Nonequilibrium statistical physics beyond the ideal heat bath approximation

Jonathan Asher Pachter[1,2,*] and Ken A. Dill[1,2,3,†]

[1]*Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, New York 11794, USA*
[2]*Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794, USA*
[3]*Department of Chemistry, Stony Brook University, Stony Brook, New York 11794, USA*

Important models of nonequilibrium statistical physics (NESP) are limited by a commonly used, but often unrecognized, near-equilibrium approximation. Fokker-Planck and Langevin equations, the Einstein and random-flight diffusion models, and the Schnakenberg model of biochemical networks suppose that fluctuations are due to an ideal equilibrium bath. But far from equilibrium, this perfect bath concept does not hold. A more principled approach should derive the rate fluctuations from an underlying dynamical model, rather than assuming a particular form. Here, using *maximum caliber* as the underlying principle, we derive corrections for NESP processes in an imperfect—but more realistic—environment, corrections which become particularly important for a system driven strongly away from equilibrium. Beyond characterizing a heat bath by the single equilibrium property of its temperature, the bath's speed and size must also be used to characterize the bath's ability to handle fast or large fluctuations.

I. INTRODUCTION

Nonequilibrium statistical physics (NESP) aims to explain force-flow relationships in situations where fluctuations and rate distributions matter. However, modeling in NESP often rests on near-equilibrium approximations that we describe below. But how do we go beyond these approximations to treat flows further from equilibrium? We desire an NESP procedure resembling *equilibrium* statistical physics (ESP).

In ESP, you assert: (i) the second law of thermodynamics, equivalent to the maximum entropy (MaxEnt) variational principle; and (ii) the existence of a "perfect bath" holding the system at constant known average quantities, like energy or particle number. This gives a principled way to *derive* the full equilibrium probability distribution. It gives the Boltzmann distribution (BD) law with exponential dependence on energy—a result of great generality—which can give macroscopic behaviors from a microscopic model of energy levels.

In contrast, traditional modeling in NESP has a more limited foundation. There is no consensus on the proper variational principle. Near equilibrium, the second law of thermodynamics is rooted in experiments through the Clausius relationship between entropy and heat, but there is no equivalent experimental anchor further from equilibrium. While some have favored using the entropy dissipation rate dS/dt, it has been shown that this is limited to near-equilibria [1]. Furthermore, common NESP models *assume* near-equilibrium rate distributions, based on equilibrium temperature baths, rather than deriving them based on information pertinent to the system at hand, such as properties of flows.

Unlike in ESP, traditional NESP gives no principled connection between macroscopic rate laws (e.g., Fick's, Fourier's, and Ohm's Laws, Newtonian viscosities, etc.) and microscopic modeling (e.g., the Langevin and Fokker-Planck equations, random walks and random flights, the Einstein diffusion relation, the Schnakenberg relations for biochemical networks, etc.). In forming the latter set, fluctuations are ascribed *as if* they are provided by an infinite equilibrium heat bath, but transport phenomena result from gradients, so it does not make sense to derive their fluctuations from the zero-gradient limit that an equilibrium bath supplies. In Langevin modeling, appending Gaussian white noise is known to cause serious issues when dynamics has nonlinearities, such as in viscoelastic materials, or in the van Kampen diode, wherein Langevin modeling leads to violations of conservation of energy [2].

Although an infinite equilibrium bath is a good approximation in many circumstances, strongly driven systems will see a bath that is finite in size and responsiveness. If a system has large fluctuations, then a finite bath will not be able to absorb or accommodate them; if a system changes too quickly, then it may generate corresponding heat faster than the bath can take it up, and the bath itself will be pushed out of equilibrium. Quantitatively, we study two dimensionless metrics for how far a bath is from its idealization: the ratio of heat capacity of the system to that of the bath, and a measure of the bath's speed in responding to any perturbation, relative to the system's speed in generating it.

To model these effects, we require a generative principle of the type indicated above as desirable: a sound variational principle that derives rate distributions from microscopic models. It has been shown that such a principle for nonequilibria is maximum caliber (MaxCal) [1,3,4]. MaxCal is essentially the maximum entropy inference procedure applied to path

*jonathan.pachter@stonybrook.edu
†dill@laufercenter.org

probabilities of flow trajectories; its validity and power rest on the mathematical bedrock of the Shore and Johnson arguments regarding MaxEnt [5], which have been recently updated and further explained by Ariel Caticha, who constructs the entire procedure of entropic inference from the ground up based on physically meaningful requirements [6]. In this paper, we employ MaxEnt and MaxCal to derive first-order corrections to baths that are not infinite in size or speed.

