7,16], where these quantities are denoted by $\chi^{(n)}$ or χ_n .
- [38] M. Baiesi, C. Maes, and B. Wynants, *Phys. Rev. Lett.* **103**, 010602 (2009).
- [39] An expression of the two forms appearing in Eq. (25) in terms of one-step probabilities may be found in Appendix A in Eq. (A5).
- [40] W. N. Findley, J. S. Lai, and K. Onaran, *Creep and Relaxation of Nonlinear Viscoelastic Materials* (North-Holland, Amsterdam, 1976), Chap. 7.
- [41] A. Alemany, M. Ribezzi-Crivellari, and F. Ritort, *New J. Phys.* **17**, 075009 (2015).
- [42] J. Lebowitz and H. A. Spohn, *J. Stat. Phys.* **95**, 333 (1999).