Entropy production in full phase space for continuous stochastic dynamics

Richard E. Spinney and Ian J. Ford
Department of Physics and Astronomy, UCL, Gower Street, London WC1E 6BT, United Kingdom and London Centre for Nanotechnology, 17-19 Gordon Street, London WC1H 0AH, United Kingdom
(Received 2 March 2012; published 10 May 2012)

Total entropy production and its three constituent components are described both as fluctuating trajectory-dependent quantities and as averaged contributions in the context of the continuous Markovian dynamics, described by stochastic differential equations with multiplicative noise, of systems with both odd and even coordinates with respect to time reversal, such as dynamics in full phase space. Two of these constituent quantities obey integral fluctuation theorems and are thus rigorously positive in the mean due to Jensen's inequality. The third, however, is not and furthermore cannot be uniquely associated with irreversibility arising from relaxation, nor with the breakage of detailed balance brought about by nonequilibrium constraints. The properties of the various contributions to total entropy production are explored through the consideration of two examples: steady-state heat conduction due to a temperature gradient, and transitions between stationary states of drift diffusion on a ring, both in the context of the full phase space dynamics of a single Brownian particle.

DOI: 10.1103/PhysRevE.85.051113 PACS number(s): 05.70.Ln, 05.40.−a

I. INTRODUCTION

The concept of entropy was introduced over 150 years ago to provide a measure of the evident irreversibility of macroscopic thermodynamic phenomena. The conflict between its monotonic increase and the underlying time-reversal symmetry of the microscopic dynamics, first pointed out by Loschmidt, is but one of its apparent mysteries. Nevertheless, in recent years significant insights into the nature of irreversibility and entropy production have emerged, partly due to the need for a framework to interpret thermodynamic processes for small systems. These developments had their beginnings in the dissipation function and fluctuation theorem in deterministic thermostatted systems considered by Evans *et al*. [1–4], and they have continued with similar concepts within the realms of chaos theory [5] and of stochastic dynamical modeling [6,7]. Some powerful results, such as the Crooks and Jarzynski relations, stand out [8–10] along with a unifying framework for overdamped Langevin dynamics [11] based on a stochastic description of the first law of thermodynamics commonly referred to as stochastic energetics [12]. In short, entropy production is a measure of the relative likelihoods of forward and reversed behavior within the context of a model dynamical framework that includes specific dissipative terms. There are certain differences in viewpoint, but the central insight is that a mechanical (or maybe dynamical) quantity can be defined that matches the behavior of thermodynamic entropy, in particular that on average it increases with time. Its fluctuating nature provides additional insight into the behavior of small systems.

More recently, it was proposed that the entropy production associated with nonequilibrium states of small systems, arising from an underlying stochastic model of the dynamics, could be divided into two components, one related to relaxation (sometimes restricted to transitions between stationary states), and the other to any fundamental constraint that maintains the system away from an equilibrium [13–21]. The two components, termed adiabatic and nonadiabatic production rates, respectively, were mapped onto earlier concepts known as excess and housekeeping heat transfers [19,22]. In a recent development [23], however, it was shown that a third component of entropy production could be conceived, arising from the nonequilibrium constraint, but associated with relaxation toward the stationary state. It only arises when odd dynamical variables play a role in the dynamics, and even then only in specific cases. Only two of the three components of entropy production satisfy an integral fluctuation theorem (IFT), making them rigorously positive in the mean, properties shared by the sum of all three; the third, however, does not satisfy an IFT, and in the mean it can take either sign. In this paper, we develop these ideas further, within a framework of full phase-space continuous dynamics modeled by stochastic differential equations, with the aim of pinning down the specific form of the three contributions, both in the mean and in fluctuations about the mean, and we go on to make use of the formalism in some instructive example systems.

II. THREE CONTRIBUTIONS TO ENTROPY PRODUCTION

We begin by considering the dynamics of a general set of variables $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ that may be odd or even under time reversal by considering the operation $\boldsymbol{\varepsilon x} = (\varepsilon_1 x_1, \varepsilon_2 x_2, \ldots, \varepsilon_n x_n)$ where $\varepsilon_i = \pm 1$ for even and odd variables x_i, respectively. Specifically, we consider continuous Markovian dynamics described by a system of arbitrary uncorrelated Ito stochastic differential equations (SDEs) such that the evolution of the coordinates $\boldsymbol{x}$ is given as

$$dx_i = A_i(\boldsymbol{x},t)dt + B_i(\boldsymbol{x},t)dW_i, \tag{1}$$

where dW_i denotes the Wiener process. Since we allow x_i to be either odd or even under time reversal, we can divide the deterministic dynamics into reversible and irreversible components [24] such that

$$dx_i = A_i^{\mathrm{rev}}(\boldsymbol{x},t)dt + A_i^{\mathrm{ir}}(\boldsymbol{x},t)dt + B_i(\boldsymbol{x},t)dW_i \tag{2}$$

by defining

$$A_i^{\mathrm{ir}}(\boldsymbol{x},t) = \tfrac{1}{2}\left[A_i(\boldsymbol{x},t) + \varepsilon_i A_i(\boldsymbol{\varepsilon x},t)\right] = \varepsilon_i A_i^{\mathrm{ir}}(\boldsymbol{\varepsilon x},t), \tag{3}$$

$$A_i^{\mathrm{rev}}(\boldsymbol{x},t) = \tfrac{1}{2}\left[A_i(\boldsymbol{x},t) - \varepsilon_i A_i(\boldsymbol{\varepsilon x},t)\right] = -\varepsilon_i A_i^{\mathrm{rev}}(\boldsymbol{\varepsilon x},t). \tag{4}$$

We briefly note that we intend our notation, $A(\boldsymbol{\varepsilon x},t)$, to imply a time reversal of all parameters that constitute A whether they be dynamical variables included in $\boldsymbol{x}$ or not. For example, a term proportional to a magnetic field appearing in A_x, where x is an even spatial coordinate, would form part of A_x^{rev} since magnetic fields are odd with respect to time reversal in contrast with, for example, a force F which would appear in A_x^{ir} since force is even with respect to time reversal.

We have specified for simplicity that all our SDEs are driven by uncorrelated noise such that we have no cross derivatives in the corresponding Fokker-Planck equation. As such, we may then represent the noise strengths as diffusion coefficients

$$D_i(\boldsymbol{x},t) = \tfrac{1}{2}B_i(\boldsymbol{x},t)^2 \tag{5}$$

that appear in a Fokker-Planck equation describing the joint probability density function of the coordinates,

$$\frac{\partial p(\boldsymbol{x},t)}{\partial t} = -\sum_i \frac{\partial}{\partial x_i}\left[A_i(\boldsymbol{x},t)p(\boldsymbol{x},t)\right] + \sum_i \frac{\partial^2}{\partial x_i^2}\left[D_i(\boldsymbol{x},t)p(\boldsymbol{x},t)\right]. \tag{6}$$

In order to proceed in the later development, we assume that the diffusion coefficient is symmetric with respect to time reversal such that $D_i(\boldsymbol{\varepsilon x}) = D_i(\boldsymbol{x})$, which puts no restriction on the dependence on even coordinates but requires that $D_i(\boldsymbol{x})$ is an even function of any odd coordinates.

It is helpful to express the Fokker-Planck equation as a continuity equation in terms of the probability density current $\boldsymbol{J}(\boldsymbol{x},t)$,

$$\frac{\partial p(\boldsymbol{x},t)}{\partial t} = -\nabla\cdot\boldsymbol{J}(\boldsymbol{x},t) = -\nabla\cdot[\boldsymbol{J}^{\mathrm{ir}}(\boldsymbol{x},t) + \boldsymbol{J}^{\mathrm{rev}}(\boldsymbol{x},t)], \tag{7}$$

which we separate into irreversible and reversible components. These take vector form $\boldsymbol{J} = (J_1,J_2,\ldots,J_n)$ as do the drift and diffusion coefficients $\boldsymbol{A} = (A_1,A_2,\ldots,A_n)$ and $\boldsymbol{D} = (D_1,D_2,\ldots,D_n)$ such that

$$\begin{aligned}\boldsymbol{J}^{\mathrm{ir}}(\boldsymbol{x},t) &= \boldsymbol{A}^{\mathrm{ir}}(\boldsymbol{x},t)p(\boldsymbol{x},t) - \nabla\cdot(\boldsymbol{D}(\boldsymbol{x},t)p(\boldsymbol{x},t)),\\ \boldsymbol{J}^{\mathrm{rev}}(\boldsymbol{x},t) &= \boldsymbol{A}^{\mathrm{rev}}(\boldsymbol{x},t)p(\boldsymbol{x},t).\end{aligned} \tag{8}$$

Having set out the dynamics we shall be using, we now consider the general procedure for producing quantities which obey IFTs. Given an interval of duration τ, such a quantity consists of a difference between the logarithmic probability density of a given trajectory under what we shall term the forward dynamics, time dependence of the dynamics (equivalent here to an external protocol), and initial distribution of starting configurations, and that of another appropriately chosen trajectory under suitable dynamics, protocol, and initial distribution. We write $\mathcal{P}[\vec{x}]$ as the probability density of the forward trajectory or path, $\vec{x} = \boldsymbol{x}(t)$ for $0 \leqslant t \leqslant \tau$, with a probability density function of starting configurations, $p(\boldsymbol{x}(0),0)$, which acts as an initial condition for the Fokker-Planck equation introduced earlier. A quantity that obeys an IFT is then of the form

$$\mathcal{A}[\vec{x}] = \ln[\mathcal{P}[\vec{x}]/\mathcal{P}^*[\vec{x}^*]], \tag{9}$$

where $\mathcal{P}^*[\vec{x}^*]$ is the probability density of a path, $\vec{x}^*$, under chosen dynamics (with specified nature and time dependence) and initial condition. Demonstrating the adherence of such a quantity to an IFT with respect to the forward dynamics and time dependence is straightforward by the reasoning

$$\begin{aligned}\langle\exp[-\mathcal{A}[\vec{x}]]\rangle &= \int d\vec{x}\ \mathcal{P}[\vec{x}]\exp[-\mathcal{A}[\vec{x}]]\\ &= \int d\vec{x}\ \mathcal{P}[\vec{x}]\frac{\mathcal{P}^*[\vec{x}^*]}{\mathcal{P}[\vec{x}]} = \int d\vec{x}^*\ \mathcal{P}^*[\vec{x}^*] = 1.\end{aligned} \tag{10}$$

Such a result requires a Jacobian of unity for the path transformation $\vec{x} \to \vec{x}^*$ in order for the path integrals to be equivalent (a result assured for any involutive transformation) and the requirement $\mathcal{P}^*[\vec{x}^*] = 0$ for all $\mathcal{P}[\vec{x}] = 0$ ensuring that for normalized $\mathcal{P}$ and $\mathcal{P}^*$ all possible paths under the dynamics that produce $\mathcal{P}^*$ are contained within the final integral. This may be seen as a version of the so-called ergodic consistency requirement [25]. Further, any quantity $\mathcal{A}[\vec{x}]$ based on a transformation with a Jacobian of unity takes the same form irrespective of whether probabilities or probability densities are used in the construction due to the equivalence of measure. Such a quantity is therefore constructed unambiguously and is the direct analog of the same quantity defined in discrete space [19,23]. The implication of the positivity in the mean, $\langle\mathcal{A}[\vec{x}]\rangle \geqslant 0$, of the quantity $\mathcal{A}[\vec{x}]$ is assured by Jensen's inequality. We point out that the path, $\vec{x}^*$, and the dynamics must be carefully chosen in order to satisfy the requirements and to produce a physically meaningful quantity. This is particularly relevant in the presence of both odd and even variables as for many physical systems the common choice of reverse path $\boldsymbol{x}^*(t) = \boldsymbol{x}(\tau - t)$ cannot be generated under the forward dynamics, and if used to define a quantity of the form in Eq. (9), it would render the final integral in Eq. (10) equal to zero. However, the choice $\boldsymbol{x}^*(t) = \boldsymbol{\varepsilon x}(\tau - t)$ typically can be generated and so leads to an IFT. However, this choice of the second path, $\boldsymbol{x}^*(t) = \boldsymbol{\varepsilon x}(\tau - t)$, is appropriate not just because of the guarantee of an IFT, but because it means the constructed quantity $A[\vec{x}]$ serves as a measure of the irreversibility of the process and thus characterizes the total entropy production.

By following the above rules, and making the definitions $\boldsymbol{x}^\dagger(t) = \boldsymbol{\varepsilon x}(\tau - t)$, $\boldsymbol{x}^{\mathrm{R}}(t) = \boldsymbol{x}(\tau - t)$, and $\boldsymbol{x}^{\mathrm{T}}(t) = \boldsymbol{\varepsilon x}(t)$, we may construct path-dependent dimensionless entropy changes, which are thermodynamically meaningful when multiplied by k_B, of the form in Eq. (9):

$$\begin{aligned}\Delta\mathcal{S}_{\mathrm{tot}} &= \ln\mathcal{P}[\vec{x}] - \ln\mathcal{P}^{\mathrm{R}}[\vec{x}^\dagger]\\ &= \ln\frac{p(\boldsymbol{x}(0),0)}{p(\boldsymbol{x}(\tau),\tau)} + \ln\frac{\mathcal{P}[\boldsymbol{x}(\tau)|\boldsymbol{x}(0)]}{\mathcal{P}^{\mathrm{R}}[\boldsymbol{\varepsilon x}(0)|\boldsymbol{\varepsilon x}(\tau)]},\end{aligned} \tag{11}$$

$$\begin{aligned}\Delta\mathcal{S}_1 &= \ln\mathcal{P}[\vec{x}] - \ln\mathcal{P}^{\mathrm{ad,R}}[\vec{x}^{\mathrm{R}}]\\ &= \ln\frac{p(\boldsymbol{x}(0),0)}{p(\boldsymbol{x}(\tau),\tau)} + \ln\frac{\mathcal{P}[\boldsymbol{x}(\tau)|\boldsymbol{x}(0)]}{\mathcal{P}^{\mathrm{ad,R}}[\boldsymbol{x}(0)|\boldsymbol{x}(\tau)]},\end{aligned} \tag{12}$$

$$\begin{aligned}\Delta\mathcal{S}_2 &= \ln\mathcal{P}[\vec{x}] - \ln\mathcal{P}^{\mathrm{ad}}[\vec{x}^{\mathrm{T}}]\\ &= \ln\frac{p(\boldsymbol{x}(0),0)}{p(\boldsymbol{x}(0),0)} + \ln\frac{\mathcal{P}[\boldsymbol{x}(\tau)|\boldsymbol{x}(0)]}{\mathcal{P}^{\mathrm{ad}}[\boldsymbol{\varepsilon x}(\tau)|\boldsymbol{\varepsilon x}(0)]},\end{aligned} \tag{13}$$

such that the total path probability densities are divided into initial probability density distributions [which also appear as initial conditions and solutions to the Fokker-Planck equation in Eq. (6)] and conditional probability densities. Here $\Delta\mathcal{S}_{\mathrm{tot}}$

amounts to the total entropy production of the universe with notation deriving from [11] while the other two consider the origins of that total entropy production. In the above, the label R designates a reversed protocol, equivalent here to reversed time dependence in the dynamics, and "ad" designates that the dynamics are so-called adjoint with respect to the forward dynamics, defined as the dynamics which reach the same stationary state but with the opposite stationary current [15,19,25]. Physically, the time-reversed dynamics would be achieved by applying a time-dependent protocol in reverse (with appropriate sign change should that protocol be odd, e.g., a magnetic field). The adjoint dynamics, however, would be achieved by alteration of the functional dependence of the $A_i(\boldsymbol{x})$ terms. For an overdamped particle in a potential subject to stationary flux in one direction, the adjoint dynamics would be such an alteration which produced the opposite flux with appropriate change to the potential so as to preserve the stationary distribution.

