Approximating the entire spectrum of nonequilibrium steady-state distributions using relative entropy: An application to thermal conduction

Puneet Kumar Patra
Advanced Technology Development Center, Indian Institute of Technology Kharagpur, India, 721302

Marc Meléndez*
Dpto. Física Fundamental, UNED, Madrid, Spain, 28040

Baidurya Bhattacharya†
Department of Civil Engineering, Indian Institute of Technology Kharagpur, India, 721302

(Received 23 September 2014; revised manuscript received 27 June 2015; published 17 August 2015)

Distribution functions for systems in nonequilibrium steady states are usually determined through detailed experiments, both in numerical and real-life settings in the laboratory. However, for a protocol-driven distribution function, it is usually prohibitive to perform such detailed experiments for the entire range of the protocol. In this article we show that distribution functions of nonequilibrium steady states (NESS) evolving under a slowly varying protocol can be accurately obtained from limited data and the closest known detailed state of the system. In this manner, one needs to perform only a few detailed experiments to obtain the nonequilibrium distribution function for the entire gamut of nonlinearity. We achieve this by maximizing the relative entropy functional (MaxRent) subject to constraints supplied by the problem definition and new measurements. MaxRent is found to be superior to the principle of maximum entropy (MaxEnt), which maximizes Shannon's informational entropy for estimating distributions but lacks the ability to incorporate additional prior information. The MaxRent principle is illustrated using a toy model of ϕ^4 thermal conduction consisting of a single lattice point. An external protocol controlled position-dependent temperature field drives the system out of equilibrium. Two different thermostatting schemes are employed: the Hoover-Holian deterministic thermostat (which produces multifractal dynamics under strong nonlinearity) and the Langevin stochastic thermostat (which produces phase-space-filling dynamics). Out of the 80 possible states produced by the protocol, we assume that four states are known to us in detail, one of which is used as input into MaxRent at a time. We find that MaxRent approximates the phase-space density functions for every value of the protocol, even when they are far from the known distribution. MaxEnt, however, is unable to capture the fine details of the phase-space distribution functions. We expect this method to be useful in other external protocol-driven nonequilibrium cases as well, making it unnecessary to perform detailed experiments for all values of the protocol when one wishes to obtain approximate distributions.

DOI: 10.1103/PhysRevE.92.023304 PACS number(s): 05.10.Gg, 05.20.Gg, 05.45.Pq

I. INTRODUCTION

One of the challenges of nonequilibrium thermodynamics lies in the unified statistical description of nonequilibrium systems [1]. These systems are usually characterized by the presence of nonvanishing currents (like heat, mass, etc.) due to the imposed thermodynamic forces (like temperature gradient, concentration gradient, etc.). A sound theoretical understanding of these nonequilibrium processes is vital for gaining insights into several important physical and biological processes. The nonequilibrium systems may be classified into systems near the equilibrium regime (local thermodynamic equilibrium) and systems far-from-equilibrium. Unlike in the equilibrium and near-equilibrium scenario [2,3], no unified theoretical framework exists for the far-from-equilibrium steady-state (NESS) cases [4]. The problem becomes all the more difficult since there is no consensus on the variables necessary to describe the nonequilibrium states [5,6]. As a result, limited statistical descriptions of NESS are usually obtained through *either numerical simulations* [7–10] *or experiments* [11–13]. Experimental techniques have matured enough to perform nanoscale experiments under externally controlled protocols, like the diffusion of a few particles in response to a variation in the concentration gradient, and single-particle reaction dynamics by varying the energy barrier between two successive equilibrium positions controlled by laser traps.

Usually, the limited information from such experiments is used to construct the probability distribution of the relevant variable through the Jaynes principle of maximum entropy (MaxEnt) [14–22]. Let a set of macroscopic observables $\langle F_1\rangle,\langle F_2\rangle,...,\langle F_N\rangle$ and its corresponding phase functions $F_1(\Gamma),F_2(\Gamma),...,F_N(\Gamma)$ be given to us. The phase-space $\Gamma = (x_1,x_2,...,x_n,p_1,p_2,...,p_n)$ represents a point in phase space, comprising the position x_i and momentum p_i of every particle. MaxEnt finds the least biased probability distribution, $\bar{\rho}(\Gamma)$, by maximizing the Shannon entropy functional (H) subjected to the constraints imposed by the set of macroscopic observables, which for continuous states can be written as:

$$H \equiv -k_B \int_\Gamma \rho(\Gamma)\log[\rho(\Gamma)]d\Gamma - \sum_{j=1}^{N}\lambda_j\left[\int_\Gamma F_j(\Gamma)\rho(\Gamma)d\Gamma - \langle F_j\rangle\right], \quad (1)$$

*mmelendez@fisfun.uned.es
†baidurya@civil.iitkgp.ernet.in

 ©2015 American Physical Society

The integration is carried out over the entire accessible phase-space Γ, k_B is the Boltzmann constant, and λ_j represents the Lagrange multiplier corresponding to the observable $\langle F_j \rangle$. The approximated probability distribution $\bar{\rho}(\Gamma)$ is obtained by setting the first variation of H, Eq. (1), equal to zero,

$$\bar{\rho}(\Gamma) = \frac{1}{Z} \exp\left[-\sum_j \lambda_j F_j(\Gamma) \right]. \tag{2}$$

In Eq. (2), Z is the generalized partition function which arises from the normalization of the probability density function and is given by

$$Z = \int_\Gamma \exp\left[-\sum_j \lambda_j F_j(\Gamma) \right] d\Gamma. \tag{3}$$

The macroscopic observables, $\langle F_j \rangle$, are related to the partition function through $\langle F_j \rangle = \partial \log Z / \partial \lambda_j$. To complete the approximated distribution, one needs to find the unknown Lagrange multipliers. This procedure is similar to the well-known problem of moments [23,24]. For the equilibrium scenario, the Lagrange multipliers can be interpreted in physical terms without any numerical computation. For example, when the external constraint is chosen as the average internal energy $\langle U \rangle$ of the system, the corresponding Lagrange multiplier can be identified as $\lambda_{\langle U \rangle} = 1/k_B T$, where T is the temperature, and one obtains a canonical distribution. Similarly other equilibrium distribution functions can be obtained by judiciously choosing the constraints. MaxEnt has been successfully employed toward constructing NESS distribution functions for steady-state thermal conduction [7–10]. The flexibility offered by the MaxEnt formalism has rendered it useful in situations beyond statistical physics [25–27].

However, the MaxEnt formalism suffers from the inability of incorporating any information other than that of constraints of the type $\langle F_i \rangle$ [28]. Let us say that we have the complete distribution function of the system at some value of the external protocol for the experiments discussed before, and suppose that the protocol is altered resulting in a new steady state for which very limited information is obtained by performing a not-so-detailed experiment. If we now wish to use MaxEnt to estimate the current NESS distribution function, we will not be able to utilize the detailed information of the previous steady state mentioned above. Rather, we will have to confine ourselves to using only the limited experimental results obtained in the current state and all previous knowledge on the system (obtained possibly at a significant cost) will go unutilized.

In this work, we show that by maximizing the relative entropy functional (MaxRent) one eliminates the problem described in the previous paragraph. One need not perform detailed experiments for all the values of the protocol. With limited data on the system at all states, and a detailed probability description of one or a few of these states, it is possible to construct the nonequilibrium distribution functions for the entire range of nonlinearity. We demonstrate the efficacy of MaxRent using a toy model of thermal conduction: a ϕ^4 thermal conducting model consisting of a single lattice point brought out of equilibrium by a position-dependent temperature field, whose strength evolves slowly through external control. We use two different schemes to impose the temperature field: the Hoover-Holian thermostat [29] and the Langevin thermostat [30]. Using the statistics at an earlier steady state and a few simple constraints at the current state that have direct physical meaning out of equilibrium, we demonstrate that, as the protocol evolves, MaxRent is able to approximate the true distribution function, $\rho_t(\Gamma)$, much more accurately than MaxEnt. The approach proposed in this study simplifies the life of the experimentalists who would no longer need to perform time-consuming detailed experiments for all values of the protocol.