It is most illuminating to focus on a particular model. Here, we focus on the Fokker-Planck (FP) evolution equation for dynamics on continuous microstates:

$$\frac{\partial p(x,t)}{\partial t} = D\beta \frac{\partial}{\partial x}[E'(x)p(x,t)] + D\frac{\partial^2 p(x,t)}{\partial x^2}, \quad (1)$$

where D is the diffusion coefficient for the microstates labeled by x, β is the inverse temperature, and $E(x)$ is the energy landscape.[1] The FP model is important due to its broad use in many disciplines, including prominent use in stochastic thermodynamics[2] [7].

II. REVIEW OF DYNAMICS IN IDEAL THERMAL RESERVOIRS

Before finding corrections to the BD and FP using MaxEnt and MaxCal, we will review how to use MaxEnt and MaxCal to derive the original, ideal-bath forms of the BD and FP. In doing so, we will note where the ideal bath assumptions arise, as a basis for going beyond them later.

A. Boltzmann distribution from maximum entropy

MaxEnt infers the microscopic probability distribution $P(x)$—given certain limited information about a system—by maximizing the Boltzmann-Gibbs-Shannon entropy,

$$S[P] = -\sum_x P(x) \ln\left[\frac{P(x)}{Q(x)}\right], \quad (2)$$

relative to some prior $Q(x)$, constrained by the known information.[3] For example, to derive the BD, you impose knowledge of the average system energy $\langle E(x)\rangle$ through use of a Lagrange multiplier β, and vary $S[P]$ with uniform prior, to get

$$P(x) = \frac{1}{Z}\exp[-\beta E(x)], \quad (3)$$

where $Z = \sum \exp[-\beta E(x)]$ is the partition function, normalizing the distribution.

---

[1]x here does not specifically refer to position, but rather to whatever labels the microstates, which could be positions, or velocities, or angles, etc.

[2]Although much of the Stochastic Thermodynamics literature assumes ideal bath behavior when postulating models—see, e.g., the review by Seifert [7]—the powerful tools used to analyze the irreversibility of those models in the Stochastic Thermodynamics papers remain valid and helpful for any model, including the models in the presence of nonideal baths that we discuss here—see, e.g., the work of Yang and Qian [8].

[3]The sum in Eq. (2) above can of course be an integral for a continuous state space.

To see the effects of finite bath size, we switch to a more general supersystem derivation. As shown in Ref. [3], consider a large system-plus-bath supersystem constrained to have a fixed total energy E_{tot}, since this supersystem is assumed isolated. In the absence of more information, MaxEnt says that the probability distribution over supersystem microstates is uniform over all allowed states. Every supersystem microstate with energy E_{tot} is equally probable. Then, we sum/integrate over the degrees of freedom of the bath, giving the system distribution as

$$P(x) \propto \Omega_B[E_{tot} - E(x)] = \exp\{\ln \Omega_B[E_{tot} - E(x)]\}, \quad (4)$$

where $\Omega_B(E_B)$ is the number of bath microstates with energy E_B.

Here is where the infinite-bath-size approximation comes in. If we now assume that fluctuations in the system energy are very small compared to the energy in the bath, we can Taylor expand $\ln \Omega_B$ around equilibrium, using the microcanonical thermodynamic definition of inverse temperature:

$$\beta = \left.\frac{d\ln \Omega_B(E_B)}{dE_B}\right|_{E_B = E_{tot} - \langle E\rangle}. \quad (5)$$

Dropping terms of higher order in system energy fluctuations, we can now write Eq. (4) as

$$P(x) \propto \exp\{\ln \Omega_B[E_{tot} - \langle E\rangle] - \beta[E(x) - \langle E\rangle]\}, \quad (6)$$

which, once properly normalized, is the BD. Later, we will include the next term in the Taylor expansion of $\ln \Omega_B$ to more accurately describe a system in contact with a large-but-finite heat bath.