All three entropy productions above are expected to obey IFTs by the nature of their form. Then, by the construction $\Delta\mathcal{S}_{\rm tot} = \Delta\mathcal{S}_1 + \Delta\mathcal{S}_2 + \Delta\mathcal{S}_3$, we define

$$\begin{aligned}\Delta\mathcal{S}_3 &= \ln\mathcal{P}^{\rm ad}[\vec{\boldsymbol{x}}^{\rm T}] + \ln\mathcal{P}^{\rm ad,R}[\vec{\boldsymbol{x}}^{\rm R}] - \ln\mathcal{P}[\vec{\boldsymbol{x}}] - \ln\mathcal{P}^{\rm R}[\vec{\boldsymbol{x}}^{\dagger}]\\ &= \ln\frac{\mathcal{P}^{\rm ad,R}[\boldsymbol{x}(0)|\boldsymbol{x}(\tau)]\mathcal{P}^{\rm ad}[\boldsymbol{\varepsilon x}(\tau)|\boldsymbol{\varepsilon x}(0)]}{\mathcal{P}[\boldsymbol{x}(\tau)|\boldsymbol{x}(0)]\mathcal{P}^{\rm R}[\boldsymbol{\varepsilon x}(0)|\boldsymbol{\varepsilon x}(\tau)]},\end{aligned} \tag{14}$$

which cannot be expressed in the form of Eq. (9) and so does not obey an IFT. By following the formalism of Seifert [11,26], we identify

$$\Delta\mathcal{S}_{\rm tot} = \ln\frac{p(\boldsymbol{x}(0),0)}{p(\boldsymbol{x}(\tau),\tau)} + \Delta\mathcal{S}_{\rm med} = \Delta\mathcal{S}_{\rm sys} + \Delta\mathcal{S}_{\rm med}, \tag{15}$$

where $\Delta\mathcal{S}_{\rm sys}$ is known as the change in dimensionless system entropy and $\Delta\mathcal{S}_{\rm med}$ is a generalization of the dimensionless entropy production in the environment, or medium. We consider it to include the effect of any external agent acting on the system such that, for example, it contains both the heat flow and "pumped" entropy discussed for velocity-dependent forcing in [27,28]. If one can consider a defined environmental temperature, $T_{\rm env}$, which we allow to be phase-space or time-dependent (a property we employ in example I), we may write this in terms of a heat flow,

$$\Delta\mathcal{S}_{\rm tot} = \ln\frac{p(\boldsymbol{x}(0),0)}{p(\boldsymbol{x}(\tau),\tau)} + \int_{t=0}^{t=\tau} d\left(\frac{\Delta Q}{k_B T_{\rm env}(\boldsymbol{x}(t),t)}\right). \tag{16}$$

The integral should be interpreted as a total medium entropy change, considered as the sum of all the incremental heat transfers to separate fixed temperature heat baths to which the particle is exposed over the course of its trajectory, divided by the appropriate temperature. Of course, this reproduces the usual $\Delta\mathcal{S}_{\rm med} = \Delta Q/k_B T_{\rm env}$ of stochastic thermodynamics [11] when the temperature is constant. Similarly, when such a temperature can be defined, we may connect the three entropy contributions with a key concept in nonequilibrium thermodynamics by dividing the heat transfer to the environment (considering here a single, constant temperature for clarity without loss of generality) into the so-called excess and housekeeping heats according to the formalism of Oono and Paniconi [22]. The housekeeping heat is usually defined as the heat transfer required to maintain a nonequilibrium steady state and the excess heat forms the remainder of the total heat flow such that for a given environmental temperature, $\Delta Q = \Delta Q_{\rm ex} + \Delta Q_{\rm hk}$. To align our quantities with such a formalism in such cases, we associate $\Delta\mathcal{S}_1$ with the excess heat,

$$\Delta Q_{\rm ex} = (\Delta\mathcal{S}_1 - \Delta\mathcal{S}_{\rm sys})k_B T_{\rm env}, \tag{17}$$

$\Delta\mathcal{S}_2$ with a so called "generalized housekeeping heat" [23],

$$\Delta Q_{\rm hk,G} = \Delta\mathcal{S}_2 k_B T_{\rm env}, \tag{18}$$

and $\Delta\mathcal{S}_3$ with the "transient housekeeping heat,"

$$\Delta Q_{\rm hk,T} = \Delta\mathcal{S}_3 k_B T_{\rm env}, \tag{19}$$

named to reflect its mean behavior, such that $\Delta Q_{\rm hk} = \Delta Q_{\rm hk,G} + \Delta Q_{\rm hk,T}$. Here, and throughout, the association with heat flows necessarily requires a well-defined environmental temperature; however, we stress that the three entropy contributions do not require such a specification and so can be applied more generally. Like previous formalisms [19,21] where the entropy production was divided into two contributions associated with relaxation, and an absence of detailed balance, respectively, both in the mean and in detail [19–21], we have a contribution $\Delta\mathcal{S}_1$ which is nonzero only in the presence of relaxation and $\Delta\mathcal{S}_2$ which is nonzero only in the absence of detailed balance both in the mean and in detail. However, we also have a quantity $\Delta\mathcal{S}_3$ which is nonzero in detail only in the absence of detailed balance, but only contributes in the mean during the course of relaxation. Such a formalism asserts that the two origins of entropy production may often be more closely related, with such a circumstance arising under the inclusion of odd variables and when the stationary distribution is asymmetric in any of those odd variables. The aim of this paper is to derive the equations of motion for each of these quantities for continuous stochastic systems and to illustrate their behavior through some simple examples.

III. REPRESENTING ENTROPY PRODUCTION FOR CONTINUOUS BEHAVIOR

A. Entropy production as an SDE

Since we are describing the dynamics using SDEs, it is sensible to seek a description of a small increment in each entropy production given an increment in the underlying variables $\boldsymbol{x}' - \boldsymbol{x} = \boldsymbol{x}(t+dt) - \boldsymbol{x}(t)$ in a time dt so that we identify from Eqs. (11)–(14)

$$d\Delta\mathcal{S}_{\rm tot} = -d(\ln p) + \ln\frac{\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)}{\mathcal{P}(\boldsymbol{\varepsilon x},t+dt|\boldsymbol{\varepsilon x}',t)}, \tag{20}$$

$$d\Delta\mathcal{S}_1 = -d(\ln p) + \ln\frac{\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)}{\mathcal{P}^{\rm ad}(\boldsymbol{x},t+dt|\boldsymbol{x}',t)}, \tag{21}$$

$$d\Delta\mathcal{S}_2 = \ln\frac{\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)}{\mathcal{P}^{\rm ad}(\boldsymbol{\varepsilon x}',t+dt|\boldsymbol{\varepsilon x},t)}, \tag{22}$$

$$d\Delta\mathcal{S}_3 = \ln\frac{\mathcal{P}^{\rm ad}(\boldsymbol{x},t+dt|\boldsymbol{x}',t)\mathcal{P}^{\rm ad}(\boldsymbol{\varepsilon x}',t+dt|\boldsymbol{\varepsilon x},t)}{\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)\mathcal{P}(\boldsymbol{\varepsilon x},t+dt|\boldsymbol{\varepsilon x}',t)}, \tag{23}$$

thereby establishing the SDEs that describe entropy production and noting the abbreviation $d[\ln(p)] = \ln[p(\boldsymbol{x}(t+dt),t+dt)/p(\boldsymbol{x}(t),t)]$.

To proceed, we require a representation of the path probabilities in these expressions that is valid over the small time interval dt. This may be achieved by considered the short time Green's function or "short time propagator" [29] which is given generally as the conditional probability of a displacement $dx = x' - x$ in a time dt subject to a δ function initial condition and is of the form

$$\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t) = \prod_i \sqrt{\frac{1}{4\pi D_i(\boldsymbol{r},t)dt}} \exp\left[-\frac{\{dx_i - A_i(\boldsymbol{r},t)dt + 2a[\partial D_i(\boldsymbol{r},t)/\partial r_i]dt\}^2}{4D_i(\boldsymbol{r})dt} - a dt \frac{\partial A_i(\boldsymbol{r},t)}{\partial r_i} + a^2 dt \frac{\partial^2 D_i(\boldsymbol{r},t)}{\partial r_i^2}\right], \quad (24)$$

where $dx_i = x_i' - x_i$ and where a is a free parameter ranging from 0 to 1, which defines the evaluation point of certain terms in the propagator $\boldsymbol{r} = a\boldsymbol{x}' + (1-a)\boldsymbol{x}$ and $r_i = a x_i' + (1-a)x_i$, and which reflects the ambiguity of a discretized interpretation of continuous stochastic behavior. We note, however, that as $dt \to 0$, all forms for the propagator are correct: they are all accurate to first order in dt and result in the same Fokker-Planck equation.

We wish to construct the increment in entropy production in the medium by a consideration of

$$d\Delta\mathcal{S}_{\text{med}} = \ln \frac{\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)}{\mathcal{P}(\boldsymbol{\varepsilon x},t+dt|\boldsymbol{\varepsilon x}',t)} \quad (25)$$

by employing the appropriate reverse short time propagator

$$\mathcal{P}(\boldsymbol{\varepsilon x},t+dt|\boldsymbol{\varepsilon x}',t) = \prod_i \sqrt{\frac{1}{4\pi D_i(\boldsymbol{\varepsilon r}',t)dt}} \exp\left[-\frac{\{-\varepsilon_i dx_i - A_i(\boldsymbol{\varepsilon r}',t)dt + 2b[\partial D_i(\boldsymbol{\varepsilon r}',t)/\partial(\varepsilon_i r_i')]dt\}^2}{4D_i(\boldsymbol{\varepsilon r}')dt} - b dt \frac{\partial A_i(\boldsymbol{\varepsilon r}',t)}{\partial(\varepsilon_i r_i')} + b^2 dt \frac{\partial^2 D_i(\boldsymbol{\varepsilon r}',t)}{\partial(\varepsilon_i r_i')^2}\right], \quad (26)$$

where b is a corresponding free parameter ranging from 0 to 1 such that $\boldsymbol{r}' = b\boldsymbol{x} + (1-b)\boldsymbol{x}'$ and $r_i' = b x_i + (1-b)x_i'$. Using Eqs. (3) and (4) along with the assumption $D_i(\boldsymbol{\varepsilon x}) = D_i(\boldsymbol{x})$ such that any diffusion constants are symmetric in odd variables, we may write

$$\mathcal{P}(\boldsymbol{\varepsilon x},t+dt|\boldsymbol{\varepsilon x}',t) = \prod_i \sqrt{\frac{1}{4\pi D_i(\boldsymbol{r}',t)dt}} \exp\Bigg[-\frac{\left\{-\varepsilon_i dx_i - \varepsilon_i\left[-A_i^{\text{rev}}(\boldsymbol{r}',t) + A_i^{\text{ir}}(\boldsymbol{r}',t)\right]dt + 2b[\partial D_i(\boldsymbol{r}',t)/\partial(\varepsilon_i r_i')]dt\right\}^2}{4D_i(\boldsymbol{r}')dt}$$
$$-b dt\left(\frac{\partial \varepsilon_i A_i^{\text{ir}}(\boldsymbol{r}',t)}{\partial(\varepsilon_i r_i')} - \frac{\partial \varepsilon_i A_i^{\text{rev}}(\boldsymbol{r}',t)}{\partial(\varepsilon_i r_i')}\right) + b^2 dt \frac{\partial^2 D_i(\boldsymbol{r}',t)}{\partial(\varepsilon_i r_i')^2}\Bigg], \quad (27)$$

which is the same as

$$\mathcal{P}(\boldsymbol{\varepsilon x},t+dt|\boldsymbol{\varepsilon x}',t) = \prod_i \sqrt{\frac{1}{4\pi D_i(\boldsymbol{r}',t)dt}} \exp\Bigg[-\frac{\{-dx_i - \left[-A_i^{\text{rev}}(\boldsymbol{r}',t) + A_i^{\text{ir}}(\boldsymbol{r}',t)\right]dt + 2b[\partial D_i(\boldsymbol{r}',t)/\partial r_i']dt\}^2}{4D_i(\boldsymbol{r}')dt}$$
$$-b\, dt\left(\frac{\partial A_i^{\text{ir}}(\boldsymbol{r}',t)}{\partial r_i'} - \frac{\partial A_i^{\text{rev}}(\boldsymbol{r}',t)}{\partial r_i'}\right) + b^2 dt \frac{\partial^2 D_i(\boldsymbol{r}',t)}{\partial r_i'^2}\Bigg]. \quad (28)$$

The mathematical details necessary for the development of Eq. (25), which due to their somewhat cumbersome nature we leave to Appendix A, reveal that for multiplicative noise one obtains a result which is dependent on the choice a and b. The resolution of this apparent arbitrariness is not related to the nature of the underlying SDEs, but rather on consistently using the equivalent evaluation point for forward and time-reversed paths on an infinitesimal scale. The normal rules of calculus would dictate no dependence, but different rules apply to SDEs and stochastic calculus (for which we direct the reader again to Appendix A). Such a consideration reveals the correct choice in Eq. (28) to be $b = 1 - a$ with a remaining as a free parameter. This yields the unambiguous Ito SDE for the medium entropy change,

$$d\Delta\mathcal{S}_{\text{med}} = \sum_i \frac{A_i^{\text{ir}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}dx_i - \frac{A_i^{\text{rev}}(\boldsymbol{x})A_i^{\text{ir}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}dt + \frac{\partial A_i^{\text{ir}}(\boldsymbol{x})}{\partial x_i}dt - \frac{\partial A_i^{\text{rev}}(\boldsymbol{x})}{\partial x_i}dt - \frac{1}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}dx_i + \frac{[A^{\text{rev}}(\boldsymbol{x}) - A^{\text{ir}}(\boldsymbol{x})]}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}dt - \frac{\partial^2 D_i(\boldsymbol{x})}{\partial x_i^2}dt + \frac{1}{D_i(\boldsymbol{x})}\left(\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\right)^2 dt, \quad (29)$$

where for brevity we use the notation $f(\boldsymbol{x}) \equiv f(\boldsymbol{x},t)$. To clarify, in such an approach choices may include Stratonovich

($a = b = 1/2$) evaluation for both propagators in Eq. (25), a choice which is implicitly used by many authors [15,26] within integrated Onsager-Machlup approaches, but does not preclude others in the construction of SDEs, such as, for example, an Ito prescription ($a = 0$) in the forward propagator and a Hanggi-Klimontovich ($b = 1$) in the backward propagator. We point out that all evaluation points lead to the correct *path probability* when supplemented with the correct multiplication scheme, but that if one has multiplicative noise, the correct representation of the *entropy production* requires the more exact relation between the evaluation points.

Proceeding, we may now construct an SDE for the total entropy production by first considering an increment in the system entropy, which under Ito rules is

$$\begin{aligned} d\Delta\mathcal{S}_{\text{sys}} &= -d[\ln p(\boldsymbol{x})] \\ &= -\frac{1}{p(\boldsymbol{x})}\frac{\partial p(\boldsymbol{x})}{\partial t}dt - \frac{1}{p(\boldsymbol{x})}\sum_i \frac{\partial p(\boldsymbol{x})}{\partial x_i}dx_i \\ &\quad - \sum_i \frac{D_i(\boldsymbol{x})}{p(\boldsymbol{x})}\left[\frac{\partial^2 p(\boldsymbol{x})}{\partial x_i^2} - \frac{1}{p(\boldsymbol{x})}\left(\frac{\partial p(\boldsymbol{x})}{\partial x_i}\right)^2\right]dt, \end{aligned} \tag{30}$$

which together with Eq. (29), and after insertion of the Fokker-Planck equation, leads to

$$\begin{aligned} d\Delta\mathcal{S}_{\text{tot}} = \sum_i & -\frac{1}{p(\boldsymbol{x})}\frac{\partial p(\boldsymbol{x})}{\partial x_i}dx_i + \frac{1}{p(\boldsymbol{x})}\frac{\partial[A_i(\boldsymbol{x})p(\boldsymbol{x})]}{\partial x_i}dt - \frac{1}{p(\boldsymbol{x})}\left[\frac{\partial^2[D_i(\boldsymbol{x})p(\boldsymbol{x})]}{\partial x_i^2} + D_i(\boldsymbol{x})\frac{\partial^2 p(\boldsymbol{x})}{\partial x_i^2} - \frac{D_i(\boldsymbol{x})}{p(\boldsymbol{x})}\left(\frac{\partial p(\boldsymbol{x})}{\partial x_i}\right)^2\right]dt \\ & + \frac{A_i^{\text{ir}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}dx_i - \frac{A_i^{\text{rev}}(\boldsymbol{x})A_i^{\text{ir}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}dt + \frac{\partial A_i^{\text{ir}}(\boldsymbol{x})}{\partial x_i}dt - \frac{\partial A_i^{\text{rev}}(\boldsymbol{x})}{\partial x_i}dt - \frac{1}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}dx_i + \frac{[A^{\text{rev}}(\boldsymbol{x}) - A^{\text{ir}}(\boldsymbol{x})]}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}dt \\ & - \frac{\partial^2 D_i(\boldsymbol{x})}{\partial x_i^2}dt + \frac{1}{D_i(\boldsymbol{x})}\left(\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\right)^2 dt. \end{aligned} \tag{31}$$

If Stratonovich rules, for example, are preferred, we can write (by definition of the Stratonovich integral, indicated by the $\circ$ notation)

$$\begin{aligned} d\Delta\mathcal{S}_{\text{tot}} = \sum_i & -\frac{1}{p(\boldsymbol{x})}\frac{\partial p(\boldsymbol{x})}{\partial x_i}\circ dx_i + \frac{1}{p(\boldsymbol{x})}\frac{\partial[A_i(\boldsymbol{x})p(\boldsymbol{x})]}{\partial x_i}dt - \frac{1}{p(\boldsymbol{x})}\left(\frac{\partial^2[D_i(\boldsymbol{x})p(\boldsymbol{x})]}{\partial x_i^2}\right)dt + \frac{A_i^{\text{ir}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}\circ dx_i - \frac{A_i^{\text{rev}}(\boldsymbol{x})A_i^{\text{ir}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}dt \\ & - D_i(\boldsymbol{x})\frac{\partial}{\partial x_i}\left(\frac{A_i^{\text{rev}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}\right)dt - \frac{1}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\circ dx_i. \end{aligned} \tag{32}$$

This is a very general and robust definition of the entropy production for continuous stochastic behavior and can be thought of as a generalization of the pioneering approach in [11] wherein the equation of motion for entropy essentially describes $d\Delta\mathcal{S}_{\text{tot}}$ for a specific system with additive noise, even variables ($\varepsilon\boldsymbol{x} = \boldsymbol{x}$), and implicitly using Stratonovich rules.