II. THE THERMAL CONDUCTION MODEL

The toy model employed in this study is a ϕ^4 model [31,32] comprising a single lattice point and governed by the potential

$$U(x) = \tfrac{1}{4} x^4. \tag{4}$$

Here, x denotes the displacement from the equilibrium position (the origin). The mass of this single quartic oscillator is taken to be one. This model is chosen because it shows all the qualitative features of a system with many degrees of freedom and yet is simple to analyze. The system is brought out of equilibrium by imposing a position-dependent temperature field [33–36]:

$$T(x) = 1 + \epsilon \tanh(x). \tag{5}$$

The parameter ϵ acts as the external protocol and defines the strength of nonlinearity. A value of zero indicates equilibrium (canonical) while a positive value indicates a nonequilibrium state. A far-from-equilibrium scenario occurs when ϵ is large (but less than 1). For simplicity we have taken Boltzmann constant to be unity. The imposed temperature field would cause the heat flow to occur in a direction opposite to the temperature gradient. In this work, ϵ is varied from 0 to 0.80 in steps of 0.01. The temperature is imposed on the system through the Langevin stochastic thermostat [30] (LT) and the Hoover-Holian deterministic thermostat [29] (HH). The LT equations of motion for this case are

$$\dot{x} = p, \quad \dot{p} = -x^3 - \phi p + \psi, \tag{6}$$

where the damping constant ϕ and the stochastic force ψ are related through

$$\langle \psi(t) \rangle = 0, \quad \langle \psi(t)\psi(t') \rangle = 2\phi T(x) \delta(t - t'). \tag{7}$$

In the present study $\phi = 0.5$. The equations of motion for the HH thermostat are

$$\begin{aligned} &\dot{x} = p, \quad \dot{p} = -x^3 - \eta p - \xi p^3, \quad \dot{\eta} = p^2 - T(x), \\ &\dot{\xi} = p^4 - 3T(x)p^2. \end{aligned} \tag{8}$$

The variables η and ξ represent reservoir variables that control the first and the second moments of kinetic energy, respectively. The equations of motion, Eqs. (6) and (8), are integrated using the classical fourth-order Runge-Kutta method for 1 billion time steps with incremental time step $\Delta t = 0.001$ for every realization of ϵ, except for four realizations where we solve 10 billion time steps (detailed later). The two thermostats result in different dynamics: LT produces a phase-space-filling dynamics even under large values of ϵ [37] as shown in Fig. 1,

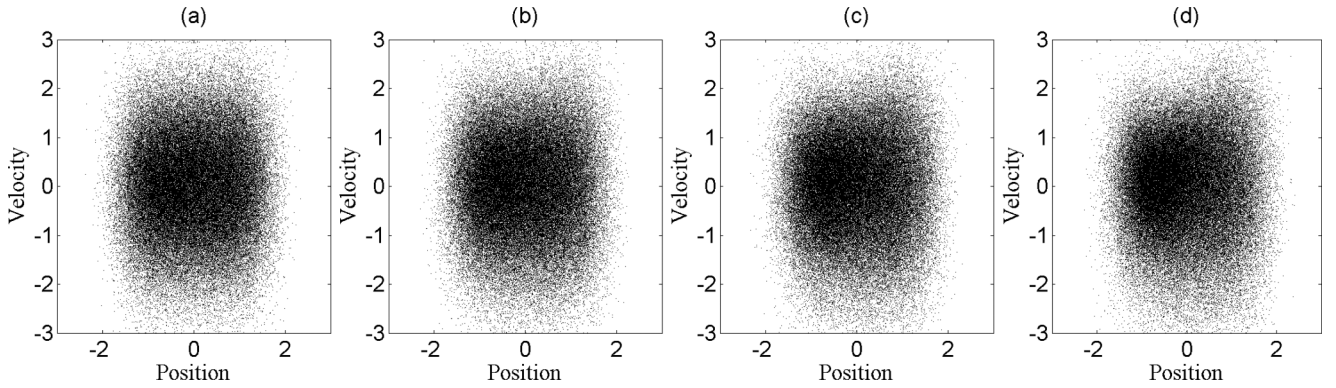

FIG. 1. Phase-space trajectory for Langevin dynamics. (a) Corresponds to $\epsilon = 0.25$; (b) corresponds to $\epsilon = 0.50$; (c) corresponds to $\epsilon = 0.74$; and (d) corresponds to $\epsilon = 0.98$. The phase-space-filling nature of the dynamics does not change with increasing the value of ϵ.

while HH results in a dynamics that show a gradual shift from being phase-space-filling to a limit cycle through intermediate multifractal dynamics [36] as shown in Fig. 2. We would like to stress that LT equations of motion cannot correctly simulate the nonequilibrium process [38], but nevertheless we use LT to show that MaxRent can approximate $\rho_t(\Gamma)$ both for phase-space-filling as well as multifractal dynamics.

It is interesting to see that in both the thermostats considered, the energy current ($\dot{E}$) is independent of the configurational variables (x), as shown in Eq. (9). Consequently, there is no contribution of potential energy transport toward heat current:

$$\begin{aligned}\dot{E} &= \frac{d}{dt}\left[\frac{1}{4}x^4 + \frac{1}{2}p^2\right] = x^3\dot{x} + p\dot{p} \\ &= \phi p^2 + \psi p \quad \text{for} \quad \text{LT} \\ &= \eta p^2 + \xi p^3 \quad \text{for} \quad \text{HH}.\end{aligned} \tag{9}$$

Since we limit the focus of this study to NESS and we make use of the ergodic hypothesis later on to estimate the density function from a single time trajectory, it is necessary to ascertain if these thermostatted dynamics can produce steady-state conditions under the imposed temperature gradient. Figure 3 plots the absolute value of heat flux for both the HH as well as LT dynamics for four different values of the protocol. It can be clearly seen that the system reaches steady state very soon after the nonequilibrium conditions are imposed. Throughout the rest of the paper, we assume that the steady state has set in so that $\partial\rho/\partial t = 0$.

Previous attempts to theoretically understand the distribution functions corresponding to the nonequilibrium states have invariably involved stochastic dynamics [39,40]. For such cases, it is possible to use the large deviation functional together with the fluctuation theorem to develop a functional form of the cumulants of heat current. One can then obtain the different moments from the theoretical cumulants and use them as constraints to construct the entire distribution function. As an outcome, the approximated distribution function is usually smooth. However, the nonequilibrium steady states (for deterministic dynamics) under strong nonlinearity are usually characterized by multifractal dynamics with information dimension smaller than the phase-space dimension [41]. As a result, the support of the distribution function and the underlying structure of the phase-space dynamics cannot be predicted beforehand. Because of this lack of knowledge, developing cumulant-based constraints is challenging. Thus, for the problem in hand, we reconstruct the nonequilibrium distribution functions through numerical techniques rather than the theoretical. The reconstructed distributions can be shown numerically to satisfy the fluctuation theorem but we do not explicitly explore this connection.

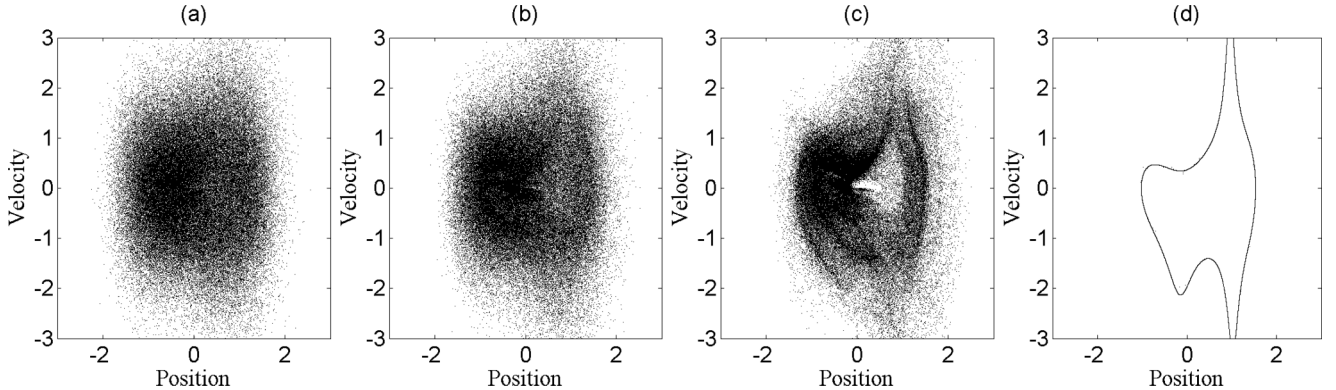

FIG. 2. Phase-space trajectory for Hoover-Holian dynamics. (a) Corresponds to $\epsilon = 0.25$; (b) corresponds to $\epsilon = 0.50$; (c) corresponds to $\epsilon = 0.74$; and (d) corresponds to $\epsilon = 0.98$. Both position and velocity are initialized at 1. The nature of dynamics changes from being phase-space-filling to a limit cycle through multifractals as ϵ increases.