B. Fokker-Planck equation from maximum caliber

Having reviewed the equilibrium properties of a system in an ideal thermal bath, we turn to the dynamics of such a system. In this section, we derive the FP equation using MaxCal. We seek the temporal evolution of a probability distribution $p(x,t)$:

$$\frac{\partial p(x,t)}{\partial t} = \lim_{\epsilon \to 0}\frac{1}{\epsilon}[p(x,t+\epsilon) - p(x,t)]$$
$$= \lim_{\epsilon \to 0}\frac{1}{\epsilon}\left[\int dx' \, p(x',t)\frac{P(x',t;x,t+\epsilon)}{P(x')} - p(x,t)\right], \quad (7)$$

where we have written the transition probability for a short time-step ϵ in terms of $P(x',t;x,t+\epsilon)$—the equilibrium probability to follow a trajectory from state x' at time t to state x at time $t+\epsilon$—and $P(x)$, the equilibrium state probability distribution. What remains is to determine the form of $P(x',t;x,t+\epsilon)$ to first order in ϵ, which calls for the use of MaxCal.

There are three relevant constraints: (i) knowledge of the equilibrium probability $P(x) = \int dx' \, P(x,t;x',t+\epsilon)$ for all values of x, which we will impose with Lagrange multiplier $\mu(x)$; (ii) microscopic reversibility $P(x',t;x,t+\epsilon) = P(x,t;x',t+\epsilon)$ for all x and x', which we will impose with Lagrange multiplier $\alpha(x,x')$; and (iii) a diffusion constraint $\int dx \, dx' \, P(x',t;x,t+\epsilon)(x-x')^2 = 2D\epsilon$, ensuring the continuity of motion, which we will impose with Lagrange

multiplier γ. Collecting this information, the unmaximized Caliber functional is[4]

$$
\begin{aligned}
\mathcal{C} = &-\int dx \int dx' \, P(x, x') \ln\left[\frac{P(x, x')}{1/\mathcal{N}}\right] \\
&+ \int dx \, \mu(x)\left[P(x) - \int dx' \, P(x, x')\right] \\
&+ \int dx \int dx' \, \alpha(x, x')[P(x, x') - P(x', x)] \\
&+ \gamma\left[2D\epsilon - \int dx \int dx' \, P(x, x') \, (x' - x)^2\right], \quad (8)
\end{aligned}
$$

where $1/\mathcal{N}$ is the uniform prior, and we note that only the antisymmetric part of $\alpha(x, x')$ will matter, so without loss of generality we just take it to be antisymmetric. Now we vary $\mathcal{C}$ with respect to the probabilities and set this variation equal to zero to derive the optimal expression for the probabilities $P(x, x')$:

$$
\begin{aligned}
0 = \frac{\delta\mathcal{C}}{\delta P(x, x')} = &-\ln\left[\frac{P(x, x')}{1/\mathcal{N}}\right] - 1 \\
&- \mu(x) + 2\alpha(x, x') - \gamma(x' - x)^2 \\
\Rightarrow P(x, x') = &\frac{e^{-1}e^{-\mu(x)}e^{2\alpha(x,x')}e^{-\gamma(x'-x)^2}}{\mathcal{N}}. \quad (9)
\end{aligned}
$$

Now, we apply constraints to elucidate the connections between the Lagrange multipliers and physical quantities. First, the symmetry and reversibility constraint easily eliminates the antisymmetric function $\alpha(x, x')$:

$$
P(x, x') = P(x', x) \Rightarrow 2\alpha(x, x') = \frac{1}{2}[\mu(x) - \mu(x')]
$$

$$
\Rightarrow P(x, x') = R(x)R(x')\sqrt{\frac{\gamma}{\pi}}e^{-\gamma(x'-x)^2}, \quad (10)
$$

where we have absorbed some constants and functions into the new function $R(x)$ without loss of generality, and we now see a normalized Gaussian for $x' - x$ contained in $P(x, x')$. For a general time span ϵ, it might be difficult to proceed, but since we are taking ϵ to zero, we know from the diffusion constraint that only very very small values of $x' - x$ should be accessible with any reasonable probability, and thus taking ϵ to zero is equivalent to taking γ to infinity. Thus we will make use of the singular expansion of a normalized Gaussian around zero variance, namely,

$$
\sqrt{\frac{\gamma}{\pi}}e^{-\gamma(x'-x)^2} = \delta(x' - x) + \frac{1}{4\gamma}\delta''(x' - x) + O(1/\gamma^2), \quad (11)
$$

where we will not need terms beyond the first two. This crucial observation allows us to easily incorporate the diffusion constraint for general $P(x)$; integrating $P(x, x')$ over x' using the expansion of Eq. (11) yields an expression for $R(x)$ in terms of $P(x)$, to zeroth and first orders in $1/\gamma$. Then, the diffusion