We point out that such a construction allows us to consider purely deterministic coordinates [$D_i(\boldsymbol{x}) = 0$] as would apply, for example, to the case of spatial coordinates within a full phase-space Langevin description. In such coordinates, $D_i(\boldsymbol{x})$ is assumed constant and taken to zero. The remaining terms then clearly diverge unless we demand $A_i^{\text{ir}}(\boldsymbol{x}) = 0$ since in these instances, for the reverse path to be a solution to the forward dynamics, the motion must be purely reversible. This condition simply amounts to the requirement that the reverse path exists. There is, however, a contribution to the medium entropy production, due to the dynamics of these coordinates, technically since path probability densities, not probabilities, are being considered in the formulation. The contribution to the medium entropy production due to the deterministic behavior of these coordinates is

$$\Delta\mathcal{S}_{\text{med,det}} = -\frac{\partial A_i^{\text{rev}}(\boldsymbol{x})}{\partial x_i}dt, \tag{33}$$

a result that provides an insight into the similarities and differences between stochastic and deterministic measures of irreversibility: it is demonstrably equal to the phase-space contraction found in nonlinear dynamical systems, which is associated with the heat transfer to the environment brought about by thermostatting terms in such approaches. This leads to a quantity that is positive in the mean for deterministic systems: the dissipation function [2]. We point out, however, that *total entropy production*, as defined here for stochastic systems, is zero for deterministic dynamics. This is because the change in the system entropy would be equal and opposite to the change in medium entropy, technically since it involves probability densities at the start and end of the process. In contrast, the dissipation function can provide a measure of irreversibility because it involves a comparison of trajectories originating *from the same* starting distribution. This contrast is to be expected as the total entropy production, as defined for the systems we consider, arises from explicit irreversibility in the dynamics, which deterministic, reversible equations do not provide.

B. The mean entropy production rate

Frequently, the average entropy production rate is argued to be proportional to the mean probability flux squared, as derived, for example, by taking the time derivative of the Gibbs entropy of a system, and identifying an evidently positive contribution as the total entropy production rate and the remainder as the (negative of) the medium entropy production rate [21,30]. We prefer, however, to derive the average contributions directly from the SDEs so that we can avoid arbitrarily identifying a positive contribution with a quantity expected to obey an IFT: strictly speaking, there is no guarantee such a division is unique, as another description shows [31]. To do so is straightforward and requires us to find the average increment in $\Delta\mathcal{S}_{\text{tot}}$ by means of the integral

$$\langle d\Delta\mathcal{S}_{\text{tot}}\rangle = \int d\boldsymbol{x}\int d\boldsymbol{x}'\; p(\boldsymbol{x},t)\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)d\Delta\mathcal{S}_{\text{tot}}. \tag{34}$$

The benefit of such a formulation is that we may characterize $d\Delta\mathcal{S}_{\text{tot}}$ using an Ito SDE based on the underlying relations $dx_i = A_i dt + B_i dW_i$ and then use the martingale property of the Ito stochastic integral $\langle B_i dW_i\rangle = 0$ since B_i is nonanticipating, such that we can simplify the integral in Eq. (34) by writing

$$\langle d\Delta\mathcal{S}_{\text{tot}}\rangle = \int d\boldsymbol{x}\; p(\boldsymbol{x})\langle d\Delta\mathcal{S}_{\text{tot}}|\boldsymbol{x}\rangle \tag{35}$$

and evaluating the conditional average $\langle d\Delta\mathcal{S}_{\text{tot}}|\boldsymbol{x}\rangle$ by replacing all occurrences of dx_i with $(A_i^{\text{ir}}+A_i^{\text{rev}})dt$ in $d\Delta\mathcal{S}_{\text{tot}}$. We thus get

$$\begin{aligned}\langle d\Delta\mathcal{S}_{\text{tot}}\rangle = \sum_i\Bigg[\int d\boldsymbol{x}\;&\frac{p(\boldsymbol{x})[A^{\text{ir}}(\boldsymbol{x})]^2}{D_i(\boldsymbol{x})} + 2p(\boldsymbol{x})\frac{\partial A_i^{\text{ir}}(\boldsymbol{x})}{\partial x_i}\\ &-2p(\boldsymbol{x})\frac{A^{\text{ir}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x_i} - p(\boldsymbol{x})\frac{\partial^2 D_i(\boldsymbol{x})}{\partial x_i^2}\\ &+\frac{p(\boldsymbol{x})}{D_i(\boldsymbol{x})}\left(\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\right)^2 - \frac{\partial^2[D_i(\boldsymbol{x})p(\boldsymbol{x})]}{\partial x_i^2}\\ &-D_i(\boldsymbol{x})\frac{\partial^2 p(\boldsymbol{x})}{\partial x_i^2} + \frac{D_i(\boldsymbol{x})}{p(\boldsymbol{x})}\left(\frac{\partial p(\boldsymbol{x})}{\partial x_i}\right)^2\Bigg]dt.\end{aligned} \tag{36}$$

By applying the product rule, integrating by parts, and assuming $p(\boldsymbol{x})$ and $\partial p(\boldsymbol{x})/\partial x_i$ either vanish or cancel at the boundaries, we may simplify to find the total entropy production rate

$$\frac{d\langle\Delta\mathcal{S}_{\text{tot}}\rangle}{dt} = \sum_i\int d\boldsymbol{x}\;\frac{\left[p(\boldsymbol{x})A_i^{\text{ir}}(\boldsymbol{x}) - D_i(\boldsymbol{x})\frac{\partial p(\boldsymbol{x})}{\partial x_i} - p(\boldsymbol{x})\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\right]^2}{p(\boldsymbol{x})D_i(\boldsymbol{x})} \tag{37}$$

or more concisely

$$\frac{d\langle\Delta\mathcal{S}_{\text{tot}}\rangle}{dt} = \sum_i\int d\boldsymbol{x}\;\frac{\left[J_i^{\text{ir}}(\boldsymbol{x})\right]^2}{p(\boldsymbol{x})D_i(\boldsymbol{x})}, \tag{38}$$

providing an expression for the mean instantaneous entropy production rate which is rigorously positive, as it must be because of the adherence of $\Delta\mathcal{S}_{\text{tot}}$ to an IFT, and is dependent on the *irreversible* flux.

IV. EXPRESSIONS FOR $\Delta\mathcal{S}_1$, $\Delta\mathcal{S}_2$, AND $\Delta\mathcal{S}_3$

In order to consider a division of the entropy production into the thermodynamically meaningful quantities outlined above, we are required to construct path probabilities using the so-called adjoint dynamics. These dynamics may not be physically realizable: for example, they may require negative positional steps to result from positive velocities [as indicated by the paths $\boldsymbol{x}^{\text{R}}(t)$ and $\boldsymbol{x}^{\text{T}}(t)$], but this is of no concern since they are only introduced for the mathematical construction of the entropy contributions. We consider an arbitrary stationary distribution of a given system which may be written in terms of a nonequilibrium potential, $\phi(\boldsymbol{x})$, such that

$$p^{\text{st}}(\boldsymbol{x}) = \exp[-\phi(\boldsymbol{x})] \tag{39}$$

and assert that the adjoint dynamics are those that result in the same stationary distribution, but have an opposite flux. As such, we require

$$\frac{\partial p^{\text{st}}(\boldsymbol{x})}{\partial t} = -\nabla\cdot\boldsymbol{J}^{\text{st}}(\boldsymbol{x}) = \nabla\cdot\boldsymbol{J}^{\text{st,ad}}(\boldsymbol{x}) = 0 \tag{40}$$

with

$$\boldsymbol{J}^{\text{st,ad}}(\boldsymbol{x}) = -\boldsymbol{J}^{\text{st}}(\boldsymbol{x}). \tag{41}$$

In order to characterize the adjoint dynamics, we follow Chernyak *et al.* [15] and construct the adjoint flux according to

$$\begin{aligned}J_i^{\text{st,ad}}(\boldsymbol{x}) &= A_i^{\text{ad}}(\boldsymbol{x})p^{\text{st}}(\boldsymbol{x}) - \frac{\partial}{\partial x_i}[D_i(\boldsymbol{x})p^{\text{st}}(\boldsymbol{x})]\\ &= A_i^{\text{ad}}(\boldsymbol{x})e^{-\phi(\boldsymbol{x})} - \frac{\partial}{\partial x_i}[D_i(\boldsymbol{x})e^{-\phi(\boldsymbol{x})}]\\ &= \left(A_i^{\text{ad}}(\boldsymbol{x}) - \frac{\partial D_i(\boldsymbol{x})}{\partial x_i} + D_i(\boldsymbol{x})\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}\right)e^{-\phi(\boldsymbol{x})}\\ &= -\left(A_i(\boldsymbol{x}) - \frac{\partial D_i(\boldsymbol{x})}{\partial x_i} + D_i(\boldsymbol{x})\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}\right)e^{-\phi(\boldsymbol{x})}.\end{aligned} \tag{42}$$

Consequently, we have the requirement

$$A_i^{\text{ad}}(\boldsymbol{x}) = -A_i(\boldsymbol{x}) + 2\frac{\partial D_i(\boldsymbol{x})}{\partial x_i} - 2D_i(\boldsymbol{x})\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}. \tag{43}$$

A. Expressions for $\Delta\mathcal{S}_1$

Let us now consider the quantity

$$d\Delta\mathcal{S}_{\text{ex}} = \ln\frac{\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)}{\mathcal{P}^{\text{ad}}(\boldsymbol{x},t+dt|\boldsymbol{x}',t)}, \tag{44}$$

where $\Delta\mathcal{S}_{\text{ex}} = \Delta Q_{\text{ex}}/k_B T_{\text{env}}$, which we have previously asserted constitutes part of the incremental contribution to the quantity $\Delta\mathcal{S}_1$ based on relations in Eq. (12) and its short time representation. We evaluate Eq. (44), taking the transition probability density in the numerator from Eq. (24) and, for convenience, choosing $a = 1/2$. We can represent the transition probability density appearing in the denominator through a similar construction, but using a substitution for the adjoint drift term from Eq. (43), together with the complementary evaluation point choice $b = 1 - a = 1/2$ such that

$$\mathcal{P}^{\rm ad}(\boldsymbol{x},t+dt|\boldsymbol{x}',t) = \prod_i \sqrt{\frac{1}{4\pi D_i(\boldsymbol{r})dt}} \exp\left[-\frac{(-dx_i + \{A_i(\boldsymbol{r}) - 2[\partial D_i(\boldsymbol{r})/\partial r_i] + 2D_i(\boldsymbol{r})[\partial\phi(\boldsymbol{r})/\partial r_i]\}dt + [\partial D_i(\boldsymbol{r})/\partial r_i]dt)^2}{4D_i(\boldsymbol{r})dt} + \frac{dt}{2}\frac{\partial}{\partial r_i}\left(A_i(\boldsymbol{r}) - 2\frac{\partial D_i(\boldsymbol{r})}{\partial r_i} + 2D_i(\boldsymbol{r})\frac{\partial\phi(\boldsymbol{r})}{\partial r_i}\right) + \frac{dt}{4}\frac{\partial^2 D_i(\boldsymbol{r})}{\partial r_i^2}\right]. \tag{45}$$

Since we have in both cases chosen evaluation at $a = b = 1/2$, we note that multiplication follows Stratonovich rules so that we have $f(\boldsymbol{r})dx_i = f(\boldsymbol{x}) \circ dx_i$. Considering the ratio of these two propagators, we find

$$d\Delta\mathcal{S}_{\rm ex} = \ln\frac{\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)}{\mathcal{P}^{\rm ad}(\boldsymbol{x},t+dt|\boldsymbol{x}',t)} = \sum_i D_i(\boldsymbol{x})\left(\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}\right)^2 dt + A_i(\boldsymbol{x})\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}dt - \frac{\partial\phi(\boldsymbol{x})}{\partial x_i}\circ dx_i - \frac{\partial A_i(\boldsymbol{x})}{\partial x_i}dt + \frac{\partial^2 D_i(\boldsymbol{x})}{\partial x_i^2}dt - D_i(\boldsymbol{x})\frac{\partial^2\phi(\boldsymbol{x})}{\partial x_i^2}dt - 2\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}dt. \tag{46}$$

However, we also have the condition

$$\nabla\cdot\boldsymbol{J}^{\rm st}(\boldsymbol{x}) = 0 = \sum_i \frac{\partial}{\partial x_i}\left[e^{-\phi(\boldsymbol{x})}\left(A_i(\boldsymbol{x}) - \frac{\partial D_i(\boldsymbol{x})}{\partial x_i} + D_i(\boldsymbol{x})\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}\right)\right]$$
$$= \left[-A_i(\boldsymbol{x})\frac{\partial\phi(\boldsymbol{x})}{\partial x_i} - D_i(\boldsymbol{x})\left(\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}\right)^2 + \frac{\partial A_i(\boldsymbol{x})}{\partial x_i} - \frac{\partial^2 D_i(\boldsymbol{x})}{\partial x_i^2} + D_i(\boldsymbol{x})\frac{\partial^2\phi(\boldsymbol{x})}{\partial x_i^2} + 2\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}\right]e^{-\phi(\boldsymbol{x})} \tag{47}$$

and so by insertion we arrive at

$$\ln\frac{\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)}{\mathcal{P}^{\rm ad}(\boldsymbol{x},t+dt|\boldsymbol{x}',t)} = \sum_i -\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}\circ dx_i, \tag{48}$$

which justifies the usual characterization of the adjoint dynamics [15,19,25] for use in continuous dynamics when written as

$$\frac{\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)}{\mathcal{P}^{\rm ad}(\boldsymbol{x},t+dt|\boldsymbol{x}',t)} = \frac{p^{\rm st}(\boldsymbol{x}')}{p^{\rm st}(\boldsymbol{x})} \tag{49}$$

through consideration of Eq. (39) and the Stratonovich rules which mimic normal calculus.

We construct an increment in $\Delta\mathcal{S}_1$, using the above result with the inclusion of a change in system entropy such that

$$d\Delta\mathcal{S}_1 = d\Delta\mathcal{S}_{\rm sys} + d\left(\Delta Q_{\rm ex}/k_B T_{\rm env}\right) = -d(\ln p) - \sum_i \frac{\partial\phi(\boldsymbol{x})}{\partial x_i}\circ dx_i = -\frac{1}{p(\boldsymbol{x})}\frac{\partial p(\boldsymbol{x})}{\partial t}dt - \sum_i \frac{1}{p(\boldsymbol{x})}\frac{\partial p(\boldsymbol{x})}{\partial x_i}dx_i - D_i(\boldsymbol{x})\left[\frac{1}{p(\boldsymbol{x})}\frac{\partial^2 p(\boldsymbol{x})}{\partial x_i^2} - \frac{1}{(p(\boldsymbol{x}))^2}\left(\frac{\partial p(\boldsymbol{x})}{\partial x_i}\right)^2\right]dt - \frac{\partial\phi(\boldsymbol{x})}{\partial x_i}dx_i - D_i(\boldsymbol{x})\frac{\partial^2\phi(\boldsymbol{x})}{\partial x_i^2}dt. \tag{50}$$

Applying the same averaging procedure used to calculate $\langle d\Delta\mathcal{S}_{\rm tot}\rangle$, we find

$$\langle d\Delta\mathcal{S}_1\rangle = \sum_i \int d\boldsymbol{x}\; p(\boldsymbol{x})\frac{\partial A_i(\boldsymbol{x})}{\partial x_i}dt - \frac{\partial^2[D_i(\boldsymbol{x})p(\boldsymbol{x})]}{\partial x_i^2}dt + \frac{D_i(\boldsymbol{x})}{p(\boldsymbol{x})}\left(\frac{\partial p(\boldsymbol{x})}{\partial x_i}\right)^2 dt - D_i(\boldsymbol{x})\frac{\partial^2 p(\boldsymbol{x})}{\partial x_i^2}dt - p(\boldsymbol{x})A(\boldsymbol{x})\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}dt - p(\boldsymbol{x})D(\boldsymbol{x})\frac{\partial^2\phi(\boldsymbol{x})}{\partial x_i^2}dt. \tag{51}$$

However, using Eq. (47), we may represent this as

$$\langle d\Delta\mathcal{S}_1\rangle = \sum_i \int d\boldsymbol{x}\; \frac{D_i(\boldsymbol{x})}{p(\boldsymbol{x})}\left(\frac{\partial p(\boldsymbol{x})}{\partial x_i^2}\right)^2 dt + p(\boldsymbol{x})D_i(\boldsymbol{x})\left(\frac{\partial\phi(\boldsymbol{x})}{\partial x_i^2}\right)^2 dt - 2p(\boldsymbol{x})\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\frac{\partial\phi(\boldsymbol{x})}{\partial x_i} - 2p(\boldsymbol{x})D_i(\boldsymbol{x})\frac{\partial^2\phi(\boldsymbol{x})}{\partial x_i^2}dt + p(\boldsymbol{x})\frac{\partial^2 D_i(\boldsymbol{x})}{\partial x_i^2}dt - \frac{\partial^2[D_i(\boldsymbol{x})p(\boldsymbol{x})]}{\partial x_i^2}dt - D_i(\boldsymbol{x})\frac{\partial^2 p(\boldsymbol{x})}{\partial x_i^2}dt. \tag{52}$$

By further integration by parts, dropping boundary terms, and rearranging, this becomes

$$\langle d\Delta\mathcal{S}_1\rangle = \sum_i \int d\boldsymbol{x}\; \frac{D_i(\boldsymbol{x})}{p(\boldsymbol{x})}\left(\frac{\partial p(\boldsymbol{x})}{\partial x_i^2}\right)^2 dt + p(\boldsymbol{x})D_i(\boldsymbol{x})\left(\frac{\partial\phi(\boldsymbol{x})}{\partial x_i^2}\right)^2 dt + 2D_i(\boldsymbol{x})\frac{\partial p(\boldsymbol{x})}{\partial x_i}\frac{\partial\phi(\boldsymbol{x})}{\partial x_i}dt, \tag{53}$$

which can be written

$$\frac{d\langle\Delta\mathcal{S}_1\rangle}{dt}=\frac{d\langle\Delta\mathcal{S}_{\rm sys}+(\Delta Q_{\rm ex}/k_BT_{\rm env})\rangle}{dt}$$
$$=\sum_i\int d\boldsymbol{x}\,\frac{p(\boldsymbol{x})}{D_i(\boldsymbol{x})}\left(\frac{J_i(\boldsymbol{x})}{p(\boldsymbol{x})}-\frac{J_i^{\rm st}(\boldsymbol{x})}{p^{\rm st}(\boldsymbol{x})}\right)^2$$
$$=\sum_i\int d\boldsymbol{x}\,\frac{p(\boldsymbol{x})}{D_i(\boldsymbol{x})}\left(\frac{J_i^{\rm ir}(\boldsymbol{x})}{p(\boldsymbol{x})}-\frac{J_i^{\rm st,ir}(\boldsymbol{x})}{p^{\rm st}(\boldsymbol{x})}\right)^2 \quad (54)$$

assuring the positivity of such a contribution. Since it can be written in terms of the total current J_i in this way, it maps precisely onto the nonadiabatic entropy production appearing in [21] and thus can be expressed as

$$\frac{d\langle\Delta\mathcal{S}_1\rangle}{dt}=-\int d\boldsymbol{x}\,\frac{\partial p(\boldsymbol{x})}{\partial t}\ln\frac{p(\boldsymbol{x})}{p^{\rm st}(\boldsymbol{x})} \quad (55)$$

as highlighted by the authors of [21]. We emphasize, however, that Eq. (55) is to be considered alongside the accompanying SDE in Eq. (50), from which it has been derived directly, rather than by a division of an observed positive contribution to the mean rate of change of Gibbs entropy into presumed unique transient and stationary terms.