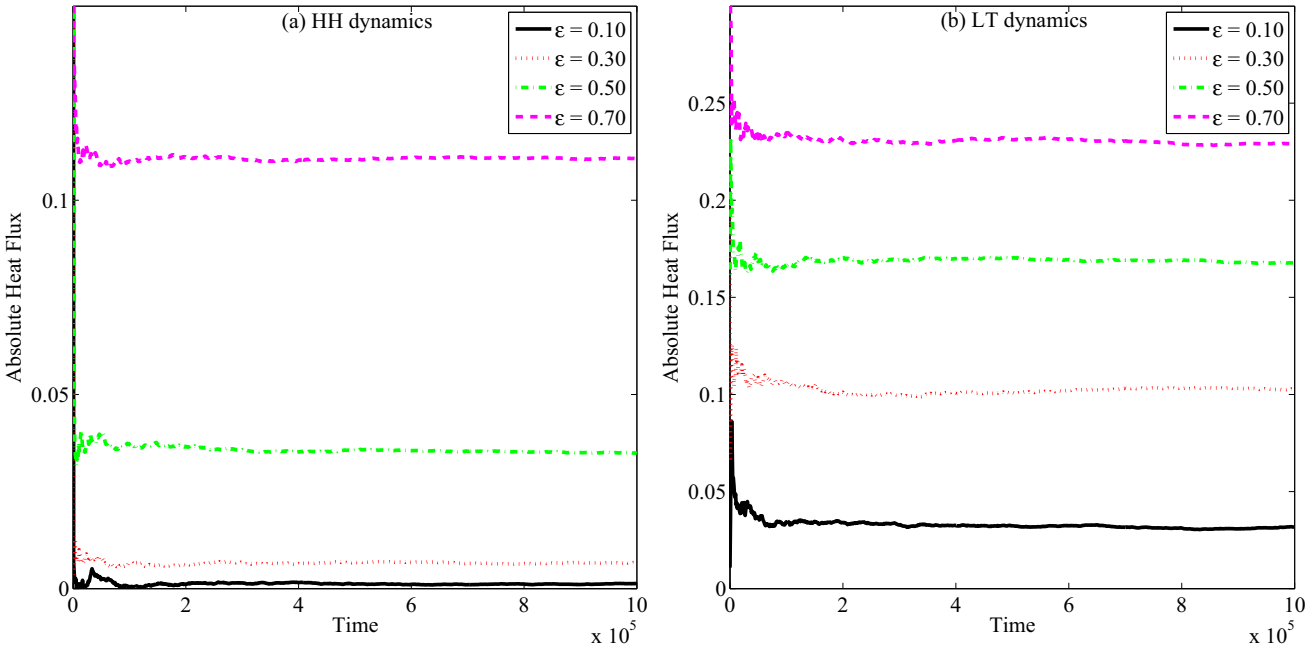

FIG. 3. (Color online) Check for steady-state condition using four different values of the protocol ϵ: (a) for HH dynamics and (b) for LT dynamics. In steady state, there is no appreciable change in the value of heat flux with time. The figures indicate that the system reaches a steady state under the imposed condition fairly early in the simulation.

III. PRINCIPLE OF MAXIMUM RELATIVE ENTROPY

Given a set of constraints $\langle F_j \rangle = F_{j,0}$ for the present state of a system and a previously known state of the same system $\rho'(\Gamma)$, the principle of maximum relative entropy (MaxRent) states that the probability density $\bar{\rho}(\Gamma)$ that best approximates the true distribution can be obtained by maximizing the relative entropy subject to the constraints on the average value. The mathematical foundations of MaxRent can be found in Refs. [16,42–45]. It has been argued that MaxRent is the only general tool for updating probabilities [46,47]. MaxRent has seen varied uses like tweaking a model system to reproduce target properties [48], ecological applications like predicting equilibrium distribution of species [49], reconstruction of genetic networks [50], inferring fatigue crack damage [51], etc. In conjunction with the constraints, their corresponding Lagrange multipliers and the known state of the system, the entropy functional to be maximized for MaxRent becomes

$$C \equiv -k_B \int_\Gamma \left(\rho(\Gamma) \log \left[\frac{\rho(\Gamma)}{\rho'(\Gamma)} \right] \right) d\Gamma - \sum_{j=1}^{N} \lambda_j \left[\int_\Gamma F_j(\Gamma) \rho(\Gamma) d\Gamma - \langle F_j \rangle \right]. \quad (10)$$

The first term (without $-k_B$) is the Kullback-Leibler divergence [47,52] that measures the discrepancy between the distributions ρ and ρ', and is always $\geqslant 0$, with the equality holding only if $\rho = \rho'$ everywhere except, perhaps, at points that form a zero measure set. The MaxRent-approximated distribution function, $\bar{\rho}(\Gamma)$, can be obtained by setting the first variation of Eq. (10) to zero:

$$\bar{\rho}(\Gamma) = \frac{1}{Z} \rho'(\Gamma) \exp \left[-\sum_j \lambda_j F_j(\Gamma) \right]. \quad (11)$$

One needs to find the unknown Lagrange multipliers, as in the MaxEnt method, to completely identify the approximated distribution function. In general, $\rho'(\Gamma)$ represents a known distribution, which in this case pertains to a known detailed state of the system. The approximated distribution function $\bar{\rho}(\Gamma)$ represents the best "guess" at the true distribution function, $\rho_t(\Gamma)$, while simultaneously satisfying the constraints $\langle F_j \rangle$ and differing as little as possible from the known state of the system. When there is no prior information available, it makes sense to preassign equal probability to each state (i.e., a uniform density function) and the formulation boils down to MaxEnt Eq. (2).

A. Selection of prior

The difference between MaxRent and MaxEnt comes from the additional ability of the former to handle prior information of any type. The success of MaxRent lies in careful selection of the prior distribution and identification of constraints not captured in the reference distribution [25]. In fact, it has recently been argued that equilibrium statistics can be inferred much more easily by using MaxRent than MaxEnt [52]. The utility of MaxRent is not just limited to equilibrium scenario. MaxRent, with an innovative prior distribution scheme, has recently been used to estimate the steady-state probability of a given gene to regulate other genes, thereby reconstructing the gene regulatory network at a coarse level [50].

However, for the problem in hand, the multifractal nature of the dynamics renders it very difficult to have any correct estimate on the prior distribution function. The nonequilibrium distribution is *never* known beforehand. Therefore, we resort to utilizing the experimental data to construct a prior distribution function. Instead of choosing an equilibrium prior distribution function, we select a nonequilibrium prior distribution function from among the states whose information is known to us in detail. We assume that the detailed information for four steady states (out of the 80 possible) corresponding to $\epsilon = 0.20, 0.40, 0.60$, and 0.80 is given to us in the form of the first 20 moments of positions and velocities. The numerical values of these moments are obtained by solving the equations of motion for 10 billion time steps. These 40 moments are then utilized in the MaxEnt framework to construct the prior distribution of position and velocity. Each of the four states represents a valid choice of prior. In the following we describe in detail the constraints on the system that have been used for the remaining 76 cases.

B. Identifying constraints for NESS thermal conduction

The performance of MaxRent depends on the choice of constraints that are usually problem dependent. For example, in Eq. (9) the functional form of energy current is different for the LT and HH thermostats. Theoretically, it is possible to recover the *exact distribution* if all joint moments of x and p, i.e., $\langle x^m p^n \rangle$ for $m,n \geqslant 0$, are known. However, it is not possible to obtain such information from either simulations or experiments and only a few moments are generally available. Additionally, higher-order moments are usually associated with large errors and take a substantial time to converge. Therefore, in this study we have used only those constraints that are intuitive and have an order less than six. Such constraints bear a direct physical relevance and can be measured in laboratory settings easily.

The simplest constraint obviously is the normalization property for density functions, i.e., $\int_\Gamma \rho(\Gamma) d\Gamma = 1$. The other constraints are listed below.