---

[4]The time coordinates are omitted during this derivation for the purpose of more concise notation.

constraint yields the relation $\gamma = 1/4D\epsilon$ as ϵ goes to zero. Eq. (10) becomes

$$
\begin{aligned}
P(x', t; x, t + \epsilon) = &P(x')\delta(x - x') - \epsilon DP'(x')\delta'(x - x') \\
&+ \epsilon DP(x')\delta''(x - x') + O(\epsilon^2). \quad (12)
\end{aligned}
$$

Plugging this back into Eq. (7) yields

$$
\frac{\partial p(x, t)}{\partial t} = -D\frac{\partial}{\partial x}\left[p(x, t)\frac{\partial}{\partial x}\ln P(x)\right] + D\frac{\partial^2 p(x, t)}{\partial x^2}, \quad (13)
$$

which completes the MaxCal derivation of the FP.

It is important to note how this derivation differs from previous derivations of the FP [2]. Traditionally, one begins with a master equation and then *assumes* the transition rates are very narrowly peaked Gaussians, possibly employing the Kramers-Moyal expansion and the Pawula Theorem—the FP then follows. Here, however, the transition rates are *derived* to be narrowly peaked Gaussians—using the MaxCal procedure with the diffusion constraint.

Note that substituting the BD [Eq. (3)] into Eq. (13) yields Eq. (1), the well-known form of the FP equation in a thermal bath. Below, we will discuss how to extend this derivation to incorporate a dynamical fluctuating local bath, albeit one which fluctuates much more quickly than the system, so that the system sees the bath as effectively almost stationary. For the derivation below, we will need to fully understand the appearance of timescales in our setup.

Although we know that it sets the timescale for a system, the diffusion coefficient contains nontemporal units—it has dimensions of [system dimension]^2/[time]. Wishing to extract just the timescale, we will divide the diffusion coefficient by an appropriate squared system dimension; specifically, we define here the time constant κ associated with a diffusion coefficient D as $\kappa = D/\sigma^2$, where σ^2 is the equilibrium variance of the system variable. Other choices exist, but this has a clear interpretation as the inverse timescale for the system to relax to equilibrium from a perturbation. Furthermore, the precise definition of time constant is not so important as long as we use the same definition when comparing two different systems, which we will be doing below with our system of interest and the local bath surrounding it. It will become evident that the effects of the dynamical bath become important when the time constant for the bath is nonnegligible compared to the time constant for the system.

III. BEYOND IDEAL THERMAL RESERVOIRS WITH MAXIMUM CALIBER

A. Limitation (A): Infinite-sized heat bath

In the previous section, we derived Eq. (4), the equilibrium distribution for a system in contact with a generic bath, about which we have no specific information. Performing a Taylor expansion of $\ln \Omega_B$ for fluctuations in the system energy that are small compared to the bath energy, we arrived at Eq. (6) keeping only the leading order term, which is equivalent to the BD. What if we include the next term, the term of order

$[E(x) - \langle E \rangle]^2$? Then we get a slight variation of the BD,

$$P(x) \propto \exp\left[\ln \Omega_B(E_{\text{tot}} - \langle E \rangle)) - \beta[E(x) - \langle E \rangle] \right.$$
$$\left. - \frac{\beta^2}{2C_B(\beta)}[E(x) - \langle E \rangle]^2 \right], \tag{14}$$

in which the variance of the energy fluctuations has also been constrained, where $C_B(\beta) = E'_B(T = 1/\beta)$ is the (microcanonical) heat capacity of the bath at inverse temperature β.[5]

Thus, for a system in contact with a finite-sized heat bath, the equilibrium probability distribution is no longer the simple BD, with its characteristic exponential energy dependence. For a very-large-but-finite bath, there is a quadratic correction,

$$P(x) \propto e^{-\beta E(x)}\left[1 - \frac{s}{2}\left(\frac{[E(x) - \langle E \rangle]^2}{\sigma_E^2} - 1 \right) + O(s^2) \right], \tag{15}$$

where we introduce the coefficient $s = C_S/C_B$—the ratio of system heat capacity (C_S) to bath heat capacity (C_B)—and σ_E^2 is the variance of the system energy.[6] Note that—although our derivation is physically illuminating—Eq. (15) could be derived directly from MaxEnt with constraints on both the average energy and energy variance of the system, so even if we cannot directly measure the heat capacity of the environment, the measured energy variance of the system can determine this correction. Equation (15) describes the effects of a finite bath on equilibria [9,10].

