

# Lawrence Berkeley National Laboratory

## Recent Work

**Title**

Multi-point nonequilibrium umbrella sampling and associated fluctuation relations

**Permalink**

<https://escholarship.org/uc/item/7601k6s4>

**Journal**

Journal of Statistical Mechanics: Theory and Experiment, 2018(6)

**ISSN**

1742-5468

**Author**

Whitelam, S

**Publication Date**

2018-06-26

**DOI**

10.1088/1742-5468/aac915

Peer reviewed

# Multi-point nonequilibrium umbrella sampling and associated fluctuation relations

Stephen Whitelam\*

*Molecular Foundry, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

We describe a simple method of umbrella trajectory sampling for Markov chains. The method allows the estimation of large-deviation rate functions, for path-extensive dynamic observables, for an arbitrary number of models within a certain family. The general relationship between probability distributions of dynamic observables of members of this family is an extended fluctuation relation. When the dynamic observable is chosen to be entropy production, members of this family include the forward Markov chain and its time reverse, whose probability distributions are related by the expected simple fluctuation relation.

## I. INTRODUCTION

One concept widely used to quantify rare behavior in computer simulations is importance sampling, in which the typical behavior of a reference system is used to infer the rare behavior of a model of interest [1]. Umbrella sampling [2] is a method of importance sampling used in equilibrium. Here, a model is modified by a bias potential that constrains it to a region of interest, and knowledge of the Boltzmann weights of the modified and original models allows the user to calculate thermodynamic properties of the latter. Away from equilibrium there exist methods similar in spirit to umbrella sampling. These methods probe the statistical properties of dynamic trajectories, using a reference dynamics whose typical behavior is in some sense equivalent to the rare behavior of the model of interest [3–15].

One such method, described in Refs. [16, 17], applies to Markov chains of a fixed number of configuration changes. The method makes use of a set of reference models to reconstruct the large-deviation rate functions, for path-extensive dynamical observables, for an original model of interest. The rate functions of the reference models can be regarded as nonequilibrium umbrella potentials, concentrating sampling in particular regions of parameter space, an idea sketched in Fig. 1(a). The concept of trajectory importance sampling is well known [3–15]; the approach of Refs. [16, 17] is different to most approaches in that it uses as a reference dynamics a simple (rather than an optimal) modification of the original dynamics, and does not use population dynamics (a.k.a. cloning) methods [8].

In general, trajectory importance sampling of the probability distribution  $\rho(a) \sim e^{-KI(a)}$  of a path-extensive quantity  $Ka$  requires evaluation of relations of the type

$$e^{-KI(a)} \sim e^{sKa} \int dq P_s(q) e^{Kq}. \quad (1)$$

Here  $I(a)$ , the large-deviation rate function for  $a$  for the original model, is the quantity we want to calculate;  $s$  is a parameter associated with the reference model;  $q$

is a fluctuating quantity associated with trajectories of the reference model; and  $K$ , the trajectory length, is a large parameter.  $P_s(q)$  is the distribution of  $q$  associated with an ensemble of reference-model trajectories. The formally optimal approach to trajectory sampling determines (or approximates) the reference model for which  $P_s(q) = \delta(q)$ , so that reference-model trajectories of given  $a$  are a constant factor less probable than those of the original model [18–22]. In this case the effort of the method lies in the construction of a (potentially complex) reference model whose properties are ideal in the stated sense. The present method uses instead a reference-model set that is a simple modification of the existing model, and the effort of the method focuses on evaluation of  $P_s(q)$  in order to recover  $I(a)$ .

Direct evaluation of (1) is impractical when  $K$  is large. If  $P_s(q)$  is Gaussian (the best case scenario outside of  $q$  being constant) with variance  $\sigma^2$ , then the error in the measurement of  $e^{Kq}$  over trajectories of the reference model is

$$\frac{\langle (e^{Kq})^2 \rangle_s - \langle e^{Kq} \rangle_s^2}{\langle e^{Kq} \rangle_s^2} \sim e^{K^2 \sigma^2}, \quad (2)$$

where  $\langle \cdot \rangle_s \equiv \int dq (\cdot) P_s(q)$ . The variance  $\sigma^2$  is typically  $\propto 1/K$ , in which case (2) diverges exponentially with  $K$ , and a number of trajectories exponentially large in  $K$  is needed to evaluate (1). However, it is not necessary to evaluate (1) or measure  $e^{Kq}$  directly (indeed, for  $K$  of even modest size such exponentials cannot be evaluated on a computer). Taking logarithms of (1) gives

$$I(a) = -sa - \bar{q} - K^{-1} \ln \int dq P_s(q) e^{K(q-\bar{q})}, \quad (3)$$

where  $\bar{q}$  is the mean value of  $q$  over the ensemble of reference-model trajectories. The first two terms of (3) can be calculated straightforwardly, using a number of trajectories that does not scale with  $K$  (see Appendix C of [16]), and provide an upper bound on the rate function  $I(a)$ . Any choice of reference dynamics produces *some* upper bound on the rate function [23]; we show that certain simple and physically-motivated choices produce meaningful (i.e. tight) bounds. Moreover, with some additional numerical effort the third term in (3) can be evaluated or closely approximated, yielding the rate function itself. For the case of Gaussian  $P_s(q)$ , the third term