B. Expressions for $\Delta\mathcal{S}_2$

We may now by similar means consider an increment in $\Delta\mathcal{S}_2$ as follows:

$$d\Delta\mathcal{S}_2=\ln\frac{\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)}{\mathcal{P}^{\rm ad}(\boldsymbol{\varepsilon x}',t+dt|\boldsymbol{\varepsilon x},t)}. \quad (56)$$

In this case, the construction of the denominator follows slightly different rules since, unlike $\Delta\mathcal{S}_{\rm tot}$ and $\Delta\mathcal{S}_1$, the alternative path, $\vec{\boldsymbol{x}}^{\rm T}$, is based on a time reversal of the coordinates, but otherwise follows the sequence of the forward path. As such, b behaves in the same manner as a rendering $r_i'=bx_i'+(1-b)x_i$, $f(\boldsymbol{r}')dx_i=f(\boldsymbol{x})dx_i+2bD(\boldsymbol{x})\partial f(\boldsymbol{x})/\partial x_i dt$, as detailed in Appendix A. In this case, the appropriate choice for the equivalence of evaluation points $\boldsymbol{r}'=\boldsymbol{r}$ is $a=b$. For continuity, we may once again choose $a=b=1/2$ with Stratonovich multiplication rules: we represent the transition probability appearing in the numerator through Eq. (24) and the denominator by a similar means using the drift term given in Eq. (43) and the path choice $\boldsymbol{x}^{\rm T}(t)=\boldsymbol{\varepsilon x}(t)$, such that

$$\mathcal{P}^{\rm ad}(\boldsymbol{\varepsilon x}',t+dt|\boldsymbol{\varepsilon x},t)$$
$$=\prod_i\sqrt{\frac{1}{4\pi D_i(\boldsymbol{\varepsilon r})dt}}\exp\Bigg[-\frac{(\varepsilon_i dx_i+\{A_i(\boldsymbol{\varepsilon r})-2[\partial D_i(\boldsymbol{\varepsilon r})/\partial(\varepsilon_i r_i)]+2D_i(\boldsymbol{\varepsilon r})[\partial\phi(\boldsymbol{\varepsilon r})/\partial(\varepsilon_i r_i)]\}dt+[\partial D_i(\boldsymbol{\varepsilon r})/\partial(\varepsilon_i r_i)]dt)^2}{4D_i(\boldsymbol{\varepsilon r})dt}$$
$$+\frac{dt}{2}\frac{\partial}{\partial(\varepsilon_i r_i)}\left(A_i(\boldsymbol{\varepsilon r})-2\frac{\partial D_i(\boldsymbol{\varepsilon r})}{\partial(\varepsilon_i r_i)}+2D_i(\boldsymbol{\varepsilon r})\frac{\partial\phi(\boldsymbol{\varepsilon r})}{\partial(\varepsilon_i r_i)}\right)+\frac{dt}{4}\frac{\partial^2 D_i(\boldsymbol{\varepsilon r})}{\partial(\varepsilon_i r_i)^2}\Bigg]. \quad (57)$$

We can utilize the usual transformation rules and assumptions for $A^{\rm ir}$, $A^{\rm rev}$, and D_i and express $\partial\phi(\boldsymbol{\varepsilon r})/\partial(\varepsilon_i r_i)=\varepsilon_i\partial\phi(\boldsymbol{\varepsilon r})/\partial r_i=\phi_i'(\boldsymbol{\varepsilon r})$ [along with $\partial^2\phi(\boldsymbol{\varepsilon r})/\partial(\varepsilon_i r_i)^2=\phi_i''(\boldsymbol{\varepsilon x})$] such that we can write the propagator as

$$\mathcal{P}^{\rm ad}(\boldsymbol{\varepsilon x}',t+dt|\boldsymbol{\varepsilon x},t)$$
$$=\prod_i\sqrt{\frac{1}{4\pi D_i(\boldsymbol{r})dt}}\exp\Bigg[-\frac{\left(dx_i+\left\{A_i^{\rm ir}(\boldsymbol{r})-A_i^{\rm rev}(\boldsymbol{r})-2[\partial D_i(\boldsymbol{r})/\partial r_i]+2\varepsilon_i D_i(\boldsymbol{r})\phi_i'(\boldsymbol{\varepsilon r})\right\}dt+[\partial D_i(\boldsymbol{r})/\partial r_i]dt\right)^2}{4D_i(\boldsymbol{r})dt}$$
$$+\frac{dt}{2}\frac{\partial}{\partial r_i}\left(A_i^{\rm ir}(\boldsymbol{r})-A_i^{\rm rev}(\boldsymbol{r})-2\frac{\partial D_i(\boldsymbol{r})}{\partial r_i}+2\varepsilon_i D_i(\boldsymbol{r})\phi_i'(\boldsymbol{\varepsilon r})\right)+\frac{dt}{4}\frac{\partial^2 D_i(\boldsymbol{r})}{\partial r_i^2}\Bigg]. \quad (58)$$

Constructing the ratio in Eq. (56), we find

$$d\Delta\mathcal{S}_2=\sum_i-\frac{A_i^{\rm ir}(\boldsymbol{x})A_i^{\rm rev}(\boldsymbol{x})}{D_i(\boldsymbol{x})}dt+\frac{A_i^{\rm ir}(\boldsymbol{x})}{D_i(\boldsymbol{x})}\circ dx_i-\frac{\partial A_i^{\rm ir}(\boldsymbol{x})}{\partial x_i}dt+D_i(\boldsymbol{x})[\phi_i'(\boldsymbol{\varepsilon x})]^2dt-2\varepsilon_i\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\phi_i'(\boldsymbol{\varepsilon x})dt$$
$$+\varepsilon_i\left[A_i^{\rm ir}(\boldsymbol{x})-A_i^{\rm rev}(\boldsymbol{x})\right]\phi_i'(\boldsymbol{\varepsilon x})dt+\frac{A_i^{\rm rev}(\boldsymbol{x})}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}dt+\varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})\circ dx_i-\frac{1}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x}\circ dx_i$$
$$+\frac{\partial^2 D_i(\boldsymbol{x})}{\partial x_i^2}dt-D_i(\boldsymbol{x})\phi_i''(\boldsymbol{\varepsilon x})dt, \quad (59)$$

which is the same as

$$d\Delta\mathcal{S}_2=d(\Delta Q_{\rm hk,G}/k_BT_{\rm env})=\sum_i-\frac{A_i^{\rm ir}(\boldsymbol{x})A_i^{\rm rev}(\boldsymbol{x})}{D_i(\boldsymbol{x})}dt+\frac{A_i^{\rm ir}(\boldsymbol{x})}{D_i(\boldsymbol{x})}dx_i+\varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})dx_i-\frac{1}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x}dx_i+\frac{1}{D_i(\boldsymbol{x})}\left(\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\right)^2dt$$
$$+D_i(\boldsymbol{x})[\phi_i'(\boldsymbol{\varepsilon x})]^2dt-2\varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})\frac{\partial D_i(\boldsymbol{x})}{\partial x}dt+\varepsilon_i\left[A_i^{\rm ir}(\boldsymbol{x})-A_i^{\rm rev}(\boldsymbol{x})\right]\phi_i'(\boldsymbol{\varepsilon x})dt-\frac{\left[A_i^{\rm ir}(\boldsymbol{x})-A_i^{\rm rev}(\boldsymbol{x})\right]}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}dt. \quad (60)$$

By employing the averaging procedure, we find

$$\frac{d\langle\Delta\mathcal{S}_2\rangle}{dt} = \sum_i \int d\boldsymbol{x}\ \frac{p(\boldsymbol{x})}{D_i(\boldsymbol{x})}\left(A_i^{\rm ir}(\boldsymbol{x}) - \frac{\partial D_i(\boldsymbol{x})}{\partial x_i} + \varepsilon_i D_i(\boldsymbol{x})\phi_i'(\boldsymbol{\varepsilon x})\right)^2, \tag{61}$$

which may be written

$$\frac{d\langle\Delta\mathcal{S}_2\rangle}{dt} = \frac{d\langle\Delta Q_{\rm hk,G}/k_B T_{\rm env}\rangle}{dt} = \sum_i \int d\boldsymbol{x}\ \frac{p(\boldsymbol{x})}{D_i(\boldsymbol{\varepsilon x})}\left(\frac{J_i^{\rm ir,st}(\boldsymbol{\varepsilon x})}{p^{\rm st}(\boldsymbol{\varepsilon x})}\right)^2. \tag{62}$$

Such a form illustrates the positivity requirement of $\Delta\mathcal{S}_2$ in the mean, resulting from its adherence to an IFT, and again Eq. (62) is to be considered alongside the complementary SDE in Eq. (60). Since it is based on an integral over the stationary irreversible flux, $d\langle\Delta\mathcal{S}_2\rangle/dt$ describes a contribution to entropy production which arises from an absence of detailed balance and is nonzero both in and out of stationarity. This quantity is to be contrasted with the adiabatic entropy production in [21], which we may now consider to be a special case when there are only even variables in the dynamics. We point out again the importance of the direct derivation of this result from the SDE in this formalism, as opposed to a division of the irreversible flux into terms with structure based solely on $p^{\rm st}(\boldsymbol{x})$, which would not obviously have led to the above expression.

We note that the integral in Eq. (62) must reduce to the total entropy production, and thus an integral over the stationary irreversible flux (i.e., without the $\boldsymbol{\varepsilon}$ factors inside the squared term in the integrand), in the stationary state, but there are other circumstances when this correspondence applies more generally. A first case is when the irreversible stationary flux is proportional to the stationary distribution, which would be the case for a nonequilibrium constraint that is independent of the phase-space variables, as illustrated later in example II, and a second case is when the total flux in each coordinate is everywhere zero $[J_i(\boldsymbol{x}) = 0]$, such as for independent variables, x_i, defined on regions with natural or reflecting boundaries.

C. Expressions for $\Delta\mathcal{S}_3$

To complete the description of all three contributions to entropy production, we now consider an increment in $\Delta\mathcal{S}_3$. By using the definition in Eq. (14),

$$d\Delta\mathcal{S}_3 = \ln\frac{\mathcal{P}^{\rm ad}(\boldsymbol{\varepsilon x}',t+dt|\boldsymbol{\varepsilon x},t)\mathcal{P}^{\rm ad}(\boldsymbol{x},t+dt|\boldsymbol{x}',t)}{\mathcal{P}(\boldsymbol{x}',t+dt|\boldsymbol{x},t)\mathcal{P}(\boldsymbol{\varepsilon x},t+dt|\boldsymbol{\varepsilon x}',t)}, \tag{63}$$

together with the previously used propagators, and employing the stationarity condition evaluated at $\boldsymbol{\varepsilon x}$,

$$\nabla\cdot\boldsymbol{J}^{\rm st}(\boldsymbol{\varepsilon x}) = 0 = \sum_i\Bigg(-\big[A_i^{\rm ir}(\boldsymbol{\varepsilon x}) + A_i^{\rm rev}(\boldsymbol{\varepsilon x})\big]\phi_i'(\boldsymbol{\varepsilon x}) + \frac{\partial A_i^{\rm ir}(\boldsymbol{\varepsilon x})}{\partial(\varepsilon_i x_i)} + \frac{\partial A_i^{\rm rev}(\boldsymbol{\varepsilon x})}{\partial(\varepsilon_i x_i)} - D_i(\boldsymbol{\varepsilon x})[\phi_i'(\boldsymbol{\varepsilon x})]^2 - \frac{\partial^2 D_i(\boldsymbol{\varepsilon x})}{\partial(\varepsilon_i x_i)^2} + D_i(\boldsymbol{\varepsilon x})\phi_i''(\boldsymbol{\varepsilon x}) + 2\frac{\partial D_i(\boldsymbol{\varepsilon x})}{\partial(\varepsilon_i x_i)}\phi_i'(\boldsymbol{\varepsilon x})\Bigg)e^{-\phi(\boldsymbol{\varepsilon x})}, \tag{64}$$

we find

$$d\Delta\mathcal{S}_3 = d(\Delta Q_{\rm hk,T}/k_B T_{\rm env}) = \sum_i \phi_i'(\boldsymbol{x})\circ dx_i - \varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})\circ dx_i = \sum_i \ln\frac{\exp[-\phi(\boldsymbol{x})]}{\exp[-\phi(\boldsymbol{x}')]}\frac{\exp[-\phi(\boldsymbol{\varepsilon x}')]}{\exp[-\phi(\boldsymbol{\varepsilon x})]}, \tag{65}$$

which maps onto the same quantity derived from a master equation approach [23]. We can then construct the average contribution by converting to Ito form and performing the path integral such that

$$\langle d\Delta\mathcal{S}_3\rangle = \sum_i\int d\boldsymbol{x}\ p(\boldsymbol{x})A_i(\boldsymbol{x})[\phi_i'(\boldsymbol{x}) - \varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})]dt + p(\boldsymbol{x})D_i(\boldsymbol{x})[\phi_i''(\boldsymbol{x}) - \phi_i''(\boldsymbol{\varepsilon x})]dt \tag{66}$$

and proceed to manipulate by integrating by parts, assuming the probability density and current vanish or cancel at boundaries, such that

$$\begin{aligned}\langle d\Delta\mathcal{S}_3\rangle &= \sum_i\int d\boldsymbol{x}\ p(\boldsymbol{x})A_i(\boldsymbol{x})[\phi_i'(\boldsymbol{x}) - \varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})]dt - \int d\boldsymbol{x}\ \frac{\partial}{\partial x_i}[p(\boldsymbol{x})D_i(\boldsymbol{x})]\,[\phi_i'(\boldsymbol{x}) - \varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})]dt\\ &= \sum_i\int d\boldsymbol{x}\ [\phi_i'(\boldsymbol{x}) - \varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})]\left(A_i(\boldsymbol{x})p(\boldsymbol{x}) - \frac{\partial}{\partial x_i}[p(\boldsymbol{x})D_i(\boldsymbol{x})]\right)dt\\ &= \sum_i\int d\boldsymbol{x}\ [\phi_i'(\boldsymbol{x}) - \varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})]J_i(\boldsymbol{x})dt = \sum_i -\int d\boldsymbol{x}\ [\phi(\boldsymbol{x}) - \phi(\boldsymbol{\varepsilon x})]\frac{\partial J_i(\boldsymbol{x})}{\partial x_i}dt\\ &= -\int d\boldsymbol{x}\ [\phi(\boldsymbol{x}) - \phi(\boldsymbol{\varepsilon x})]\left(\sum_i\frac{\partial J_i(\boldsymbol{x})}{\partial x_i}\right)dt = -\int d\boldsymbol{x}\ [\phi(\boldsymbol{x}) - \phi(\boldsymbol{\varepsilon x})]\,[\nabla\cdot\boldsymbol{J}(\boldsymbol{x})]\,dt.\end{aligned} \tag{67}$$

By substituting the original Fokker-Planck equation, we may also write this as

$$\begin{aligned}\frac{d\langle\Delta\mathcal{S}_3\rangle}{dt} &= \frac{d\langle\Delta Q_{\mathrm{hk,T}}/k_B T_{\mathrm{env}}\rangle}{dt}\\ &= \int d\boldsymbol{x}\,\frac{\partial p(\boldsymbol{x})}{\partial t}\,[\phi(\boldsymbol{x})-\phi(\boldsymbol{\varepsilon x})]\\ &= -\int d\boldsymbol{x}\,\frac{\partial p(\boldsymbol{x})}{\partial t}\ln\frac{p^{\mathrm{st}}(\boldsymbol{x})}{p^{\mathrm{st}}(\boldsymbol{\varepsilon x})}. \end{aligned}\tag{68}$$

This has a form similar to Eq. (55) and is clearly a contribution to the mean total entropy production rate that behaves transiently in a manner similar to $\Delta\mathcal{S}_1$, but explicitly vanishes when there are no odd variables in the dynamics. The quantity $\Delta\mathcal{S}_1$ appears in the Hatano-Sasa relation which describes the entropy production associated with a transition between different stationary states. However, in light of Eq. (68) we suggest that $\Delta\mathcal{S}_1$, and thus the Hatano-Sasa relation and nonadiabatic entropy production, do not represent the entire entropy production associated with transitions between stationary states (or more generally relaxation) since, in the mean, we can construct a new quantity that comprises all contributions which are nonzero only during relaxation, by combining Eqs. (55) and (68) giving

$$\begin{aligned}\frac{d\langle\Delta\mathcal{S}_1+\Delta\mathcal{S}_3\rangle}{dt} &= \frac{d\langle\Delta\mathcal{S}_{\mathrm{sys}}+(\Delta Q_{\mathrm{ex}}+\Delta Q_{\mathrm{hk,T}})/k_B T_{\mathrm{env}}\rangle}{dt}\\ &= -\int d\boldsymbol{x}\,\frac{\partial p(\boldsymbol{x})}{\partial t}\ln\frac{p(\boldsymbol{x})}{p^{\mathrm{st}}(\boldsymbol{\varepsilon x})}.\end{aligned}\tag{69}$$

This describes a contribution to the mean entropy production rate which occurs when the system is out of stationarity, but it does not obey an IFT and thus has no guarantee of positivity.