No net mass current: Absence of net mass current in this system, i.e., average velocity of the system is zero:

$$\langle p \rangle = \int_\Gamma p \rho(x,p) dx dp = 0. \tag{12}$$

Fixed averaged position: An implication of Eq. (12) is that the averaged position of the system must be time invariant and must equal some constant. Depending on the steady-state position of the system, this constraint can be written as

$$\langle x \rangle = \int_\Gamma x \rho(x,p) dx dp = x_0. \tag{13}$$

Heat flux and flow of potential energy: A set of constraints can be obtained from the definition of steady-state systems: macroscopic properties of the system relax to fixed average values [53]. In the present context, the relevant macroscopic property is the heat current, which can be decomposed into two parts [54,55]: (i) transport of kinetic energy, and (ii) transport of potential energy. As has been shown before, the energy current is not dependent on configurational variables in this problem [see Eq. (9)], and consequently, the rate at which kinetic energy is transferred is a proxy for average heat flux (J_0) and is a valid choice of constraint:

$$\int_\Gamma \frac{p^3}{2} \rho(x,p) dx dp = J_0. \tag{14}$$

In fact, in conventional molecular dynamics, Eq. (14) represents the "heat flux." Since the contribution of configurational variables toward energy current is absent, the averaged rate at which potential energy is transported from left to right (or vice versa) must be zero, and therefore,

$$\int_\Gamma \frac{x^4 p}{4} \rho(x,p) dx dp = 0. \tag{15}$$

Higher-order derivatives of steady-state quantities: One can once again invoke the definition of steady-state systems to argue that if Eq. (14) is time invariant, then its subsequent time derivatives must be zero; i.e.,

$$\begin{aligned} HH &: \int_\Gamma (-p^2 x^3 - \eta p^3 - \xi p^5) \rho(x,p) dx dp = 0, \\ LT &: \int_\Gamma (-p^2 x^3 - \phi p^3 + \psi p^2) \rho(x,p) dx dp = 0. \end{aligned} \tag{16}$$

One can simplify Eq. (16) by recognizing that ϕ is a constant (=0.5) and ψ is independent of p, with its average being equal to zero. Many subsequent time derivatives can also be equated to be zero. However, we do not consider these constraints because they are usually associated with substantial errors and are slow to converge.

Flux fluctuations: The average fluctuations of heat flux from its mean value, once the steady state is reached, is constant. In our particular problem, the flux fluctuations assume a simple form and can be implemented easily. Let the instantaneous flux be denoted by $J(=p^3/2)$, then these fluctuations may be represented as

$$\begin{aligned} &\int_\Gamma [J - J_0]^2 \rho(x,p) dx\, dp = J_1, \\ &\Rightarrow \int_\Gamma J^2 \rho(x,p) dx\, dp - J_0^2 = J_1, \\ &\Rightarrow \int_\Gamma p^6 \rho(x,p) dx\, dp = J_2. \end{aligned} \tag{17}$$

The J_is are obtained from the numerical experiments. Similarly, the skewness and kurtosis of flux may also be evaluated. But due to the poor convergence of these higher-order constraints, we keep ourselves limited to Eq. (17).

Averaged work done by the system: Another constraint can be obtained by recognizing that in the steady state averaged work done (w_0) by the external tethering potential must be time invariant; i.e.,

$$\langle x^4 \rangle = \int_\Gamma x^4 \rho(x,p) dx dp = w_0. \tag{18}$$

Energy constraint: So far we have only chosen those constraints that can be measured easily in laboratory settings. We now place a constraint on the spatial variation of energy:

$$U(x) = \int_\Gamma H(q,p) \delta(q-x) \rho(q,p) dp dx. \tag{19}$$

Equation (19) is the limiting case for $U(a \leqslant x \leqslant b)$ as $a \to b$. In the next subsection, we show that this constraint is associated with a Lagrange multiplier, $\beta(x)$, whose functional form is given by $1/T(x)$. This constraint is essential for MaxEnt to give sensible results even though local thermodynamic equilibrium (LTE) may not be valid. For MaxRent, this constraint is not suitable for LT dynamics since it does not constrain the temperature accurately. Imposing this constraint results in poor performance.

C. Solving for the Lagrange multipliers

In this section we obtain the functional form of the Lagrange multiplier $\beta(x)$ for both the MaxEnt and MaxRent methods. We assume that the thermodynamic temperature is applicable locally:

$$\frac{1}{T(x)} = \left[\frac{\delta S}{\delta U}\right]_x, \tag{20}$$

where S is the Gibbs entropy. The Gibbs entropy corresponding to the MaxEnt-approximated distribution Eq. (2) becomes

$$\begin{aligned} S[\bar{\rho}] &= -\int \bar{\rho}(x,p)\ln(\bar{\rho}(x,p))dxdp \\ &= \int \bar{\rho}(x,p)\left(\sum_j \lambda_j F_j + \beta(x)H(x,p)\right)dxdp \\ &= \int \bar{\rho}(x,p)\left(\sum_j \lambda_j F_j\right)dxdp \\ &\quad + \int_x\int_\Gamma \bar{\rho}(q,p)[\beta(q)H(q,p)]\delta(q-x)dqdpdx \\ &= \int \bar{\rho}(x,p)\left(\sum_j \lambda_j F_j\right)dxdp + \int \beta(x)U(x)dx. \end{aligned} \tag{21}$$

In Eq. (21), we have omitted the integration constant. Taking variation of Eq. (21) with respect to $U(x)$ we get

$$\beta(x) = \frac{1}{T(x)}. \tag{22}$$

Now, let us consider the case where ρ' is not uniform. We assume that ρ' has the functional form of $\rho' = \exp\left[\sum_i \lambda_{\rho',i} Y_i(x,p) - \beta_{\rho'}(x)H(x,p)\right]$, where $Y_i(x,p)$ is a polynomial function of (x,p). Using the approximated distribution Eq. (11), the Gibbs entropy becomes

$$\begin{aligned} S(\bar{\rho}) &= -\int \bar{\rho}(x,p)\ln\bar{\rho}(x,p) \\ &= -\int \bar{\rho}(x,p)\left\{\sum_i \lambda_{\rho',i}Y_i + \sum_j(\lambda_j + \lambda_{\rho',j})F_j \right. \\ &\quad \left. + [\beta(x) + \beta_{\rho'}(x)]H(x,p)\right\}dxdp. \end{aligned} \tag{23}$$

The Lagrangian multipliers $\lambda_{\rho',j}$ and $\beta_{\rho'}(x)$ correspond to the prior distribution. Proceeding like before, we can see that the relationship between $\beta(x)$ and $T(x)$ becomes

$$\beta(x) = \frac{1}{T(x)} - \beta_{\rho'}(x). \tag{24}$$

The remaining unknown Lagrange multipliers λ_j are solved using the Newton-Raphson technique. All necessary integrals are performed numerically using the trapezoidal rule. Sampling was confined to the $[-3,3]$ interval for both the position and velocity.

IV. APPROXIMATING THE NESS DISTRIBUTIONS UNDER AN EVOLVING PROTOCOL

We now show that MaxRent can be used to approximate the entire spectrum of the nonequilibrium distribution functions for the quartic oscillator under the position-dependent temperature field. We proceed with maximizing Eq. (10) subject to the constraints highlighted before. Out of the four detailed states, whichever lies closest to the present state is taken as the prior distribution function. In simple terms it means that if we want to approximate the distribution function for $\epsilon_0 = 0.49$, then we choose the state that is closest to ϵ_0, i.e., $\epsilon = 0.40$ as the prior distribution function. The distributions at each value of the protocol ϵ are compared to the one obtained through the MaxEnt framework. The spatial resolution is taken to be 0.01, and the true distributions (wherever computed) are calculated using 100,000 position-velocity data points.

For simplicity, in the case of HH thermostat we have assumed that the variables η, ξ, and p are independent of each other (just like in equilibrium). Consequently, we replace $\langle \eta p^3 \rangle$ with $\langle \eta \rangle \langle p^3 \rangle$ and $\langle \xi p^5 \rangle$ with $\langle \xi \rangle \langle p^5 \rangle$, respectively, for the constraint Eq. (16). Since in this work we are not interested in the distribution of the thermostat variables, we replace their averages with numeric constants. This replacement does not alter the nature of position and velocity distributions drastically. The approximated distributions obtained by replacing $\langle \eta \rangle = \langle \xi \rangle = 1.0$ and $\langle \eta \rangle = \langle \xi \rangle = 2.0$ overlap each other, suggesting that the distributions are independent of the constant value chosen. The reason may be attributed to the negligible contribution of the terms $\langle \eta p^3 \rangle$ and $\langle \xi p^5 \rangle$. For the rest of the analysis, we proceed with $\langle \eta \rangle = \langle \xi \rangle = 1.5$.

A. Marginal position and velocity distributions

We begin by studying the ability of MaxRent to correctly describe the probability distributions of position and velocity (see Figs. 4 and 5) in three different regimes for both the thermostats: (i) near equilibrium (small temperature gradient with $\epsilon = 0.10$), (ii) moderately away from equilibrium (moderately large temperature gradient with $\epsilon = 0.30$), and (iii) far-from-equilibrium (large temperature gradient with $\epsilon = 0.70$).