Now, what is the effect of a finite bath on *nonequilibria*? As discussed above, for a system relaxing to the BD, we recover the familiar gradient of the energy landscape drift term in the FP. However, in a large-but-finite heat bath, to first order in s, substituting Eq. (15) into Eq. (13) with $D = \kappa\sigma^2$—as explained above—yields

$$\frac{\partial p(x,t)}{\partial t} = \kappa\sigma^2 \frac{\partial^2 p(x,t)}{\partial x^2}$$
$$+ \kappa\sigma^2\beta \frac{\partial}{\partial x}\left[E'(x)\left(1 + s\frac{[E(x) - \langle E \rangle]}{\beta\sigma_E^2} \right)p(x,t) \right]. \tag{16}$$

This can be interpreted as an FP equation with a microstate-dependent temperature perturbation.

So far, x has been used as a generic microstate label. Let us now consider the specific case one-dimensional Brownian motion, in which the relevant microstate variable is the velocity v for a particle of mass m diffusing through a heat bath of inverse temperature β and size ratio s. The velocity probability distribution diffuses with coefficient $D = \kappa\sigma^2 =$
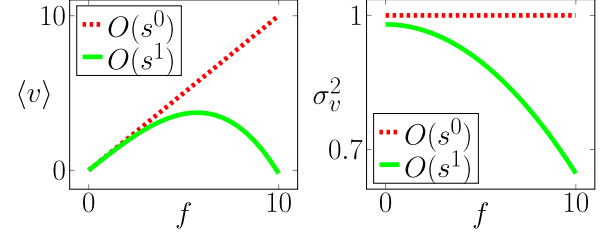

FIG. 1. Finite bath size: effect of large fluctuations. Dimensionless units are employed, and s is the ratio of system heat capacity to bath heat capacity—see text for more details. (Left) Mean velocity vs driving force for $s = 1/100$. (Right) Variance of velocity vs. driving force for $s = 1/100$.

$\kappa/\beta m$,[7] the energy is $E(v) = \frac{1}{2}mv^2$, and we add an extra drift term $-(F/m)p'(v,t)$ to our probability evolution equation corresponding to acting on the particle with a constant external force F, i.e., the particle would accelerate at rate F/m if it were not influenced by the heat bath. The equilibrium probability distribution with an infinite bath is a Gaussian with variance $1/\beta m$ and zero mean. Furthermore, the nonequilibrium steady state (NESS) with an infinite bath is a Gaussian having the same variance, but a different mean value, $\langle v \rangle_F = F/m\kappa$. Physically, an infinite bath implies a perfectly linear drag force with fluctuations that are always symmetric about the mean and independent of the driving force.

As an aside, we note a troublesome aspect of the concept of a finite bath for NESS. If the bath is finite, then an externally driven process will eventually drain or saturate the bath; the system can never truly be stationary, since the bath properties will be changing constantly. We will consider here only sufficiently large finite baths, in which there can be a relatively long transient period resembling a NESS.[8]

The NESS in the presence of a finite bath—i.e., the steady-state solution to Eq. (16) with added external force term—is proportional to $\exp(\beta F v/\kappa)$ times the finite-bath equilibrium distribution Eq. (15). We introduce a dimensionless version of the force $f = (F\sqrt{\beta})/(\kappa\sqrt{m})$, as well as absorbing a factor of $\sqrt{\beta m}$ into velocities v to make them dimensionless. To first order in s, we can write the NESS probability distribution as

$$P_f(v) = \frac{e^{-\frac{1}{2}(v-f)^2}}{\sqrt{2\pi}}\left\{ 1 - \frac{s}{4}[(v-f)^4 + 4f(v-f)^3 \right.$$
$$+ 2(f^2 - 2)(v-f)^2 + 4f(f^2 - 1)(v-f)$$
$$\left. -(6f^2 + 1)] \right\}. \tag{17}$$

In this form, it is simple to calculate central moments,

$$\langle v \rangle_f = f - sf(2 + f^2) + O(s^2),$$

$$\langle (v - \langle v \rangle_f)^2 \rangle_f = 1 - \frac{s}{3}(6 + f^2) + O(s^2), \tag{18}$$

---

[5]The form of Eq. (14) has appeared previously in the literature, see, e.g., Ref. [9] or Ref. [10].