Our central results, therefore, are expressions for three contributions to entropy production for arbitrary systems with odd and even dynamical variables evolving according to Ito SDEs with multiplicative noise. These expressions apply to individual trajectories Eqs. (32), (50), (60), and (65) and in the mean Eqs. (38), (54), (62), and (68). Such a demonstration shows the additional complexity introduced by the inclusion of odd variables if one insists on considering entropy production to be due to relaxation or to nonequilibrium constraints with particular reference to Eq. (69). One may think of $\langle\Delta\mathcal{S}_1+\Delta\mathcal{S}_3\rangle$ as describing a transient contribution to entropy production in the same manner as $\langle\Delta\mathcal{S}_1\rangle$, but with the further specification of the nature of the coordinates: the entropy production depends on whether the variables being described are odd or even. The additional complexity of $\Delta\mathcal{S}_3$ arises because Eq. (69) can only differ from Eq. (55) when the stationary state is out of equilibrium, such that $p^{\mathrm{st}}(\boldsymbol{x})\neq p^{\mathrm{st}}(\boldsymbol{\varepsilon x})$.

We of course expect and require that the contributions detailed here are related such that

$$\frac{d\langle\Delta\mathcal{S}_{\mathrm{tot}}\rangle}{dt}=\frac{d\langle\Delta\mathcal{S}_1\rangle}{dt}+\frac{d\langle\Delta\mathcal{S}_2\rangle}{dt}+\frac{d\langle\Delta\mathcal{S}_3\rangle}{dt}\tag{70}$$

yet their forms derived above do not obviously lend themselves to such a demonstration immediately. For completeness, this is shown in Appendix B.

V. EXAMPLE I: STATIONARY HEAT TRANSPORT

We provide as a first example of usage of the above formalism a physical situation which necessitates the use of odd variables in order to describe entropy production adequately: heat transport due to diffusion in one spatial dimension in the presence of a spatially dependent temperature field. Mathematically this system may be modeled without odd (velocity) variables by employing the overdamped limit and constructing a multiplicative SDE and Fokker-Planck equation of the form

$$dx=\frac{F(x)}{m\gamma}dt+\sqrt{\frac{2k_BT(x)}{m\gamma}}dW\tag{71}$$

and

$$\frac{\partial p(x,t)}{\partial t}=-\frac{\partial}{\partial x}\left(\frac{F(x)p(x,t)}{m\gamma}\right)+\frac{\partial^2}{\partial x^2}\left(\frac{k_BT(x)p(x,t)}{m\gamma}\right),\tag{72}$$

where m is the particle mass, γ is the damping coefficient, and $F(x)$ is the force operating on the particle. We note the Ito form of both (for a discussion of the resolution of the Ito-Stratonovich dilemma in this case, see, for example, [32–34]). This Fokker-Planck equation has a stationary distribution

$$p^{\mathrm{st}}(x)=\frac{\mathcal{N}m}{k_BT(x)}\exp\left[\int_0^x dx'\,\frac{F(x')}{k_BT(x')}\right],\tag{73}$$

where $\mathcal{N}$ is a normalization constant. We can quite readily identify the terms $A_x^{\mathrm{ir}}=F(x)/m\gamma$, $A_x^{\mathrm{rev}}=0$, and $D_x(x)=k_BT(x)/m\gamma$. However, when we come to construct the dimensionless entropy production in the stationary state from Eq. (32) as

$$\begin{aligned} d\Delta\mathcal{S}_{\mathrm{tot}} &= \frac{A_x^{\mathrm{ir}}(x)}{D_x(x)}\circ dx-\frac{1}{D_x(x)}\frac{\partial D_x(x)}{\partial x}\circ dx-\frac{1}{p^{\mathrm{st}}(x)}\frac{\partial p^{\mathrm{st}}(x)}{\partial x}\circ dx\\ &= \left[\frac{F(x)}{k_BT(x)}-\frac{1}{T(x)}\frac{\partial T(x)}{\partial x}-\frac{1}{p^{\mathrm{st}}(x)}\left(-\frac{1}{T(x)}\frac{\partial T(x)}{\partial x}p^{\mathrm{st}}(x)\right.\right.\\ &\quad\left.\left.+\frac{F(x)}{k_BT(x)}p^{\mathrm{st}}(x)\right)\right]\circ dx\\ &= 0,\end{aligned}\tag{74}$$

we find that there is zero entropy production for all trajectories. This may be understood either physically by recognizing that in the overdamped limit one demands that the velocity distribution relaxes instantaneously, thereby preventing any heat transfer due to temperature inhomogeneities, or geometrically by recognizing the impossibility of having stationary flow, and thus entropy production, for a system in one dimension with natural boundaries.

To provide a satisfactory representation and to understand the entropy production in such a system, we need to consider the more realistic underdamped dynamics in full phase space where we retain both position and velocity coordinates, x and v, which are even and odd under time reversal, respectively.

The SDEs and Fokker-Planck equation are now given as

$$dx = v\,dt,$$
$$dv = -\gamma v\,dt + \frac{F(x)}{m}dt + \sqrt{\frac{2k_B T(x)\gamma}{m}}dW, \tag{75}$$

and

$$\frac{\partial p(x,v,t)}{\partial t} = -v\frac{\partial p(x,v,t)}{\partial x} - \frac{\partial}{\partial v}\left[\left(\frac{F(x)}{m} - \gamma v\right)p(x,v,t)\right] + \frac{k_B T(x)\gamma}{m}\frac{\partial^2 p(x,v,t)}{\partial v^2}. \tag{76}$$

We may then identify the terms $A_x^{\text{ir}} = 0$, $A_x^{\text{rev}} = v$, $A_v^{\text{ir}} = -\gamma v$, $A_v^{\text{rev}} = F(x)/m$, $D_x = 0$, and $D_v = k_B T(x)\gamma/m$. By Eq. (32), the dimensionless entropy production is

$$\begin{aligned} d\Delta\mathcal{S}_{\text{tot}} &= -d[\ln p(x,v,t)] - \frac{mv}{k_B T(x)}\circ dv + \frac{Fv}{k_B T(x)}dt \\ &= -d[\ln p(x,v,t)] - \frac{1}{k_B T(x)}d\left(\frac{mv^2}{2}\right) + \frac{F}{k_B T(x)}dx \end{aligned} \tag{77}$$

using $v\circ dv = (1/2)(v'+v)(v'-v)$ and $v\,dt = dx$, and noting that x is now deterministic, meaning the integration rules are irrelevant. The second and third terms correctly reproduce the form of the change in medium entropy as heat transfer to the environment, equal to negative heat transfer to the particle (in agreement with the result found in stochastic energetics [35]), divided by the instantaneous temperature, and do so only by virtue of the consideration of odd and even variables. We can use this SDE to produce distributions of entropy production and verify relevant fluctuation theorems. To do so, however, requires knowledge of the solution to the Fokker-Planck equation, for which there is no simple analytical form. To proceed, we restrict ourselves to the stationary state and utilize the expansion found in [34] and [36], which expresses the stationary solution as a series expansion about the overdamped distribution:

$$p^{\text{st,over}}(x,v) = \frac{\mathcal{N}m}{k_B T(x)}\exp\left[\int_0^x dx'\,\frac{F(x')}{k_B T(x')}\right] \times\sqrt{\frac{m}{2\pi k_B T(x)}}\exp\left[-\frac{mv^2}{2k_B T(x)}\right], \tag{78}$$

where $\mathcal{N}$ is determined by normalization, such that

$$p^{\text{st}}(x,v) = p^{\text{st,over}}(x,v) + \sum_{i=1}^{\infty}(1/\gamma)^i p_i(x,v), \tag{79}$$

$p_i(x,v)$ has a general form

$$p_i(x,v) = \sum_{k=a_i}^{k=b_i}\frac{c_{i,k}(x)H_k[v\sqrt{m/k_B T(x)}]}{\sqrt{2\pi k_B T(x)/m}}\exp\left[-\frac{mv^2}{2k_B T(x)}\right], \tag{80}$$

where constants a_i,b_i and functions $c_{i,k}(x)$ are found by an iterative procedure, and $H_k(y)$ are Hermite polynomials defined as

$$H_k(y) = (-1)^k e^{\frac{y^2}{2}}\frac{d^k}{dy^k}e^{\frac{-y^2}{2}}. \tag{81}$$

While the expansion has the formal deficiency that the expansion parameter is not unitless, it suffices for a theoretical illustration in which we can consider it in a limit where it is appropriate. We consider units $k_B = T = m = 1$, a harmonic confining potential such that $F(x) = -x$, a temperature profile

$$T(x) = 1 + \tfrac{1}{2}\tanh(x), \tag{82}$$

and we approximate the stationary distribution by considering the expansion in Eq. (79) to fourth order in γ^{-1}, applying the formalism numerically.

We first demonstrate that this approach yields a result which maps onto the expected phenomenological expression for dimensionless internal entropy generation [37],

$$\frac{d\langle\Delta\mathcal{S}_{\text{tot}}\rangle}{dt} = \int_{-\infty}^{+\infty}dx\;J_Q(x)\frac{\partial}{\partial x}\left(\frac{1}{k_B T(x)}\right), \tag{83}$$

where $J_Q(x)$ is the stationary heat current defined as

$$J_Q(x) = \int_{-\infty}^{+\infty}dv\;\frac{1}{2}mv^3 p^{\text{st}}(x,v). \tag{84}$$

Figure 1 shows the dimensionless entropy production obtained by performing the integral in Eq. (83) using a numerically calculated $p^{\text{st}}(x,v)$, compared with that obtained by averaging the SDE in Eq. (77) by Monte Carlo simulation of the underlying particle dynamics, for a range of damping coefficients, alongside a demonstrably positive first-order approximation based on the first correction term in Eq. (79)

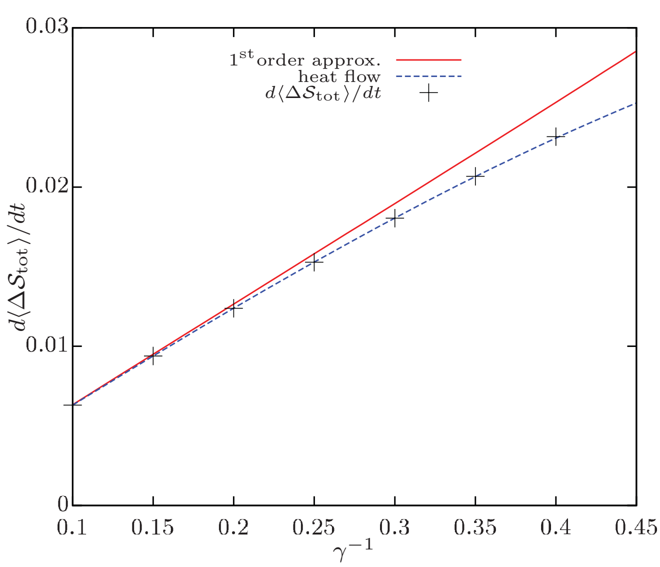

FIG. 1. (Color online) Mean dimensionless entropy production for example I for a range of damping coefficients as predicted by a first-order approximation in Eq. (85) (solid line), an integral over the heat current, Eq. (83) (dashed line), and a Monte Carlo average based on the SDE in Eq. (77) (crosses). Simulations were performed by initialization of particles into the stationary distribution using a simple reject/accept algorithm along with a burn-in time of $t = 10$. We performed 3×10^7 Monte Carlo runs utilizing a forward Euler discretization method with time step $dt = 1.0\times10^{-3}$ to solve the SDE in Eq. (77). The same method was utilized for all subsequent figures in example I.

given by [34]

$$\frac{d\langle\Delta\mathcal{S}_{\rm tot}\rangle}{dt} \simeq \int_{-\infty}^{+\infty} dx\, \frac{k_B p^{\rm st}(x)}{2m\gamma T(x)} \left(\frac{\partial T(x)}{\partial x}\right)^2, \quad (85)$$

where

$$p^{\rm st}(x) = \int_{-\infty}^{+\infty} dv\; p^{\rm st}(x,v). \quad (86)$$

Our formalism for the rate of change of $\langle\Delta\mathcal{S}_{\rm tot}\rangle$ agrees with Eq. (83) and both are consistent with Eq. (85) in the $\gamma \to \infty$ limit where the system's proximity to local equilibrium brings it within the scope of linear irreversible thermodynamics as evidenced by the dependence on the square of the temperature gradient.

We point out that total entropy production decreases as coupling to the environment increases, which may seem counterintuitive, but we emphasize that with increased coupling, despite greater heat transfer to and from the environment, there is highly diminished spatial heat *transport* (the latter being the cause of entropy production) as the system is brought closer to a local equilibrium.

We can use the SDE for entropy production Eq. (77) to move beyond a classical description of mean entropy production to one described by Jarzynski, Seifert, Sekimoto, and others [8,26,38], where we can identify entropy generating and destroying trajectories. We can explicitly calculate the distribution of total entropy production which is shown for $\gamma = 10$ in Fig. 2 for various process intervals, along with a demonstration that it adheres to an IFT throughout. Additionally, since we consider the stationary state, we can demonstrate a detailed fluctuation theorem of the form $p(\Delta\mathcal{S}_{\rm tot})/p(-\Delta\mathcal{S}_{\rm tot}) = \exp(\Delta\mathcal{S}_{\rm tot})$ [11] as shown in Fig. 3.

Finally we point out that, being in the stationary state, $d\langle\Delta\mathcal{S}_3\rangle/dt = 0$, but since it is a nonequilibrium stationary state that is asymmetric in the odd velocity variable, we have $\Delta\mathcal{S}_3 \neq 0$ in detail, as is clear in Eq. (65). We can demonstrate the increasing range of values of $\Delta\mathcal{S}_3$ as γ is reduced and the system is taken further away from local equilibrium, with its symmetric velocity distribution, by generating the distribution of $\Delta\mathcal{S}_3$ using Eq. (65) for a given time interval, as shown in Fig. 4. Such a result highlights the fact that although a nonzero $d\langle\Delta\mathcal{S}_3\rangle/dt$ is only possible during relaxation, as shown by Eq. (68), the specific evolution of $\Delta\mathcal{S}_3$ for each trajectory is brought about by nonequilibrium constraints that cause the stationary solution to depart from equilibrium.

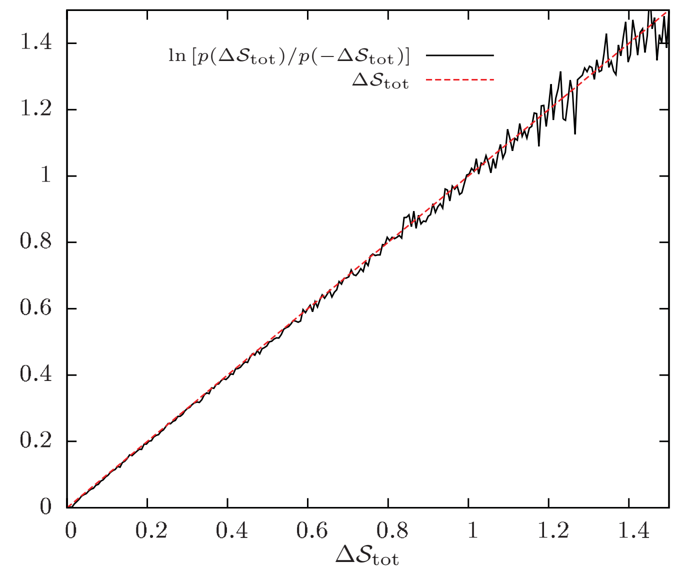

FIG. 3. (Color online) Verification of a detailed fluctuation theorem for example I using data from simulation for $\gamma = 10$ at time $t = 8$.