The black lines in Figs. 4 and 5 indicate the true distributions, $\rho_t(x)$ and $\rho_t(p)$ [henceforth, referred to as $\rho_t(.)$], obtained directly from the simulation. The true position distribution function, $\rho_t(x)$, in Fig. 4 loses its symmetric nature even under a small temperature gradient. As ϵ increases, the asymmetry in $\rho_t(x)$ also increases. It occurs because of the higher velocity of the oscillator when $x > 0$, due to higher temperature associated with this region. Consequently, the

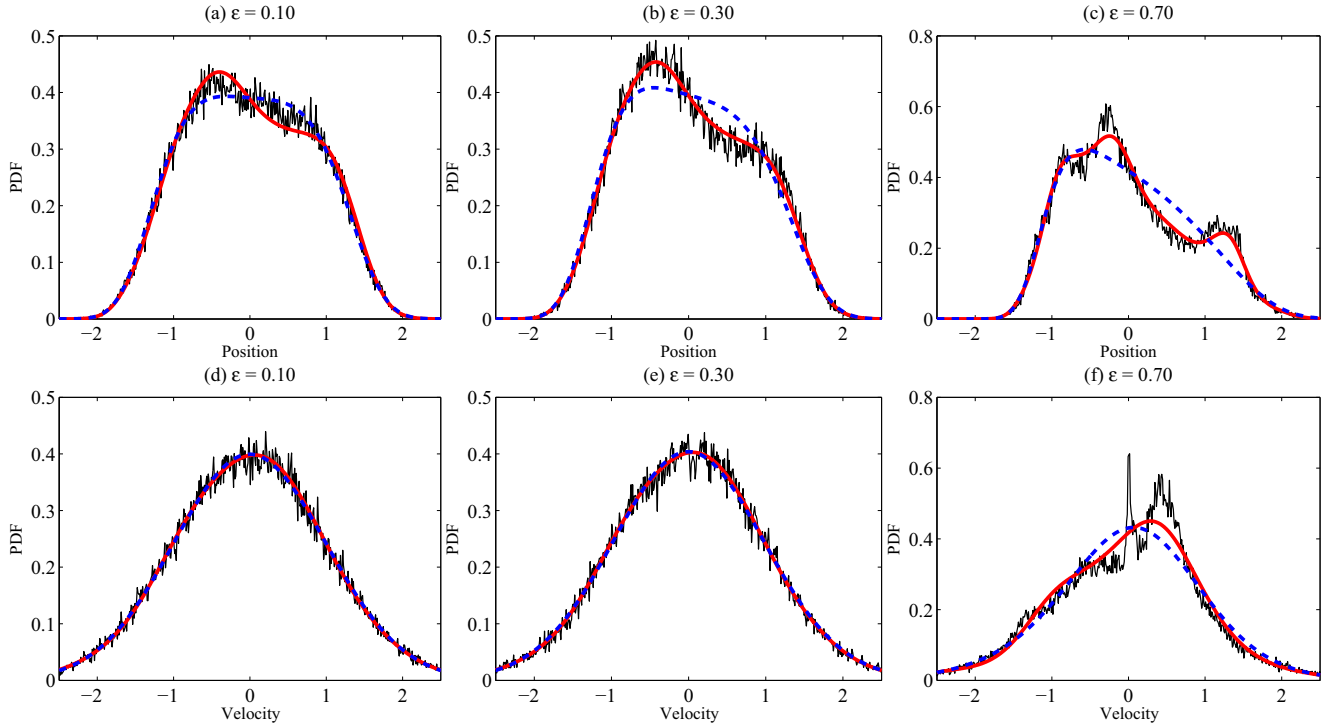

FIG. 4. (Color online) True (black thin line), MaxRent (red thick line), and MaxEnt (blue dashed line) distribution functions of position (top row) and velocity (bottom row) for HH dynamics. Figures (a) and (d) correspond to $\epsilon = 0.10$. Figures (b) and (e) correspond to $\epsilon = 0.30$, and figures (c) and (f) correspond to $\epsilon = 0.70$. MaxRent accurately approximates and preserves the important features of the true distribution function. MaxEnt always returns a smooth distribution.

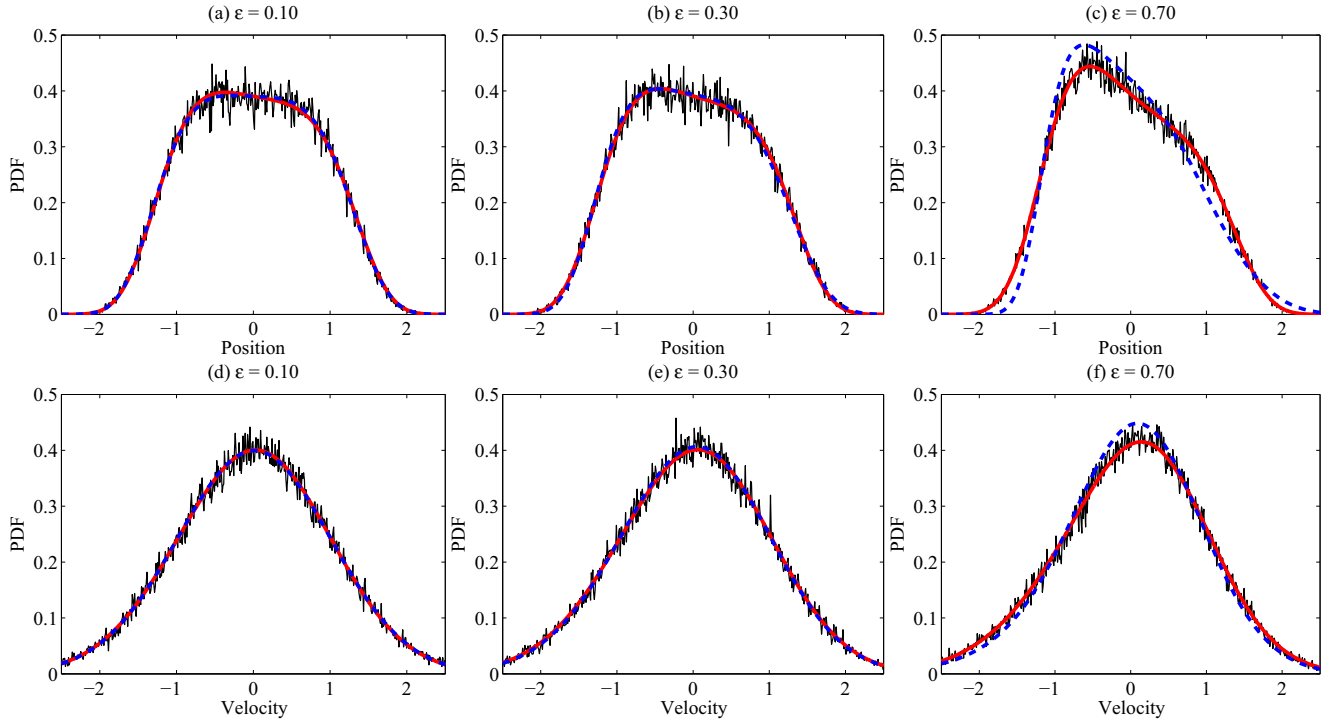

FIG. 5. (Color online) True (black thin line), MaxRent (red thick line), and MaxEnt (blue dashed line) distribution functions of position (top row) and velocity (bottom row) for LT dynamics. Figures (a) and (d) correspond to $\epsilon = 0.10$. Figures (b) and (e) correspond to $\epsilon = 0.30$, and figures (c) and (f) correspond to $\epsilon = 0.70$. MaxRent accurately approximates and preserves the important features of the true distribution function.

oscillator spends significantly less time in this region and hence, the asymmetry.

At larger values of ϵ, several kinks in $\rho_t(x)$ can be observed. These kinks are a characteristic of multifractal dynamics where some of the regions of phase space are not visited as frequently as others [see Fig. 2(c)]. On the other hand, $\rho_t(p)$ in Fig. 4 are associated with negligible asymmetry until moderately large values of ϵ. At larger ϵ, however, $\rho_t(p)$ shows the presence of kinks along with asymmetry. A faithful approximation to the true distribution functions must be able to capture these essential features. In Fig. 4, we can clearly observe that MaxRent is able to approximate the true distribution functions faithfully for all strengths of nonlinearity. Under large temperature gradient [see Figs. 4(b) and 4(c)], MaxRent accurately captures the peaks and the dips of the distribution as well as the asymmetry. On the other hand, MaxEnt washes out all the important features of $\rho_t(x)$ and $\rho_t(p)$, including the kinks. Its performance is rather poor when dealing with multifractal dynamics.