[6]The system heat capacity C_S is defined as the derivative of the expectation value of the system energy with respect to temperature $T = 1/\beta$, i.e., $C_S(\beta) = d\langle E \rangle/dT = -\beta^2 d\langle E \rangle/d\beta$. To zeroth order in s, i.e., for the original BD, $C_S = \beta^2\sigma_E^2$; since we only care about $P(x)$ to first order in s here, no further s-dependent terms in C_S are relevant.

[7]This is *not* the same as the oft-used diffusion coefficient for the *position* of a particle; it is the analogous quantity that appears in Eq. (13) when the generic microstate label x refers to velocity rather than position.

[8]Alternatively, a true finite bath NESS can be achieved by an external agent cycling energy continuously between the system and the bath, rather than simply supplying it.
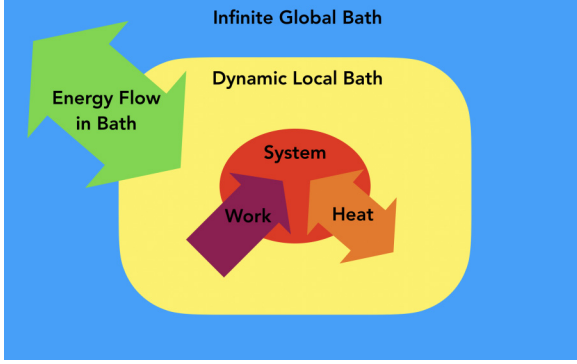
FIG. 2. Energy flow in dynamic local bath model.

and the skewness of the distribution is $-6sf + O(s^2)$. We see nonlinear corrections to the drag force, as well as asymmetric noise which decreases in strength for stronger forcing (see Fig. 1).

One important and strange feature is that the variance in velocity seems to go negative for high enough force, which is nonphysical. This signals a phase transition, beyond which a NESS is no longer supported—when the system is driven too strongly, eventually the dissipation into the finite bath is so great that the bath cannot take it. In an infinite-sized bath, it would never occur, but this sort of breaking-point behavior is known to appear in nature, such as when a spring is pulled too far, or in dielectric breakdown.

B. Limitation (B): Stationary heat bath

A second approximation for nonequilibrium models is that baths respond very rapidly to fluctuations in the system, such that they are practically stationary.[9] We begin to mitigate this limitation with a framework we call the dynamic local bath (DLB) model (see Fig. 2). The system is only coupled directly to an intermediate local bath, which itself couples to a truly infinite global bath. This introduces a finite time constant for the environment to respond to fluctuations.

In the DLB, we track not only the microstate x of the system, but also one extra variable: the total energy E_{tot} in the local-bath-plus-system super-system. Since the super-system is immersed in a perfect bath, its equilibrium probability distribution is the BD:

$$P(x, E_{\text{tot}}) \propto \Omega_B[E_{\text{tot}} - E(x)]e^{-\beta E_{\text{tot}}}, \qquad (19)$$

where $\Omega_B(E_B)$ is the energy degeneracy of the local bath.

Performing a very similar analysis to the MaxCal derivation of the FP found above, we study the dynamics of probability for the local-bath-plus-system super-system. We use the time constant κ associated with diffusion and fluctuations in the system, as well as introducing a new time constant κ_E associated with diffusion and fluctuations in the

super-system energy E_{tot}.[10] However, changes in x as a result of interactions with the DLB are assumed uncorrelated with changes in E_{tot} as a result of interactions at the far-away border between the DLB and the global bath, so there will be no cross-derivative term. MaxCal then yields the primary evolution equation for the DLB model:

$$
\frac{\partial p(x, E_{\text{tot}}, t)}{\partial t}
$$
$$
= \kappa\sigma^2 \frac{\partial}{\partial x}[\beta_B(E_{\text{tot}} - E(x))E'(x)p(x, E_{\text{tot}}, t)]
$$
$$
+ \kappa\sigma^2 \frac{\partial^2 p(x, E_{\text{tot}}, t)}{\partial x^2}
$$
$$
+ \kappa_E \sigma_B^2 \frac{\partial}{\partial E_{\text{tot}}}([\beta - \beta_B(E_{\text{tot}} - E(x))]p(x, E_{\text{tot}}, t))
$$
$$
+ \kappa_E \sigma_B^2 \frac{\partial^2 p(x, E_{\text{tot}}, t)}{\partial E_{\text{tot}}^2}, \qquad (20)
$$

where σ^2 and σ_B^2 refer to the equilibrium variance of x and E_{tot}, respectively, and we have defined the function $\beta_B(E_B) = \Omega'_B(E_B)$. The resemblance to Eq. (13)—to which this equation provides a minor generalization—should be clear: there are now simply diffusion and drift terms for each variable, and the drift terms came from Eq. (19), the equilibrium distribution for the super-system.