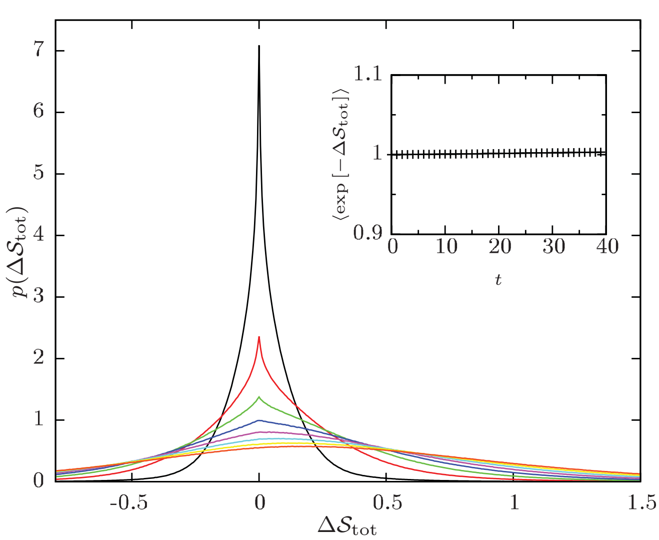

FIG. 2. (Color online) Distributions of dimensionless total entropy production $\Delta\mathcal{S}_{\rm tot}$ for example I for $\gamma = 10$ together with a demonstration of adherence to an IFT. Distributions shown are for process intervals from $t = 2$ (narrowest) to $t = 44$ (widest) in steps of six units.

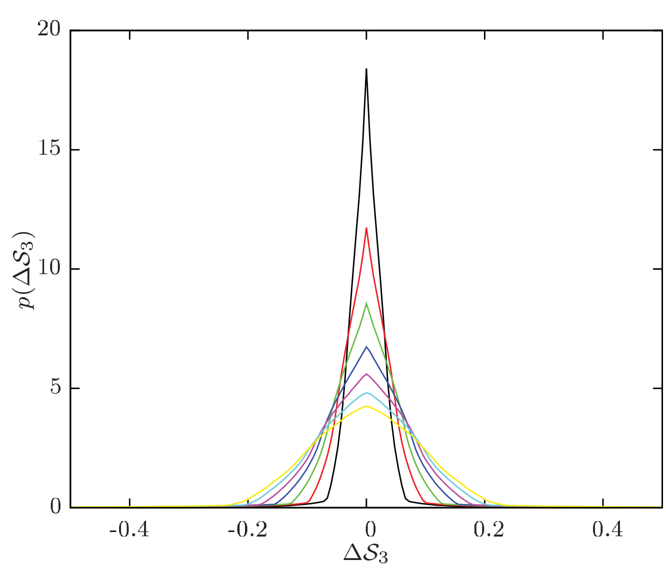

FIG. 4. (Color online) Distributions of $\Delta\mathcal{S}_3$ for example I evaluated at $t = 8$ for a range of γ from $\gamma^{-1} = 0.1$ (narrowest) to $\gamma^{-1} = 0.4$ (widest).

VI. EXAMPLE II: PARTICLE DRIVEN BY A NONCONSERVATIVE FORCE

Once again utilizing the full phase-space Langevin description of the dynamics, we consider diffusion of a particle on a ring driven by a spatially independent nonconservative force and spatially independent (additive) noise such that

$$dx = v\,dt, \qquad dv = -\gamma v\,dt + \frac{F(t)}{m}dt + \sqrt{\frac{2k_B T\gamma}{m}}dW, \tag{87}$$

thus giving $A_x^{\rm ir}=0$, $A_x^{\rm rev}=v$, $A_v^{\rm ir}=-\gamma v$, $A_v^{\rm rev}=F(t)/m$, $D_x=0$, and $D_v=k_BT\gamma/m$. For any nonzero value of $F(t)$, there will exist a stationary solution with an asymmetric Gaussian distribution in v and a uniform distribution in x due to the symmetry of the problem. Any relaxation from a given stationary state caused by changes to the nonconservative force will then also result in a uniform distribution in x for all time by the translational symmetry. As such, we may proceed by considering the marginalized velocity distribution when starting from a stationary state. Exploiting the fact that the initial Gaussian solution will remain Gaussian for any $F(t)$, we can parametrize a transient solution to the Fokker-Planck equation

$$p(x,v,t) \propto \sqrt{\frac{m}{2\pi k_B T}}\exp\left[-\frac{m(v-\langle v\rangle)^2}{2k_BT}\right] \tag{88}$$

with

$$\frac{d\langle v\rangle}{dt} = \left(\frac{F}{m} - \gamma\langle v\rangle\right) \tag{89}$$

such that

$$\langle v\rangle^{\rm st} = \frac{F}{m\gamma}. \tag{90}$$

A scenario in which closed-form solutions exist for all contributions to entropy production is that of an instantaneous step change in the driving force $F(t)$, so that we have

$$F(t) = \begin{cases} F_0, & t < t_0, \\ F_1, & t \geqslant t_0 \end{cases} \tag{91}$$

and

$$\langle v\rangle(t) = \begin{cases} F_0/m\gamma, & t < t_0, \\ [F_1 + e^{-\gamma(t-t_0)}(F_0-F_1)]/m\gamma, & t \geqslant t_0. \end{cases} \tag{92}$$

Performing the relevant integrals in Eqs. (38), (54), (62), and (68), we then obtain

$$\frac{d\langle\Delta\mathcal{S}_{\rm tot}\rangle}{dt} = \begin{cases} F_0^2/m\gamma k_BT, & t < t_0, \\ [F_0 + F_1(e^{\gamma(t-t_0)}-1)]^2 & \\ \times e^{-2\gamma(t-t_0)}/m\gamma k_BT, & t \geqslant t_0, \end{cases} \tag{93}$$

$$\frac{d\langle\Delta\mathcal{S}_1\rangle}{dt} = \begin{cases} 0, & t < t_0, \\ e^{-2\gamma(t-t_0)}(F_0-F_1)^2/m\gamma k_BT, & t \geqslant t_0, \end{cases} \tag{94}$$

$$\frac{d\langle\Delta\mathcal{S}_2\rangle}{dt} = \begin{cases} F_0^2/m\gamma k_BT, & t < t_0, \\ F_1^2/m\gamma k_BT, & t \geqslant t_0, \end{cases} \tag{95}$$

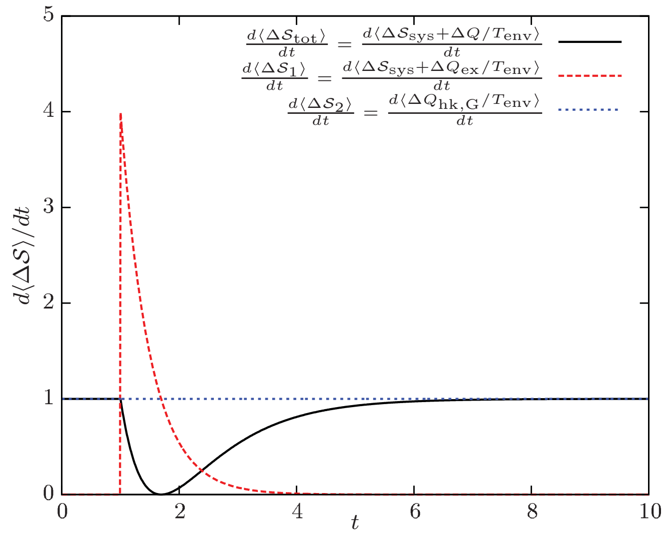

FIG. 5. (Color online) Positive mean rates of dimensionless entropy change against time for example II, where we consider the transition between stationary states of a driven particle on a ring with $F_0=1$, $F_1=-1$, $t_0=1$, and $k_B=m=\gamma=T=1$.

and

$$\frac{d\langle\Delta\mathcal{S}_3\rangle}{dt} = \begin{cases} 0, & t < t_0, \\ -2e^{-\gamma(t-t_0)}F_1(F_1-F_0)/m\gamma k_BT, & t \geqslant t_0. \end{cases} \tag{96}$$

Choosing the specific case of a reversal of the driving force such that it changes from $F_0=1$ to $F_1=-1$ at time $t_0=1$ and employing units $k_B=m=\gamma=T=1$, we can generate the results shown in Figs. 5 and 6.

We note first that the mean rates of change of all three contributions $\Delta\mathcal{S}_{\rm tot}$, $\Delta\mathcal{S}_1$, and $\Delta\mathcal{S}_2$ are positive, reflecting their adherence to an IFT. All three mean rates of change are constant for $t < t_0 = 1$, are perturbed by the change in direction of the force, and relax back to constant values

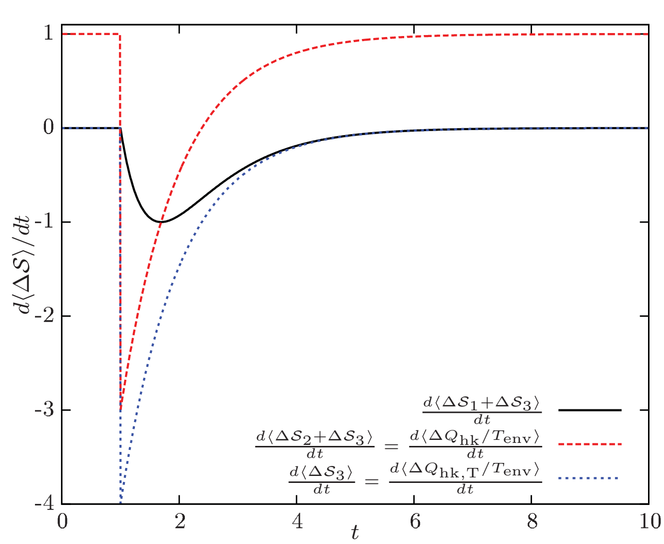

FIG. 6. (Color online) Unbounded mean rates of dimensionless entropy change for example II, the driven system on a ring with $F_0=1$, $F_1=-1$, $t_0=1$, and $k_B=m=\gamma=T=1$.

consistent with the transition between the stationary states. A key feature of this behavior is that upon perturbation, the total entropy production rate *decreases*, which would not emerge using an overdamped description of the dynamics. This feature can be explained by the existence of the $d\langle\Delta\mathcal{S}_3\rangle/dt$ contribution to the mean entropy production rate, which may take negative values depending on the relationship between the instantaneous distribution and the stationary distribution. In this specific case, the large negative value for $d\langle\Delta\mathcal{S}_3\rangle/dt$ indicates that upon reversal of the force, the instantaneous distribution corresponds to particle motion, on average, in a direction counter to that expected to result from the new value of the force. The velocity distribution does relax, of course, to the distribution that corresponds to the new value of the force and so the mean rate of change of $\Delta\mathcal{S}_3$ decays away. An important point to draw from Fig. 6 is that $\Delta\mathcal{S}_3$, $\Delta\mathcal{S}_1+\Delta\mathcal{S}_3$, and $\Delta\mathcal{S}_2+\Delta\mathcal{S}_3$ cannot be expected, in general, to be positive, reflecting that they cannot be expressed in the form of Eq. (9) and thus do not obey IFTs. This means previous approaches in which the entropy production can always be divided into two positive quantities [19–21] and where the housekeeping heat can be expected to obey an IFT [14] do not extend to the systems considered here.

We consider this example to be a helpful illustration of how entropy production cannot always be divided into two contributions which derive from relaxation and an absence of detailed balance due to a nonequilibrium constraint, respectively. Explicitly, the nonequilibrium constraint here is the constant force which produces entropy in the stationary state by inducing a constant flux around the ring. The mean rate of entropy production in the stationary state is characterized by $d\langle\Delta\mathcal{S}_2\rangle/dt$, which remains constant throughout the process due to the constant magnitude of the force that is applied. However, both $\Delta\mathcal{S}_2$ and $\Delta\mathcal{S}_3$ are nonzero only in the presence of a nonequilibrium constraint which breaks detailed balance. At the same time, the *mean* rate of change of $\Delta\mathcal{S}_3$ is nonzero only when the distribution is relaxing to a new stationary solution, in the same manner as $\Delta\mathcal{S}_1$. While $\Delta\mathcal{S}_1$ describes the entropy production that arises from an evolution of the probability distribution of a general set of variables, $\Delta\mathcal{S}_3$ expresses what $\Delta\mathcal{S}_1$ explicitly leaves out: the additional impact of relaxation on entropy production that relates to the *a priori* physical specification of the variables as odd or even. Clearly, given that the nonequilibrium constraint is a force of constant magnitude, reflected by the constant $d\langle\Delta\mathcal{S}_2\rangle/dt$, it is reasonable to consider the sum of $\Delta\mathcal{S}_1$ and $\Delta\mathcal{S}_3$ as the contribution that arises due to relaxation to a new stationary state, particularly when the form of its mean rate of change in Fig. 6 is contrasted with that of $\Delta\mathcal{S}_{\rm tot}$, $\Delta\mathcal{S}_1$, and $\Delta\mathcal{S}_2$ in Fig. 5. We may make the analysis complete by considering the SDEs for all contributions. The explicit Ito forms of Eqs. (32), (50), (60), and (65) are given as

$$d\Delta\mathcal{S}_{\rm tot}=-\frac{m}{k_BT}\langle v\rangle dv-\frac{m}{k_BT}(v-\langle v\rangle)\frac{d\langle v\rangle}{dt}dt+\frac{F(t)}{k_BT}dx, \tag{97}$$

$$d\Delta\mathcal{S}_1=\frac{1}{k_BT}\left(\frac{F(t)}{\gamma}-m\langle v\rangle\right)dv-\frac{m}{k_BT}(v-\langle v\rangle)\frac{d\langle v\rangle}{dt}dt, \tag{98}$$

$$d\Delta\mathcal{S}_2=\frac{F(t)}{\gamma k_BT}dv+\frac{F(t)}{k_BT}dx, \tag{99}$$

$$d\Delta\mathcal{S}_3=-\frac{2F(t)}{\gamma k_BT}dv \tag{100}$$

and illustrate the behavior of all the contributions. $d\Delta\mathcal{S}_{\rm tot}$ is only zero when $\langle v\rangle=0$, $F=0$, and $d\langle v\rangle/dt=0$, meaning the system is in the equilibrium state. $d\Delta\mathcal{S}_1$ is zero whenever $\langle v\rangle=F/m\gamma$ and $d\langle v\rangle/dt=0$ corresponding to any stationary state, equilibrium or otherwise, while $d\Delta\mathcal{S}_2$ and $d\Delta\mathcal{S}_3$ contribute independently of properties of the distribution (namely $\langle v\rangle$), but only when the nonequilibrium constraint is present such that $F(t)\neq 0$. $d\Delta\mathcal{S}_3$, however, has a mean contribution of zero at stationarity since $\langle dv\rangle=0$ for any stationary state. We can calculate distributions of all the contributions, as measured from the force reversal, numerically using the above SDEs, and we demonstrate the validity of IFTs, where appropriate, in Figs. 7 and 8. We observe that all distributions take Gaussian form, to be expected, as the model is essentially a recasting of the overdamped dragged oscillator found in [39] where the further, but nongeneral, detailed fluctuation theorem symmetry $p(\Delta\mathcal{S}_{\rm tot})/p(-\Delta\mathcal{S}_{\rm tot})=\exp(\Delta\mathcal{S}_{\rm tot})$ has been noted to hold over finite times [39], but is stressed elsewhere [40] to be coincidental. Further insight into this coincidence can be derived from the form of the SDEs which yield Gaussian distributions (for the given initial conditions) since they comprise only drift and additive noise terms [that is, no terms of the form $f(v)dv$]. Such properties, however, do not distract attention from the nature of the contributions which can be readily observed: the distributions in $\Delta\mathcal{S}_1$ and $\Delta\mathcal{S}_3$ develop fastest at first, reflecting the initially fast response of the distribution to the change in force. However, distributions for both $\Delta\mathcal{S}_2$ and $\Delta\mathcal{S}_{\rm tot}$ develop steadily, owing to their

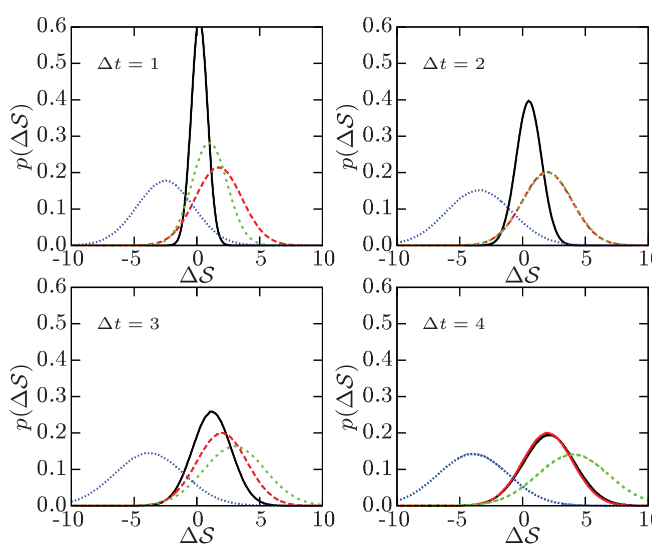

FIG. 7. (Color online) Distributions of entropy productions $\Delta\mathcal{S}_{\rm tot}$ (solid line), $\Delta\mathcal{S}_1$ (wide dashed line), $\Delta\mathcal{S}_2$ (narrow dashed line), and $\Delta\mathcal{S}_3$ (dotted line) measured at times $\Delta t=t-t_0=1$, $\Delta t=2$, $\Delta t=3$, and $\Delta t=4$ after the reversal of the force for $F_0=1$, $F_1=-1$, $t_0=1$, and $k_B=m=\gamma=T=1$. We performed 7.5×10^6 Monte Carlo runs with time step $dt=1\times10^{-3}$ to generate the contents of all results in example II.