For LT dynamics (see Fig. 5), there are no kinks owing to the phase-space-filling dynamics. At larger values of ϵ, we again observe asymmetry in the position-distribution function. In this case as well, MaxRent performs better than MaxEnt in approximating $\rho_t(.)$. For HH dynamics (Fig. 4), all three distribution functions overlap at large absolute value of the independent variable. Interestingly, the same does not occur for LT dynamics. In Fig. 5(c), there is a significant deviation between the distribution functions at large values of position. This is because the LT thermostat does not sample the imposed temperature accurately, and consequently, the use of energy constraint in MaxEnt leads to inconsistent distributions. At this stage we would like to point out that the performance of MaxEnt (as well as MaxRent) will improve if more constraints were chosen.

B. Quantifying the difference from the true distribution function

The superiority of MaxRent becomes clear when we calculate the difference between the approximated distributions and $\rho_t(\Gamma)$. We use two functions for this purpose: (i) modified form of the Kullback-Leibler divergence, with $\rho_t(i)$ being the true marginal distribution corresponding to the variable i and $\bar{\rho}_i$ being the approximated marginal distribution corresponding to the variable i,

$$D_{\mathrm{KLD}}(\bar{\rho}||\rho_t) = \sum_x \bar{\rho}(x) \log\left[\frac{\bar{\rho}(x)}{\rho_t(x)}\right] + \sum_p \bar{\rho}(p) \log\left[\frac{\bar{\rho}(p)}{\rho_t(p)}\right], \tag{25}$$

and (ii) the absolute difference defined through

$$D_{\mathrm{ABS}}(\bar{\rho}||\rho_t) = \sqrt{\sum_x [\bar{\rho}(x) - \rho_t(x)]^2 + \sum_p [\bar{\rho}(p) - \rho_t(p)]^2}. \tag{26}$$

The comparison is shown in Fig. 6. The deviation of MaxRent-approximated distributions from the true distributions is quite consistent for the better part of the spectrum. At larger gradients, though the difference increases, the results are still better than the MaxEnt-distribution functions. The results suggest that if slight errors are permissible, MaxRent can be used to approximate the unknown true distributions with limited data with better accuracy than MaxEnt. We would

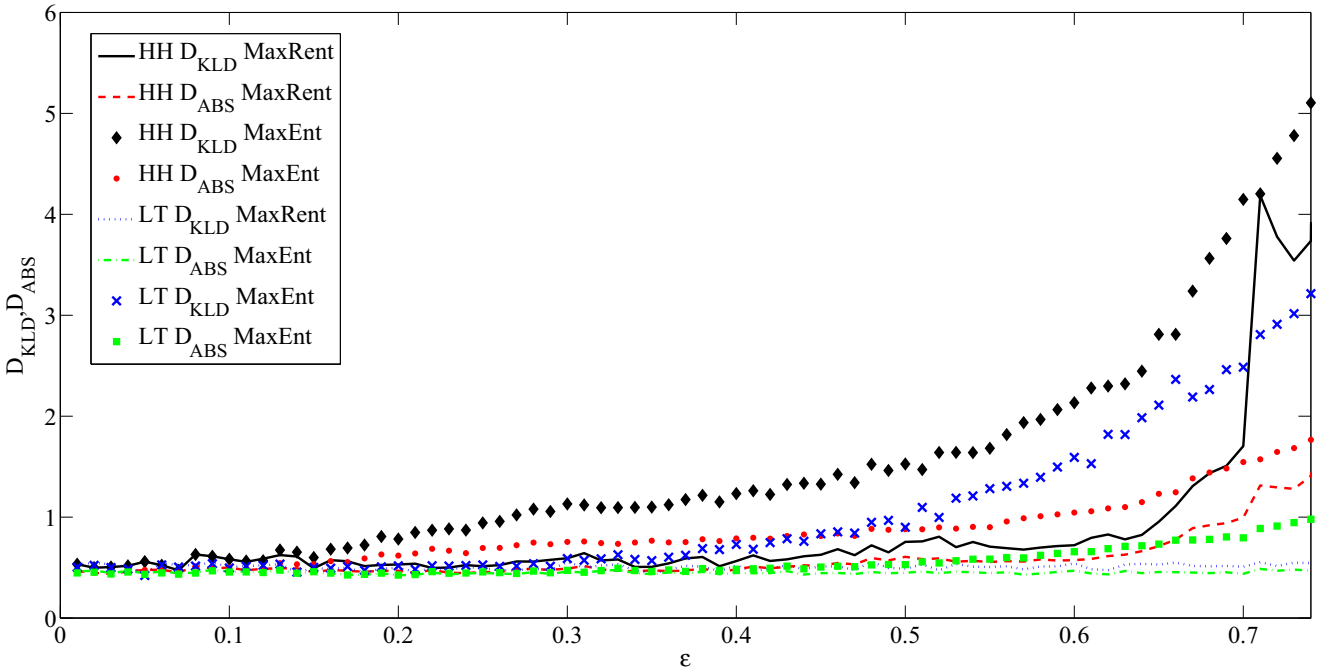

FIG. 6. (Color online) Comparison of Kullback-Leibler divergence Eq. (25) and absolute error Eq. (26) for the MaxRent- and MaxEnt-approximated distribution functions. Lines indicate the value of the metrics for MaxRent, while dots indicate the values for MaxEnt. Apart from cases very close to equilibrium cases ($\epsilon \leqslant 0.04$), the deviation from true distribution is greater for MaxEnt than MaxRent. The deviation increases with increasing value of ϵ for MaxEnt, while it remains fairly constant over a large range of ϵ for MaxRent.

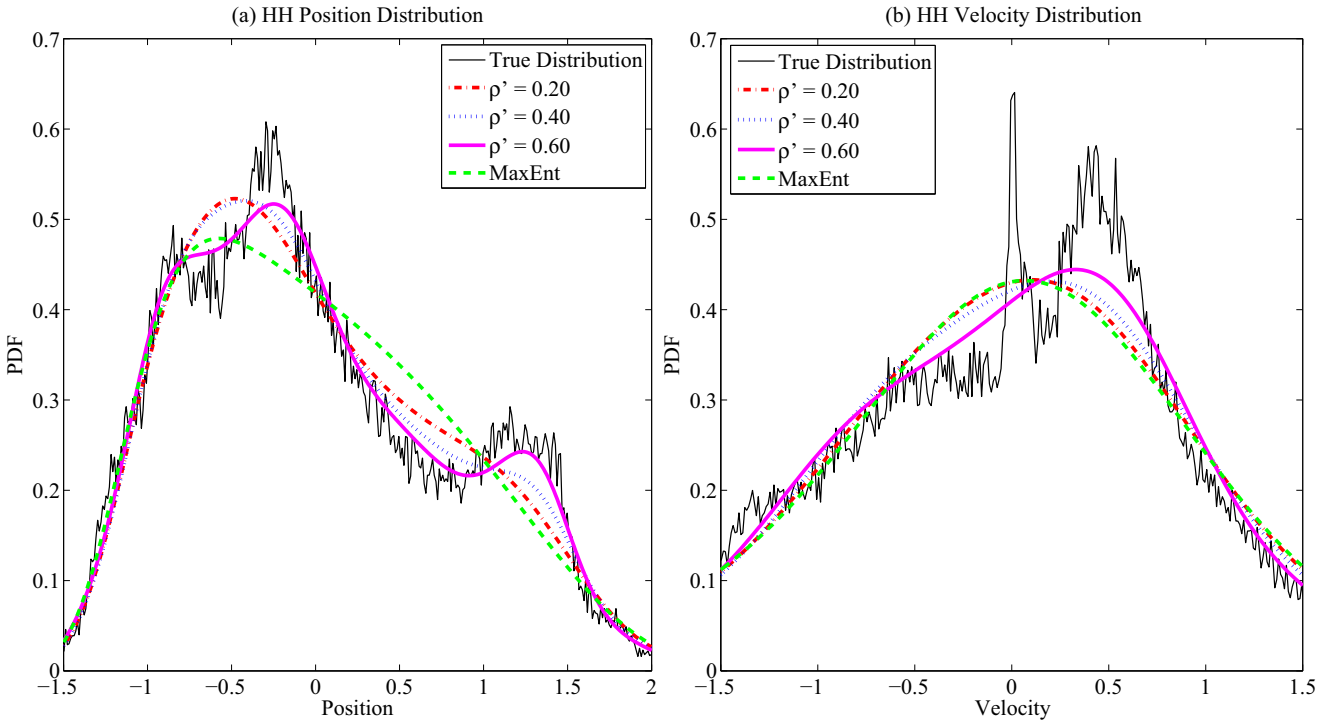

FIG. 7. (Color online) Distribution functions for HH dynamics with external protocol $\epsilon = 0.70$ using different prior distributions: (a) marginal distribution of position and (b) marginal distribution of velocity. MaxRent results in a better approximation than MaxEnt for all the cases considered. But, the performance of MaxRent decreases as the prior distribution moves further away from the unknown true distribution.