As a consistency check, Eq. (20) reduces to Eq. (13)—the FP for the system alone—when the dimensionless response ratio $r = \kappa/\kappa_E$ goes to zero, i.e., when the local bath is perfectly stationary. Importantly, we can go further and find corrections to Eq. (13) for an almost-but-not-perfectly-stationary bath. To first order in r, we compute the expectation value:

$$
\int dE_{\text{tot}}\, p(E_{\text{tot}}|x, t)\, \beta_B[E_{\text{tot}} - E(x)]
$$
$$
= \beta - r\frac{\sigma^2}{\sigma_B^2}E'(x)\left[\beta E'(x) + \frac{1}{p(x, t)}\frac{\partial p(x, t)}{\partial x}\right] + O(r^2). \qquad (21)
$$

The effective temperature depends not only on the microstate x but also on the macrostate, since it depends on the distribution $p(x, t)$. This correction is zero at equilibrium, i.e., when the probability distribution is the BD, and it will be small when the probability distribution is close to the BD.

Using Eq. (21), we integrate Eq. (20) over E_{tot} to find a corrected FP for fast-but-not-infinitely-responding baths, to

---

[9] By this we mean that the probability distribution of bath microstates is assumed to not change over time, even though the system may be fluctuating, because the bath is assumed to be unaffected by—i.e., uncorrelated with—the system.

[10] As a reminder of how these time constants are defined, they are the diffusion coefficients for some degree of freedom, with the equilibrium variance of that degree of freedom extracted, such that they only have units of inverse time. In this case, $\kappa_E \sigma_E^2 = \lim_{t \to 0}\langle \Delta E_{\text{tot}}^2 \rangle/t$, where σ_E^2 is the equilibrium variance of E_{tot}. This definition also relates to heat conductivity across the barrier between the local bath and the global bath: If the local bath and global bath are initially both in equilibrium at temperatures differing by an infinitesimal dT, then the responding heat current across the barrier when they come into contact will be $\kappa_E dT$.

first order in r:

$$\frac{1}{\kappa\sigma^2}\frac{\partial p(x,t)}{\partial t}$$
$$= \frac{\partial^2}{\partial x^2}\left\{\left[1 - r\frac{\sigma^2}{\sigma_B^2}E'(x)^2\right]p(x,t)\right\}$$
$$+ \frac{\partial}{\partial x}\left(\left\{\beta + r\frac{\sigma^2}{\sigma_B^2}[2E''(x) - \beta E'(x)^2]\right\}E'(x)p(x,t)\right).$$
$$(22)$$

This equation is the central result of this section. Note that previously nonexistent x-dependence appears in the diffusion term, slowing it down when energy flow is greater in magnitude. Another key result is the effect on the Einstein relation, wherein the diffusion coefficient is proportional to temperature. For Brownian motion in a DLB, we find

$$D = \frac{\kappa}{\beta m}\left[1 - \frac{r}{C_B(\beta)} + O(r^2)\right]. \quad (23)$$

Unless the heat capacity of the DLB $C_B(\beta)$ is independent of temperature, the diffusion coefficient will no longer be purely proportional to temperature.

Now, consider the effect of applied force. The dimensionless velocity and force and are defined as before, and a quantity s_L is defined analogously to the quantity s above, except s_L is the ratio of system heat capacity to heat capacity of just the local bath. For arbitrary force f, we can add a force-induced drift term to Eq. (22) and solve for the NESS:

$$P_f(v) = \frac{e^{-\frac{1}{2}(v-f)^2}}{\sqrt{2\pi}}\left\{1 + \frac{1}{3}rs_Lf[(v-f)^3\right.$$
$$\left. + 3f(v-f)^2 + 3f^2(v-f) - 3f] + O(r^2)\right\}. \quad (24)$$

From this we can calculate the moments,

$$\langle v \rangle_f = f + rs_Lf(1+f^2) + O(r^2),$$
$$\langle (v - \langle v \rangle_f)^2 \rangle_f = 1 + 2rs_Lf^2 + O(r^2), \quad (25)$$

and skewness $2rs_Lf + O(r^2)$. Again we see nonlinear drag, as well as asymmetric noise that depends on the driving force (see Fig. 3).