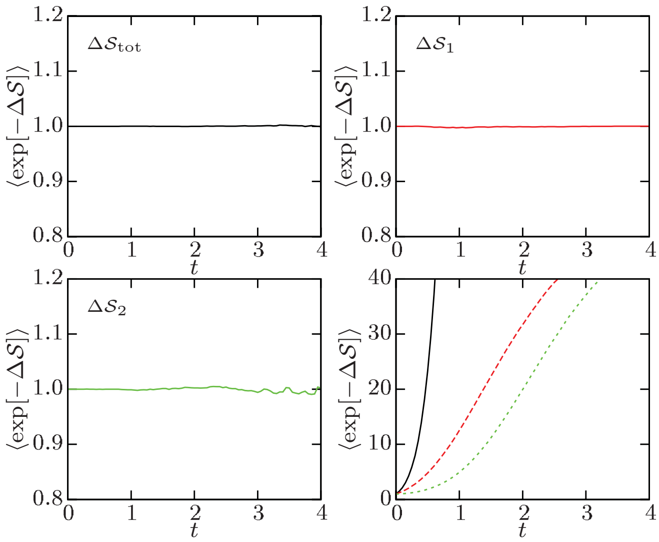

FIG. 8. (Color online) Illustration of adherence to IFTs by consideration of the average $\langle\exp[-\Delta\mathcal{S}]\rangle$ against time $\Delta t = t - t_0$ after the force reversal, for $\Delta\mathcal{S}_{\rm tot}$, $\Delta\mathcal{S}_1$, and $\Delta\mathcal{S}_2$ (indicated) and the failure to adhere to an IFT of $\Delta\mathcal{S}_3$ (solid line, fourth subplot), $\Delta\mathcal{S}_1 + \Delta\mathcal{S}_3$ (wide dashed line, fourth subplot), and $\Delta\mathcal{S}_2 + \Delta\mathcal{S}_3$ (narrow dashed line, fourth subplot) for $F_0 = 1$, $F_1 = -1$, $t_0 = 1$, and $k_B = m = \gamma = T = 1$.

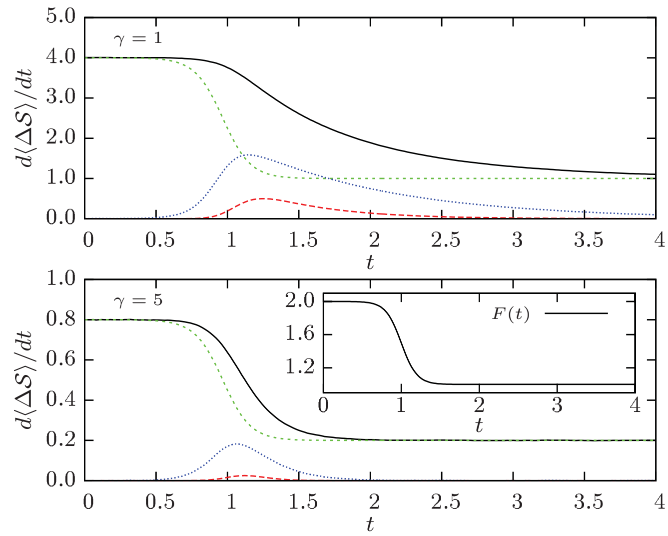

FIG. 9. (Color online) Mean rates of change of $\Delta\mathcal{S}_1$ (wide dashed line), $\Delta\mathcal{S}_2$ (narrow dashed line), $\Delta\mathcal{S}_3$ (dotted line), and their sum $\Delta\mathcal{S}_{\rm tot}$ (solid line) for example II with a time-dependent force given by Eq. (101) with units $k_B = m = T = 1$, and for $\gamma = 1$ (top) and $\gamma = 5$ (bottom).

contributions being characterized by steady heat dissipation. As such, as time progresses, the distribution of $\Delta\mathcal{S}_1$ ceases to develop as the system reaches the new stationary state and the distributions of $\Delta\mathcal{S}_2$ and $\Delta\mathcal{S}_{\rm tot}$ continue to shift to the right until they eventually dominate. Similarly for $\Delta\mathcal{S}_3$, we observe here that the distribution stops evolving despite receiving nonzero contributions.

For completeness, we investigate the same model with a less trivial time dependence in the nonconservative force, along with its approach to the overdamped limit where such systems have been considered previously [21,41]. We employ the force protocol

$$F(t) = 1.5 - 0.5\tanh[-5(t-1)] \qquad (101)$$

and perform the calculations numerically for two values of damping coefficient, $\gamma = 1$ and 5. We point out again that the meaning of $d\langle\Delta\mathcal{S}_2\rangle/dt$ for this system is easily elucidated since the nonequilibrium constraint, $F(t)$, being phase-space-independent, leads to $J_v^{\rm ir,st} \propto p^{\rm st}$ so that

$$\frac{d\langle\Delta\mathcal{S}_2\rangle}{dt} = \frac{d\langle\Delta\mathcal{S}_{\rm tot}\rangle^{\rm st}}{dt} = \frac{F(t)^2}{m\gamma k_B T}. \qquad (102)$$

The mean contributions for such a protocol for two values of the damping coefficient, again starting from the stationary state, are shown in Fig. 9. Note that in this case, the contribution $d\langle\Delta\mathcal{S}_3\rangle/dt$ is positive, reflecting that as the nonconservative force decreases, the instantaneous distribution corresponds to a greater average particle flux than would be expected from the instantaneous value of the force, thus producing more entropy in the process of relaxation from one stationary state to the other than would be expected if the relaxation were instantaneous. As γ increases, the asymmetry of the stationary state (in velocity) decreases and the contribution from $\Delta\mathcal{S}_3$ diminishes. Consequently, the two stationary distributions become increasingly similar, meaning the contribution $\Delta\mathcal{S}_1$ also diminishes rendering the total entropy production almost entirely comprised of the contribution from $\Delta\mathcal{S}_2$. When the full overdamped limit is taken, $\Delta\mathcal{S}_2$ is the only contribution and the results map onto those found in [21].

VII. DISCUSSION AND CONCLUSIONS

We have derived SDEs describing the fluctuating evolution of three contributions to entropy production, along with expressions for their mean behavior, and we demonstrated that two of these contributions obey IFTs and thus are rigorously positive in the mean. Furthermore we have demonstrated that while these two naturally align themselves with the irreversibility associated with relaxation and nonequilibrium constraints, respectively, the inclusion of odd dynamical variables can give rise to a third term, which has no bounds on its sign and which cannot be so readily associated with one origin of entropy production or the other. We have sought to make these expressions as general as possible, within reason, with the intention that they may be applied to any system (physical or otherwise) described by stochastic differential equations, providing a framework for the discussion of entropy production, as defined here, within as wide a range of applications as may be relevant. To this end, we have considered a simple heat conduction problem, and after demonstrating that a full phase-space representation of the dynamics is crucial to the treatment of its entropy production, we have examined specifically how it may be evaluated. The second example, that of a transition between stationary states of drift and diffusion on a ring, demonstrates the need for a third contribution to entropy production in the analysis, and provides some intuitive understanding of its nature.

We suggest that the division of the total entropy production into $\Delta\mathcal{S}_1$, $\Delta\mathcal{S}_2$, and $\Delta\mathcal{S}_3$, as we propose, is always helpful for three main reasons. The first is the ability to identify the physical origins of irreversibility in a process (relaxation and nonequilibrium constraints) and the interplay between them. The second is related to the identification of IFTs with the subsequent positivity requirements and restrictions on the statistics for the two contributions, $\Delta\mathcal{S}_1$ and $\Delta\mathcal{S}_2$, which unambiguously align themselves with the two causes of irreversibility. And thirdly, the more delicate reason is that all three contributions are constructed from total path probability densities with equivalent measures such that they align themselves unambiguously with the same contributions found in master equation approaches which are necessarily formed from path *probabilities* [23]. This third point may be contrasted with an alternative division of the total entropy production into a system and medium contribution, neither of which can be expressed as ratios of total path probabilities. Being formed from a probability density, the system entropy, as defined in [11] and employed here [Eq. (16) implies $\mathcal{S}_{\rm sys} = -\ln p(\boldsymbol{x}(t),t)$], is strictly not dimensionally correct (even if one argues that the relative entropy change is well defined [26]), but moreover it does not share the same form as the system entropy that appears in master equation approaches, which if followed would imply $\mathcal{S}_{\rm sys} = -\ln p(\boldsymbol{x}(t),t)d\boldsymbol{x}(t)$. Such ambiguity then also enters into the definition of the medium entropy, but can be avoided altogether by considering the total entropy production in terms of explicit measures of irreversibility such that $\Delta\mathcal{S}_{\rm tot}$ is comprised of $\Delta\mathcal{S}_1$, $\Delta\mathcal{S}_2$, and $\Delta\mathcal{S}_3$, all four of which do not suffer from such issues.

We note that despite restricting ourselves to uncorrelated processes for the sake of clarity, the method utilized extends naturally to multidimensional Ito processes of the form $d\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x},t)dt + \boldsymbol{B}(\boldsymbol{x},t)d\boldsymbol{W}$ in vector form and where $d\boldsymbol{W}$ is a vector of independent uncorrelated Wiener processes and $\boldsymbol{B}(\boldsymbol{x}) \equiv \boldsymbol{B}(\boldsymbol{x},t)$ is a matrix such that the generalized Fokker-Planck equation is governed by the diffusion matrix $\boldsymbol{D}(\boldsymbol{x}) = (1/2)\boldsymbol{B}(\boldsymbol{x})\boldsymbol{B}(\boldsymbol{x})^{\rm T}$ with inverse $\boldsymbol{D}^{-1}(\boldsymbol{x})$. This can be achieved by utilizing the full form of the propagator found in [29] and recognizing that the conversion from Stratonovich to Ito form for the averaging introduces an additional term for every independent Wiener process with which the increment is correlated. The condition $D_i(\boldsymbol{\varepsilon x}) = D_i(\boldsymbol{x})$ is extended to be $D_{ij}(\boldsymbol{\varepsilon x}) = \varepsilon_i\varepsilon_j D_{ij}(\boldsymbol{x})$ (which in some instances, where an odd and even variable are entirely correlated, is necessary for the reverse path to exist) and renders the mean contribution $d\langle\Delta\mathcal{S}_3\rangle/dt$ unchanged and the remainder of the form

$$\frac{d\langle\Delta\mathcal{S}_{\rm tot}\rangle}{dt} = \int d\boldsymbol{x}\ p(\boldsymbol{x})\left[\frac{\boldsymbol{J}^{\rm ir}(\boldsymbol{x})}{p(\boldsymbol{x})}\right]^{\rm T}\boldsymbol{D}^{-1}(\boldsymbol{x})\left[\frac{\boldsymbol{J}^{\rm ir}(\boldsymbol{x})}{p(\boldsymbol{x})}\right], \tag{103}$$

$$\frac{d\langle\Delta\mathcal{S}_1\rangle}{dt} = \int d\boldsymbol{x}\ p(\boldsymbol{x}) \times\left[\frac{\boldsymbol{J}^{\rm ir}(\boldsymbol{x})}{p(\boldsymbol{x})}-\frac{\boldsymbol{J}^{\rm ir,st}(\boldsymbol{x})}{p^{\rm st}(\boldsymbol{x})}\right]^{\rm T}\boldsymbol{D}^{-1}(\boldsymbol{x})\left[\frac{\boldsymbol{J}^{\rm ir}(\boldsymbol{x})}{p(\boldsymbol{x})}-\frac{\boldsymbol{J}^{\rm ir,st}(\boldsymbol{x})}{p^{\rm st}(\boldsymbol{x})}\right], \tag{104}$$

and

$$\frac{d\langle\Delta\mathcal{S}_2\rangle}{dt} = \int d\boldsymbol{x}\ p(\boldsymbol{x})\left[\frac{\boldsymbol{J}^{\rm ir,st}(\boldsymbol{\varepsilon x})}{p^{\rm st}(\boldsymbol{\varepsilon x})}\right]^{\rm T}\boldsymbol{D}^{-1}(\boldsymbol{\varepsilon x})\left[\frac{\boldsymbol{J}^{\rm ir,st}(\boldsymbol{\varepsilon x})}{p^{\rm st}(\boldsymbol{\varepsilon x})}\right], \tag{105}$$

all three of which are non-negative since $\boldsymbol{D}^{-1}(\boldsymbol{x})$ is positive semidefinite. We must acknowledge the restriction $D_{ij}(\boldsymbol{\varepsilon x}) = \varepsilon_i\varepsilon_j D_{ij}(\boldsymbol{x})$ [and the equivalent $D_i(\boldsymbol{\varepsilon x}) = D_i(\boldsymbol{x})$] used in this work, but we note that its relaxation, in general, may not be straightforward. However, we point out that systems without such a symmetry are unlikely to be physically meaningful. It would be natural, however, to explore examples involving odd dynamical variables such as angular momentum and magnetic dipole moments, and to include driving by external forces that are themselves odd under time reversal, such as torques and magnetic fields. We expect to find further richness in the phenomenology of entropy production associated with stochastic dynamical behavior.

ACKNOWLEDGMENT

R.E.S. acknowledges financial support from the UK Engineering and Physical Sciences Research Council.

APPENDIX A: THE USE OF SHORT TIME PROPAGATORS WITH MULTIPLICATIVE NOISE

Here we consider one of the terms in Eq. (25) and derive Eq. (29). By utilizing Eqs. (24) and (28), we can describe an increment in the medium entropy production according to the formalism of Seifert [11] as

$$\begin{aligned}
d\Delta\mathcal{S}_{\rm med} = \sum_i \frac{1}{2}\ln D_i(\boldsymbol{r}')-\frac{1}{2}\ln D_i(\boldsymbol{r})+\frac{dx_i^2}{4D_i(\boldsymbol{r}')dt}-\frac{dx_i^2}{4D_i(\boldsymbol{r})dt}+\frac{dx_i}{2}\left(\frac{A_i^{\rm rev}(\boldsymbol{r})}{D_i(\boldsymbol{r})}+\frac{A_i^{\rm ir}(\boldsymbol{r})}{D_i(\boldsymbol{r})}+\frac{A_i^{\rm ir}(\boldsymbol{r}')}{D_i(\boldsymbol{r}')}-\frac{A_i^{\rm rev}(r')}{D_i(\boldsymbol{r}')}\right.\\
\left.-2a\frac{1}{D_i(\boldsymbol{r})}\frac{\partial D_i(\boldsymbol{r})}{\partial r_i}-2b\frac{1}{D_i(\boldsymbol{r}')}\frac{\partial D_i(\boldsymbol{r}')}{\partial r_i'}\right)-\frac{dt}{4}\left(\frac{\left[A_i^{\rm rev}(\boldsymbol{r})+A_i^{\rm ir}(\boldsymbol{r})\right]^2}{D_i(\boldsymbol{r})}-\frac{\left[A_i^{\rm rev}(\boldsymbol{r}')-A_i^{\rm ir}(\boldsymbol{r}')\right]^2}{D_i(\boldsymbol{r}')}\right)\\
-adt\left[D_i(\boldsymbol{r})\frac{\partial}{\partial r_i}\left(\frac{A_i^{\rm ir}(\boldsymbol{r})}{D_i(\boldsymbol{r})}\right)+D_i(\boldsymbol{r})\frac{\partial}{\partial r_i}\left(\frac{A_i^{\rm rev}(\boldsymbol{r})}{D_i(\boldsymbol{r})}\right)\right]-bdt\left[-D_i(\boldsymbol{r}')\frac{\partial}{\partial r_i'}\left(\frac{A_i^{\rm ir}(\boldsymbol{r}')}{D_i(\boldsymbol{r}')}\right)+D_i(\boldsymbol{r}')\frac{\partial}{\partial r_i'}\left(\frac{A_i^{\rm rev}(\boldsymbol{r}')}{D_i(\boldsymbol{r}')}\right)\right]\\
+a^2dt\left[\frac{\partial^2 D_i(\boldsymbol{r})}{\partial r_i^2}-\frac{1}{D_i(\boldsymbol{r})}\left(\frac{\partial D_i(\boldsymbol{r})}{\partial r_i^2}\right)^2\right]-b^2dt\left[\frac{\partial^2 D_i(\boldsymbol{r}')}{\partial r_i'^2}-\frac{1}{D_i(\boldsymbol{r}')}\left(\frac{\partial D_i(\boldsymbol{r}')}{\partial r_i'^2}\right)^2\right],
\end{aligned} \tag{A1}$$

where time dependence in variables $A^{\rm ir}$, $A^{\rm rev}$, and D_i is assumed but not explicitly written for brevity. We may proceed by understanding that the quantity dx_i is an increment in an underlying SDE, meaning that we must consider all multiplications of the form $f(r)dx$ as infinitesimal stochastic integrals with a summation rule defined by the evaluation point r. For example, $r = x$ would imply an Ito integration, $r = (1/2)(x + x')$ would imply Stratonovich, and so on. To consolidate the above, it is sensible to convert all multiplications into one type, for which we choose Ito in order to apply the Ito stochastic calculus transparently using the heuristic rules $(dW_i)^2 = dt$ and $dW_i dW_j = 0$, and to drop all terms of order $dt^{3/2}$ and higher. To do so, we apply the following reasoning. For a suitably smooth function, $f(\boldsymbol{r})$, and for infinitesimal dt, with $\boldsymbol{r}$ constructed from $\boldsymbol{x}$ and $\boldsymbol{x}'$ using a parameter a, and for the case of a diagonal diffusion matrix, we may write

$$\begin{aligned} f(\boldsymbol{r}) &= f[(1-a)\boldsymbol{x} + a\boldsymbol{x}'] \simeq (1-a)f(\boldsymbol{x}) + af(\boldsymbol{x}') = f(\boldsymbol{x}) + a[f(\boldsymbol{x}') - f(\boldsymbol{x})] = f(\boldsymbol{x}) + adf(\boldsymbol{x}) \\ &= f(\boldsymbol{x}) + a\left(\frac{\partial f(\boldsymbol{x})}{\partial t}dt + \nabla f(\boldsymbol{x})\cdot d\boldsymbol{x} + \sum_i \frac{B_i(\boldsymbol{x})^2}{2}\nabla^2 f(\boldsymbol{x})dt\right) \\ &= f(\boldsymbol{x}) + a\left(\frac{\partial f(\boldsymbol{x})}{\partial t}dt + \sum_i \frac{1}{2}[B_i(\boldsymbol{x})]^2\nabla^2 f(\boldsymbol{x})dt + \sum_i \frac{\partial f(\boldsymbol{x})}{\partial x_i}[A_i(\boldsymbol{x})dt + B_i(\boldsymbol{x})dW_i]\right). \end{aligned} \tag{A2}$$