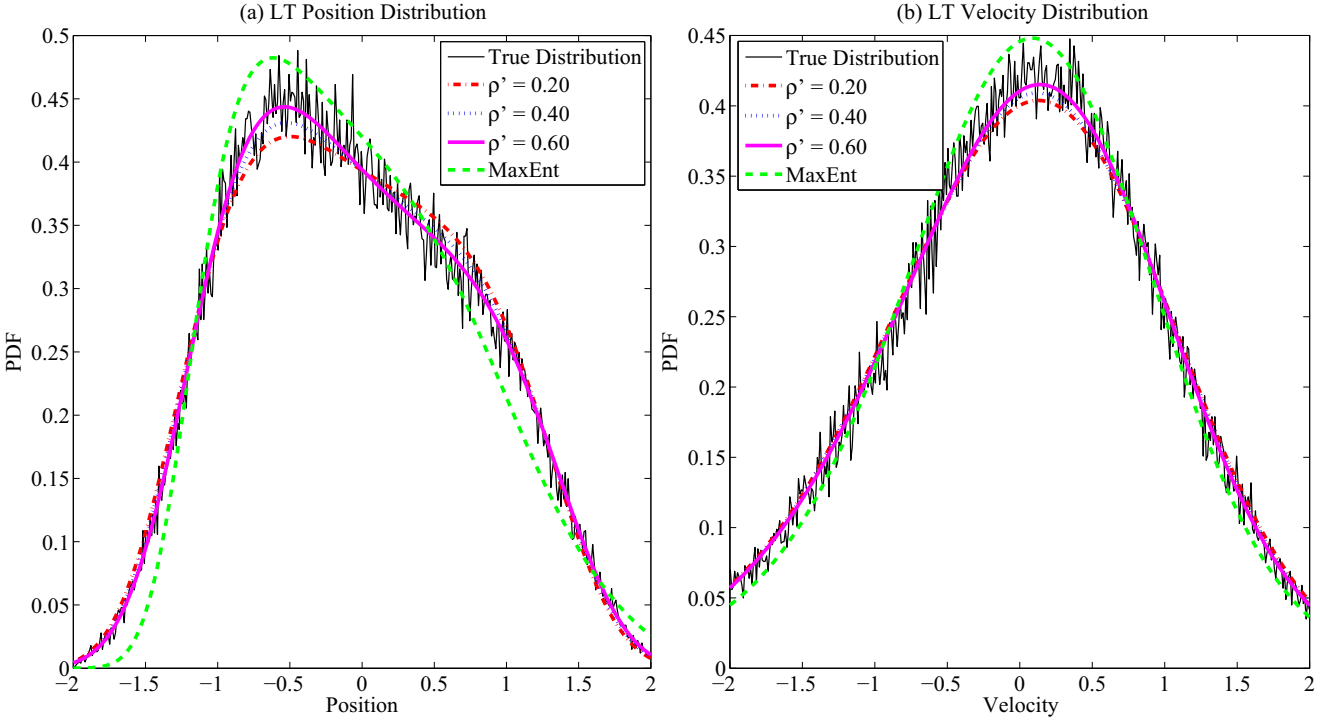

FIG. 8. (Color online) Distribution functions for LT dynamics with external protocol $\epsilon = 0.70$ using different prior distributions: (a) marginal distribution of position and (b) marginal distribution of velocity. MaxRent results in a better approximation than MaxEnt for all the cases considered. MaxEnt approximation is substantially further away from the true distributions at larger values of x.

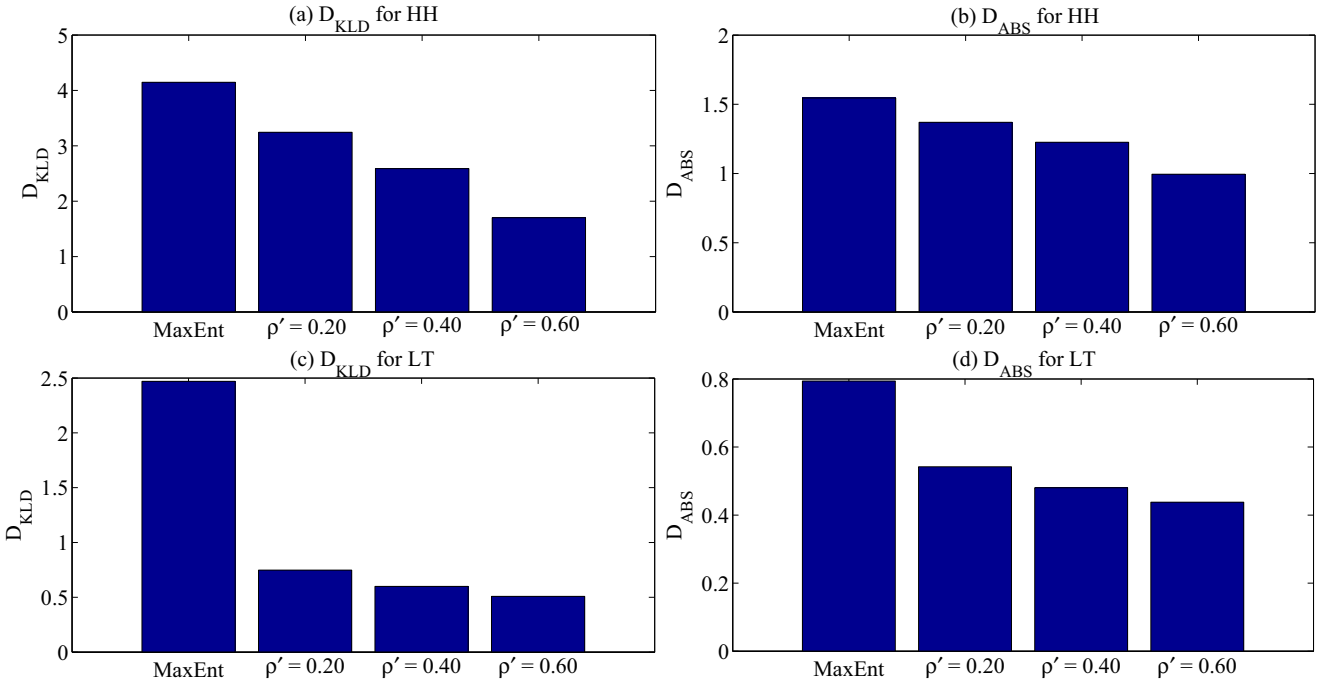

FIG. 9. (Color online) Comparison of (left) D_{KLD} and (right) D_{ABS} for different prior distributions due to (top row) HH dynamics and (bottom row) LT dynamics with $\epsilon = 0.70$. The closer the bars are to zero, the better the approximation to the true distribution function. MaxEnt results in the worst approximation.

again like to stress that the performance will improve upon inclusion of more constraints for both MaxRent and MaxEnt.

C. Effect of changing the prior distribution

So far we have used the closest among the four known states for calculating the approximated distribution function. It could be argued that the efficacy of MaxRent is due to selecting a prior distribution close to the unknown true distribution. Selection of a good prior distribution is crucial for the performance of MaxRent. In this section, we study how the approximated distributions change by changing the prior distribution. The effect is illustrated at $\epsilon = 0.70$. The resulting approximated distributions for the HH and LT thermostatted dynamics are shown in Figs. 7 and 8, respectively.

As expected, for HH dynamics, the closer the prior distribution is to $\rho_t(\Gamma)$, the better the approximation due to MaxRent. However, even with prior distributions as far away as $\epsilon = 0.20$, MaxRent results in a better approximation than MaxEnt and is able to capture the features of the true distribution. A similar behavior was observed for LT dynamics as well, although the difference among the different approximated distributions was not too large to begin with. The approximated distributions due to the different prior distributions (for both HH and LT) are compared in Fig. 9 using the two functions, Eqs. (25) and (26). The error due to MaxRent decreases monotonically as the protocol comes closer to the current state.