It is interesting to note, as well, that the effects of a DLB depend not only on the responsiveness of the local bath—through r—but also on how large the local bath is—through s_L. In future work, we would like to extend the DLB concept to multiple layers of intermediate baths, each with its own size and characteristic response time, to more accurately describe certain real systems; the work here is an important first step in that direction.

Fluctuation relations in nonequilibria, such as those of Crooks and Jarzynski, are based on ratios of forward to backward trajectory probabilities [11,12]. In the NESS studied
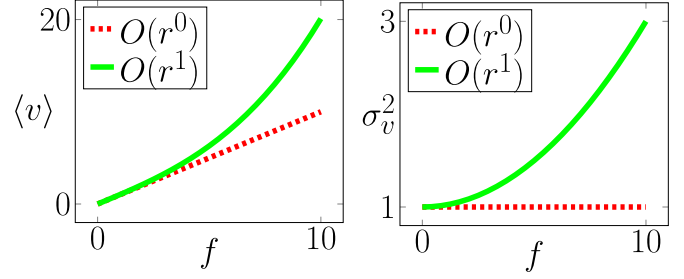

FIG. 3. Finite bath speed: effects of fast system fluctuations. Dimensionless units are employed, r is the ratio of system response speed to local bath response speed, and s_L is the ratio of system heat capacity to local bath heat capacity—see text for more details. (Left) Mean velocity vs driving force for $rs_L = 1/100$. (Right) Variance of velocity vs driving force for $rs_L = 1/100$.

here, we find

$$\ln\left[\frac{P_f(v)}{P_f(-v)}\right] = 2fv + \frac{2}{3}rs_Lfv^3 + O(r^2). \quad (26)$$

This logarithmic asymmetry factor is directly proportional to dissipated power for a perfect bath, but we see a nonlinear correction for a realistic environment. Although this is not exactly the form of the canonical Crooks formulation, we see that fluctuation relations can depend on bath speed (as well as bath size [10]).

IV. SUMMARY AND CONCLUSIONS: THE MEANING OF "FAR-FROM-EQUILIBRIUM"

How should we define *far from equilibrium* (FFE)? Often this means that forces and flows are large, compared to some relevant scales, and then force-flow relations are nonlinear. But we note here two new important considerations: the ratio of system size to bath size (here quantified as s), to account for the bath's ability to handle large fluctuations, and the ratio of system speed to bath speed (here quantified as r), to account for how quickly the bath can respond to fast changes in the system.

We saw that, although these quantities can have negligible effects near-equilibrium, they are increasingly important under larger driving forces. Here, we have derived quantitative first-order corrections for nonequilibrium statistical processes in the presence of nonideal thermal baths. We have also extracted insight into significant qualitative differences in the example case of driven Brownian motion, including fundamentally different steady states with breaking-point phase changes in baths of finite size and dependence of diffusion on energy flow rate in baths of finite speed.

[1] K. Ghosh, P. D. Dixit, L. Agozzino, and K. A. Dill, Annu. Rev. Phys. Chem. **71**, 213 (2020).

[2] N. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland personal library (North-Holland, Amsterdam, 1981).

[3] S. Pressé, K. Ghosh, J. Lee, and K. A. Dill, Rev. Mod. Phys. **85**, 1115 (2013).

[4] P. D. Dixit, J. Wagoner, C. Weistuch, S. Pressé, K. Ghosh, and K. A. Dill, J. Chem. Phys. **148**, 010901 (2018).

[5] J. Shore and R. Johnson, IEEE Trans. Info. Theory **26**, 26 (1980).

[6] A. Caticha, Entropy **23**, 895 (2021).

[7] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).

[8] Y.-J. Yang and H. Qian, J. Stat. Phys. **182**, 46 (2021).

[9] J. D. Ramshaw, Phys. Rev. E **98**, 020103(R) (2018).

[10] M. Campisi, P. Talkner, and P. Hänggi, Phys. Rev. E **80**, 031145 (2009).

[11] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).

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