Considering all instances of multiplication along with the definitions of $\boldsymbol{r}$, $\boldsymbol{r}'$, a, and b, we find the following heuristic rules:

$$f(\boldsymbol{r})dt = f(\boldsymbol{x})dt + O(dt^{3/2}), \tag{A3}$$

$$f(\boldsymbol{r}')dt = f(\boldsymbol{x})dt + O(dt^{3/2}), \tag{A4}$$

$$f(\boldsymbol{r})dx_i = f(\boldsymbol{x})dx_i + 2aD_i(\boldsymbol{x})\frac{\partial f(\boldsymbol{x})}{\partial x_i}dt + O(dt^{3/2}), \tag{A5}$$

$$f(\boldsymbol{r}')dx_i = f(\boldsymbol{x})dx_i + 2(1-b)D_i(\boldsymbol{x})\frac{\partial f(\boldsymbol{x})}{\partial x_i}dt + O(dt^{3/2}), \tag{A6}$$

giving us a method for converting all multiplications into Ito form. We use a similar reasoning to approximate

$$D_i(\boldsymbol{r}) \simeq D_i(\boldsymbol{x}) + ad[D_i(\boldsymbol{x})] \tag{A7}$$

$$D_i(\boldsymbol{r}') \simeq D_i(\boldsymbol{x}) + (1-b)d[D_i(\boldsymbol{x})] \tag{A8}$$

which along with the approximations to *second* order in dx_i, and therefore $d[D_i(\boldsymbol{x})]$, of the form

$$\{1 + d[D_i(\boldsymbol{x})]\}^{-1} \simeq 1 - d[D_i(\boldsymbol{x})] + d[D_i(\boldsymbol{x})]^2, \tag{A9}$$

$$\ln\{1 + d[D_i(\boldsymbol{x})]\} \simeq d[D_i(\boldsymbol{x})] - \frac{d[D_i(\boldsymbol{x})]^2}{2}, \tag{A10}$$

and an Ito definition of $d[D_i(\boldsymbol{x})]$, allow us to write the first four terms in Eq. (A1) to first order in dt as

$$\begin{aligned} &\frac{1}{2}\ln D_i(\boldsymbol{r}') - \frac{1}{2}\ln D_i(\boldsymbol{r}) + \frac{dx_i^2}{4D_i(\boldsymbol{r}')dt} - \frac{dx_i^2}{4D_i(\boldsymbol{r})dt} \\ &\quad\simeq \frac{[(1-b)^2 - a^2]}{2D_i(\boldsymbol{x})}\left(\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\right)^2 dt. \end{aligned} \tag{A11}$$

Using the above, and the heuristic rules in Eqs. (A3)–(A6), we obtain

$$\begin{aligned} d\Delta\mathcal{S}_{\text{med}} = \sum_i &\frac{A_i^{\text{ir}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}dx_i - \frac{A_i^{\text{rev}}(\boldsymbol{x})A_i^{\text{ir}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}dt \\ &+ \frac{\partial A_i^{\text{ir}}(\boldsymbol{x})}{\partial x_i}dt - \frac{\partial A_i^{\text{rev}}(\boldsymbol{x})}{\partial x_i}dt \\ &+ \frac{A_i^{\text{rev}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}dt - \frac{A_i^{\text{ir}}(\boldsymbol{x})}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}dt \\ &- (a+b)\frac{1}{D_i(\boldsymbol{x})}\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}dx_i \\ &+ (b^2 - 2b - a^2)\left[\frac{\partial^2 D_i(\boldsymbol{x})}{\partial x_i^2} - \frac{1}{D_i(\boldsymbol{x})}\left(\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\right)^2\right]dt \\ &+ \frac{[(1-b)^2 - a^2]}{2D_i(\boldsymbol{x})}\left(\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\right)^2 dt, \end{aligned} \tag{A12}$$

which depends on the choice of a and b. We note, however, that without multiplicative noise [i.e., for $(\partial/\partial x_i)D_i(\boldsymbol{x}) = 0$], where the inherent mathematical ambiguity in stochastic integrals is absent, the dependence on evaluation points (a and b) disappears. So this dependence is evidently related to the ambiguity of the evaluation point in a stochastic integral. However, the underlying SDEs and entropy are not, and should not be, ambiguous since we have specified Ito SDEs and have used the short time propagator appropriate for their corresponding Fokker-Planck equation. Since all evaluation points lead to the correct path probability density, we are not obliged to consider, for example, only Ito-type multiplication ($a = b = 0$) simply because the underlying SDEs are of Ito form. Rather, to proceed we recognize that with multiplicative noise, we must ensure that we evaluate the two transition probability densities at precisely the same coordinates and not just at the same time (which would suffice for additive noise), for the same reasons that make stochastic integration sensitive to the specific integration scheme when the integrand has a dependence on the integrating variable. Alternatively, it may be reasoned that as $dt \to 0$, we require the short time propagators to approach jump transition probabilities of a master equation. Under such a description, the entropy production can be unambiguously described by the ratios of probabilities appearing in a master equation approach [19,23]. Therefore, if we represent such a quantity using the short time propagators, we require the transition rates in both numerator and denominator to be evaluated equivalently. Since these are characterized by our system variables A_i, B_i, etc., to effect such a condition we require $\boldsymbol{r}' = \boldsymbol{r}$, which is equivalent to making the choice $b = 1 - a$, noting that a is still a free parameter. One may think of this as insisting that the path transformation

$\vec{x}^{\dagger}(t) = \boldsymbol{\varepsilon}\vec{x}(\tau - t)$ persists on a (sub) infinitesimal scale, which would not matter in normal calculus, so that the noise is experienced in precisely the right way. Inserting $b = 1 - a$ into the above yields Eq. (29), which has no dependence on the choice a, an indication that it is defined in a sound fashion.

APPENDIX B: CONSISTENCY OF ENTROPY CONTRIBUTIONS

By construction, we have $\Delta\mathcal{S}_{\rm tot} = \Delta\mathcal{S}_1 + \Delta\mathcal{S}_2 + \Delta\mathcal{S}_3$, which according to the expressions in Eqs. (38), (54), (62), and (68) means we require

$$\begin{aligned}\frac{d\langle\Delta\mathcal{S}_{\rm tot}\rangle}{dt} &= \sum_i\int d\boldsymbol{x}\,\frac{[J_i^{\rm ir}(\boldsymbol{x})]^2}{p(\boldsymbol{x})D_i(\boldsymbol{x})}\\ &= \sum_i\int d\boldsymbol{x}\,\frac{p(\boldsymbol{x})}{D_i(\boldsymbol{x})}\left(\frac{J_i^{\rm ir}(\boldsymbol{x})}{p(\boldsymbol{x})} - \frac{J_i^{\rm st,ir}(\boldsymbol{x})}{p^{\rm st}(\boldsymbol{x})}\right)^2\\ &\quad + \int d\boldsymbol{x}\,\frac{p(\boldsymbol{x})}{D_i(\boldsymbol{\varepsilon x})}\left(\frac{J_i^{\rm ir,st}(\boldsymbol{\varepsilon x})}{p^{\rm st}(\boldsymbol{\varepsilon x})}\right)^2\\ &\quad + \int d\boldsymbol{x}\,J_i(\boldsymbol{x})[\phi_i'(\boldsymbol{x}) - \varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})].\end{aligned}\tag{B1}$$

Thus in order that everything is consistent, we require

$$\begin{aligned}0 = \sum_i\int d\boldsymbol{x}\Bigg[&-2\frac{J_i^{\rm ir}(\boldsymbol{x})}{D_i(\boldsymbol{x})}\frac{J_i^{\rm ir,st}(\boldsymbol{x})}{p^{\rm st}(\boldsymbol{x})} + \frac{p(\boldsymbol{x})}{D_i(\boldsymbol{x})}\left(\frac{J_i^{\rm ir,st}(\boldsymbol{x})}{p^{\rm st}(\boldsymbol{x})}\right)^2\\ &+\frac{p(\boldsymbol{x})}{D_i(\boldsymbol{\varepsilon x})}\left(\frac{J_i^{\rm ir,st}(\boldsymbol{\varepsilon x})}{p^{\rm st}(\boldsymbol{\varepsilon x})}\right)^2 + J_i(\boldsymbol{x})[\phi_i'(\boldsymbol{x}) - \varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})]\Bigg].\end{aligned}\tag{B2}$$

By substitution of the definitions of the fluxes, this reduces to

$$\begin{aligned}0 = \sum_i\int d\boldsymbol{x}\Bigg[&2A_i^{\rm ir}(\boldsymbol{x})\frac{\partial p(\boldsymbol{x})}{\partial x_i} - 2\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\frac{\partial p(\boldsymbol{x})}{\partial x_i}\\ &+\frac{\partial[D_i(\boldsymbol{x})p(\boldsymbol{x})]}{\partial x_i}[\phi_i'(\boldsymbol{x}) + \varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})]\\ &+p(\boldsymbol{x})D_i(\boldsymbol{x})\{[\phi_i'(\boldsymbol{x})]^2 + [\phi_i'(\boldsymbol{\varepsilon x})]^2\}\\ &-2p(\boldsymbol{x})\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}[\phi_i'(\boldsymbol{x}) + \varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})]\\ &+p(\boldsymbol{x})\big[A_i^{\rm ir}(\boldsymbol{x}) + A_i^{\rm rev}(\boldsymbol{x})\big]\phi_i'(\boldsymbol{x})\\ &+\varepsilon_i p(\boldsymbol{x})\big[A_i^{\rm ir}(\boldsymbol{x}) - A_i^{\rm rev}(\boldsymbol{x})\big]\phi_i'(\boldsymbol{\varepsilon x})\Bigg].\end{aligned}\tag{B3}$$

Integrating by parts, dropping or canceling boundary terms, and using the definition of the irreversible and reversible drift terms yield the condition

$$\begin{aligned}0 = \sum_i\int d\boldsymbol{x}\;p(\boldsymbol{x})\Bigg[&-2\frac{\partial A_i^{\rm ir}(\boldsymbol{x})}{\partial x_i} + 2\frac{\partial^2 D_i(\boldsymbol{x})}{\partial x_i^2}\\ &-D_i(\boldsymbol{x})[\phi_i''(\boldsymbol{x}) + \phi_i''(\boldsymbol{\varepsilon x})] + D_i(\boldsymbol{x})\{[\phi_i'(\boldsymbol{x})]^2 + [\phi_i'(\boldsymbol{\varepsilon x})]^2\}\\ &-2\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}[\phi_i'(\boldsymbol{x}) + \varepsilon_i\phi_i'(\boldsymbol{\varepsilon x})] + \big[A_i^{\rm ir}(\boldsymbol{x}) + A_i^{\rm rev}(\boldsymbol{x})\big]\phi_i'(\boldsymbol{x})\\ &+\big[A_i^{\rm ir}(\boldsymbol{\varepsilon x}) + A_i^{\rm rev}(\boldsymbol{\varepsilon x})\big]\phi_i'(\boldsymbol{\varepsilon x})\Bigg].\end{aligned}\tag{B4}$$

The divergenceless stationary distribution condition, however, yields

$$\begin{aligned}0 = \Bigg(&-\big[A_i^{\rm ir}(\boldsymbol{x}) + A_i^{\rm rev}(\boldsymbol{x})\big]\phi_i'(\boldsymbol{x}) - D_i(\boldsymbol{x})[\phi_i'(\boldsymbol{x})]^2\\ &+\frac{\partial A_i^{\rm ir}(\boldsymbol{x})}{\partial x_i} + \frac{\partial A_i^{\rm rev}(\boldsymbol{x})}{\partial x_i} - \frac{\partial^2 D_i(\boldsymbol{x})}{\partial x_i^2}\\ &+D_i(\boldsymbol{x})\phi_i''(\boldsymbol{x}) + 2\frac{\partial D_i(\boldsymbol{x})}{\partial x_i}\phi_i'(\boldsymbol{x})\Bigg)e^{-\phi(\boldsymbol{x})},\end{aligned}\tag{B5}$$

but also

$$\begin{aligned}0 = \Bigg(&-\big[A_i^{\rm ir}(\boldsymbol{\varepsilon x}) + A_i^{\rm rev}(\boldsymbol{\varepsilon x})\big]\phi_i'(\boldsymbol{\varepsilon x}) - D_i(\boldsymbol{\varepsilon x})[\phi_i'(\boldsymbol{\varepsilon x})]^2\\ &+\frac{\partial A_i^{\rm ir}(\boldsymbol{\varepsilon x})}{\partial(\varepsilon_i x_i)} + \frac{\partial A_i^{\rm rev}(\boldsymbol{\varepsilon x})}{\partial(\varepsilon_i x_i)} - \frac{\partial^2 D_i(\boldsymbol{\varepsilon x})}{\partial(\varepsilon_i x_i)^2}\\ &+D_i(\boldsymbol{\varepsilon x})\phi_i''(\boldsymbol{\varepsilon x}) + 2\frac{\partial D_i(\boldsymbol{\varepsilon x})}{\partial(\varepsilon_i x_i)}\phi_i'(\boldsymbol{\varepsilon x})\Bigg)e^{-\phi(\boldsymbol{\varepsilon x})}.\end{aligned}\tag{B6}$$

Combining the two conditions in Eqs. (B5) and (B6) yields the contents of the large square brackets in Eq. (B4), and so the result is proved.

[1] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
[2] D. J. Evans and D. J. Searles, Phys. Rev. E **52**, 5839 (1995).
[3] D. J. Evans and D. J. Searles, Adv. Phys. **51**, 1529 (2002).
[4] D. M. Carberry, J. C. Reid, G. M. Wang, E. M. Sevick, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **92**, 140601 (2004).
[5] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995).
[6] J. Kurchan, J. Phys. A **31**, 3719 (1998).
[7] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999).
[8] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[9] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[10] G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998).
[11] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
[12] K. Sekimoto, Prog. Theor. Phys. Suppl. **130**, 17 (1998).
[13] T. Hatano and S. I. Sasa, Phys. Rev. Lett. **86**, 3463 (2001).
[14] T. Speck and U. Seifert, J. Phys. A **38**, L581 (2005).
[15] V. Y. Chernyak, M. Chertkov, and C. Jarzynski, J. Stat. Mech. (2006) P08001.
[16] M. Esposito, U. Harbola, and S. Mukamel, Phys. Rev. E **76**, 031132 (2007).
[17] H. Ge, Phys. Rev. E **80**, 021137 (2009).
[18] H. Ge and H. Qian, Phys. Rev. E **81**, 051133 (2010).
[19] M. Esposito and C. Van den Broeck, Phys. Rev. Lett. **104**, 090601 (2010).
[20] M. Esposito and C. Van den Broeck, Phys. Rev. E **82**, 011143 (2010).

[21] C. Van den Broeck and M. Esposito, Phys. Rev. E **82**, 011144 (2010).
[22] Y. Oono and M. Paniconi, Prog. Theor. Phys. Suppl. **130**, 29 (1998).
[23] R. E. Spinney and I. J. Ford, Phys. Rev. Lett. **108**, 170603 (2012).
[24] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed. (Springer, New York, 1989).
[25] R. J. Harris and G. M. Schütz, J. Stat. Mech. (2007) P07020.
[26] U. Seifert, Eur. Phys. J. B **64**, 423 (2008).
[27] K. H. Kim and H. Qian, Phys. Rev. Lett. **93**, 120602 (2004).
[28] K. H. Kim and H. Qian, Phys. Rev. E **75**, 022102 (2007).
[29] C. Wissel, Z. Phys. B **35**, 185 (1979).
[30] T. Tomé and M. J. de Oliveira, Phys. Rev. E **82**, 021120 (2010).
[31] D. Daems and G. Nicolis, Phys. Rev. E **59**, 4000 (1999).
[32] M. Matsuo and S. Sasa, Physica A **276**, 188 (2000).
[33] N. van Kampen, J. Phys. Chem. Solids **49**, 673 (1988).
[34] G. Stolovitzky, Phys. Lett. A **241**, 240 (1998).
[35] K. Sekimoto, *Stochastic Energetics*, Lecture Notes in Physics, Vol. 799 (Springer, Berlin, 2010).
[36] M. Widder and U. Titulaer, Physica A **154**, 452 (1989).
[37] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, 1995).
[38] K. Sekimoto, Prog. Theor. Phys. Suppl. **130**, 17 (1998).
[39] A. Saha, S. Lahiri, and A. M. Jayannavar, Phys. Rev. E **80**, 011117 (2009).
[40] B. H. Shargel, J. Phys. A **43**, 135002 (2010).
[41] T. Speck, V. Blickle, C. Bechinger, and U. Seifert, Europhys. Lett. **79**, 30002 (2007).