V. CONCLUSIONS

In this work, we show through numerical examples that in a protocol controlled NESS it is not mandatory to perform detailed experiments for every value of the protocol to obtain statistics for the entire range of nonlinearity. We demonstrate this procedure for the toy model of a ϕ^4 chain comprising a single lattice point. The system is brought out of equilibrium through a protocol-driven and position-dependent temperature field. This temperature field is imposed using two different thermostats. Using limited data in the form of simple constraints and detailed description (obtained from simulations) of a few of the states is sufficient to approximate the distribution functions for the entire range of nonlinearity. To do so, we maximize the relative entropy functional, Eq. (10), with the method of Lagrange multipliers using the Newton-Raphson technique. The detailed description is assumed to be given in the form of the distribution function for 4 of the 80 states that the protocol explores.

The MaxRent-approximated distribution functions are compared with those of the true and MaxEnt-approximated distribution functions. We observe that MaxRent represents a marked improvement over the MaxEnt framework. For smaller values of ϵ, both the MaxRent and MaxEnt perform comparably. However, at larger ϵ, MaxRent approximates the true distributions in a much better way. In the absence of any prior information available, MaxRent defaults to MaxEnt. For the problem in hand, we observe that if MaxEnt were to perform comparably with MaxRent, then many more constraints would be needed, implying that one must perform detailed experiments for every value of the protocol.

MaxRent has the ability to reproduce the true distribution in an accurate way because of the additional information that gets incorporated through the prior distribution function. However, its choice is subjective. The MaxRent-approximated distribution functions have a remarkable dependence on the

choice of the prior distribution. Likewise, the choice of constraints also impacts the performance of MaxRent. One must be careful to use appropriate constraints and prior distributions with MaxRent. Lastly, we would like to stress that although the method highlighted in this paper is for the simple case of a quartic oscillator, we believe that the method is useful in other nonequilibrium cases and would ease the life of the experimentalists, who now need not perform costly experiments for marginal changes in the external agent.

ACKNOWLEDGMENT

The authors thank Professor William G. Hoover for providing stimulating ideas and critically appraising the first draft of this work.

[1] D. Jou, C. Perez-Garcia, and J. Casas-Vázquez, J. Phys. A: Math. Gen. **17**, 2799 (1984).
[2] S. R. De Groot and P. Mazur, *Nonequilibrium Thermodynamics* (Courier Dover Publications, New York, 2013).
[3] D. Jou, J. Casas-Vázquez, and G. Lebon, *Extended Irreversible Thermodynamics* (Springer, Berlin, 1996).
[4] I. Vavruch, Chemické listy **96**, 271 (2002).
[5] J. Casas-Vázquez and D. Jou, Rep. Progr. Phys. **66**, 1937 (2003).
[6] S. Goldstein and J. L. Lebowitz, Phys. D: Nonlin. Phenom. **193**, 53 (2004).
[7] A. B. Corbet and H. J. Morowitz, Phys. Rev. A **6**, 2298 (1972).
[8] A. B. Corbet, Phys. Rev. A **9**, 1371 (1974).
[9] B. N. Miller and P. M. Larson, Phys. Rev. A **20**, 1717 (1979).
[10] R. M. Nisbet and W. S. C. Gurney, Phys. Rev. A **10**, 720 (1974).
[11] K. Ghosh, K. A. Dill, M. M. Inamdar, E. Seitaridou, and R. Phillips, Am. J. Phys. **74**, 123 (2006).
[12] E. Seitaridou, M. M. Inamdar, R. Phillips, K. Ghosh, and K. Dill, J. Phys. Chem. B **111**, 2288 (2007).
[13] D. Wu, K. Ghosh, M. Inamdar, H. J. Lee, S. Fraser, K. Dill, and R. Phillips, Phys. Rev. Lett. **103**, 050603 (2009).
[14] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).
[15] E. T. Jaynes, Phys. Rev. **108**, 171 (1957).
[16] J. Shore and R. Johnson, IEEE Trans. Info. Theory **26**, 26 (1980).
[17] J. Skilling, Nature **309**, 748 (1984).
[18] S. Pressé, K. Ghosh, J. Lee, and K. A. Dill, Rev. Modern Phys. **85**, 1115 (2013).
[19] P. Bamberg and S. Shlomo, *A Course in Mathematics for Students of Physics*, Vol. 1 (Cambridge University Press, Cambridge, 1991).
[20] L. R. Mead and N. Papanicolaou, J. Math. Phys. **25**, 2404 (1984).
[21] M. Grendár and M. Grendár, Entropy **3**, 58 (2001).
[22] G. Emch and C. Liu, *The Logic of Thermo-Statistical Physics* (Springer, Berlin, 2002).
[23] N. Agmon, Y. Alhassid, and R. D. Levine, J. Comput. Phys. **30**, 250 (1979).
[24] K. Bandyopadhyay, A. K. Bhattacharya, P. Biswas, and D. A. Drabold, Phys. Rev. E **71**, 057701 (2005).
[25] J. Banavar, A. Maritan, and I. Volkov, J. Phys.: Condensed Matter **22**, 063101 (2010).
[26] J. Skilling and R. K. Bryan, *Maximum Entropy Image Reconstruction: General Algorithm*, Vol. 211 (Wiley-Blackwell, Malden, MA, 1984).
[27] S. F. Gull and J. Skilling, Commun. Radar Signal Process. IEE Proc. F **131**, 646 (1984).
[28] J. E. Shore, arXiv:1304.3423 (2013).
[29] W. G. Hoover and B. L. Holian, Phys. Lett. A **211**, 253 (1996).
[30] T. Schneider and E. Stoll, Phys. Rev. B **17**, 1302 (1978).
[31] K. Aoki and D. Kusnezov, Phys. Lett. B **477**, 348 (2000).
[32] K. Aoki and D. Kusnezov, Phys. Lett. A **309**, 377 (2003).
[33] H. A. Posch and W. G. Hoover, Phys. Rev. E **55**, 6803 (1997).
[34] W. G. Hoover, C. Hoover, H. Posch, and J. Codelli, Commun. Nonlin. Sci. Numer. Simul. **12**, 214 (2007).
[35] W. G. Hoover and O. Kum, Phys. Rev. E **56**, 5517 (1997).
[36] J. C. Sprott, W. G. Hoover, and C. G. Hoover, Phys. Rev. E **89**, 042914 (2014).
[37] W. G. Hoover, H. A. Posch, and C. G. Hoover, Chaos: Interdisc. J. Nonlin. Sci. **2**, 245 (1992).
[38] A. Jones and B. Leimkuhler, J. Chem. Phys. **135**, 084125 (2011).
[39] M. Baiesi, C. Maes, and K. Netocn, J. Stat. Phys. **135**, 57 (2009).
[40] B. Derrida, J. Stat. Mech.: Theory Exp. (2007) P07023.
[41] H. Posch and W. Hoover, in *Molecular Liquids: New Perspectives in Physics and Chemistry*, Vol. 379, edited by J. Teixeira-Dias (Springer, Netherlands, 1992), pp. 527–547.
[42] J. Shore and R. Johnson, IEEE Trans. Info. Theory **27**, 472 (1981).
[43] A. Caticha, AIP Conf. Proc. **707**, 75 (2004).
[44] A. Caticha and A. Giffin, AIP Conf. Proc. **872**, 31 (2006).
[45] A. Giffin, *Maximum Entropy: The Universal Method for Inference* (ProQuest, Ann Arbor, MI, 2008).
[46] A. Caticha, AIP Conf. Proc. **1305**, 20 (2011).
[47] J. Banavar and A. Maritan, arXiv:cond-mat/0703622 (2007).
[48] M. S. Shell, J. Chem. Phys. **129**, 144108 (2008).
[49] B. Haegeman and R. S. Etienne, Am. Naturalist **175**, E74 (2010).
[50] L. Diambra, Phys. A: Stat. Mech. Appl. **390**, 2198 (2011).
[51] X. Guan, A. Giffin, R. Jha, and Y. Liu, Probabil. Eng. Mech. **29**, 157 (2012).
[52] M. Meléndez and P. Español, J. Stat. Phys. **155**, 93 (2014).
[53] D. J. Evans and G. Morriss, *Statistical Mechanics of Nonequilibrium Liquids*, 2nd ed. (Cambridge University Press, Cambridge, 2008).
[54] S. Maruyama, Adv. Numer. Heat Transfer **2**, 189 (2000).
[55] A. Guajardo-Cuéllar, D. B. Go, and M. Sen, J. Chem. Phys. **132**, 104111 (2010).