\* swhitelam@lbl.gov

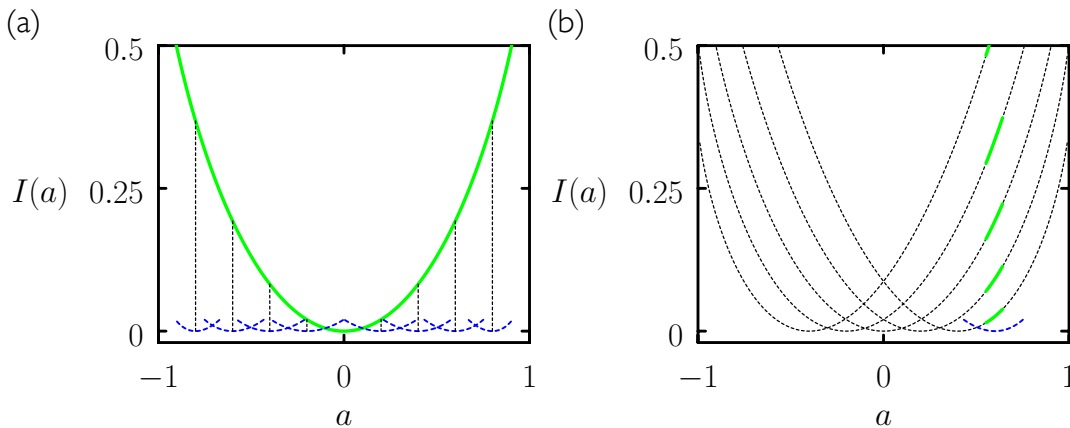


FIG. 1. Nonequilibrium umbrella sampling. (a) Estimation of the rate function  $I(a)$  of a model  $s$  (green), a member of the family (5), can be done using a set of reference models  $s'$ . Typical reference-model trajectories, i.e. trajectories whose values of  $a$  concentrate near the minima  $a_{s'}$  of the reference-model rate functions (blue dashed lines), each allow reconstruction of one point  $I(a_{s'})$  on the green curve. Unlike conventional umbrella sampling, each piece of  $I(a_{s'})$  is evaluated independently, and values of  $s'$  can be as close or as far apart as desired. (b) Multi-point sampling. In this paper we show that each set of reference-model simulations can be used to recover one point (middle of the green interval) on the rate functions of any set  $\{s\}$  of models of the family (5). Thus a collection of reference-model simulations allows reconstruction of the rate functions of the entire family of models.

is equal to  $K\sigma^2/2$ , which can be evaluated by direct simulation using a number of trajectories that scales less quickly than linearly with  $K$  (see Appendix C of [16]). Higher-order cumulants of the integral become progressively more expensive to evaluate – and the method will fail if  $P_s(q)$  has long tails that cannot be sampled efficiently by direct simulation – but in what follows we show that a set of simple reference-model choices can yield short-tailed  $P_s(q)$ , in which case the rare behavior of the original model can be evaluated by direct simulation of a set of simple reference models.

Thus the method described in Refs. [16, 17] is a particular implementation of trajectory importance sampling, and is non-optimal in the formal sense. However, it is simple and it works: it requires no special techniques beyond direct simulation, and can be used to bound or closely approximate the rate function in a variety of settings. In addition, the method directly computes the logarithmic probability distribution for path-extensive observables, rather than computing the Legendre transform of the distribution. This feature is useful when the probability distribution has linear or non-convex portions, in which case its Legendre transform is singular [4].

In this paper we describe two extensions of this method. First, we point out that each set of reference dynamical trajectories can be used to generate pieces of the probability distribution of *any* stochastic model within a particular family. We refer to this procedure as *multi-point* sampling, an idea sketched in Fig. 1(b). Second, we extend the method to Markov chains of fixed time.

The paper is organized as follows. In Section II we describe the multi-point sampling procedure. We show that the general relationship between probability distributions of dynamic observables of any two members of

this family is an extended fluctuation relation. The sampling method proceeds by evaluating many distribution-pairs at the concentration point of one distribution, so allowing evaluation of the other distribution away from its concentration point. In Section III we show that when the dynamical observable is entropy production, models in this family include forward and time-reversed trajectory ensembles, whose distributions satisfy the expected fluctuation relation [24–30]. In Section IV we show how to carry out the sampling method for trajectories of fixed time. We conclude in Section V.

## II. IMPORTANCE SAMPLING OF DYNAMIC TRAJECTORIES

The reference-model method of Refs. [16, 17] is designed to calculate probability distributions of dynamic observables for Markov chains of a fixed number of configuration changes. Consider a family of models, parameterized by a variable  $s$ , that move between microstates  $C$  and  $C'$  with probability

$$p_s(C \rightarrow C') = \frac{W_s(C \rightarrow C')}{R_s(C)}. \quad (4)$$

Here

$$W_s(C \rightarrow C') = e^{-s\alpha(C \rightarrow C')} W_0(C \rightarrow C') \quad (5)$$

is a rate, and  $R_s(C) \equiv \sum_{C'} W_s(C \rightarrow C')$  is a normalization factor that ensures probability conservation, i.e.  $\sum_{C'} p_s(C \rightarrow C') = 1$ . The quantity  $\alpha(C \rightarrow C')$  is the change of an arbitrary dynamic observable  $A$  upon moving from  $C$  to  $C'$ . The form (5) is motivated by the exponential tilting of the probability distribution commonly

done in the mathematical literature [4, 31], although this choice of rates does not in general produce the tilted distribution.

Our aim is to calculate the probability distribution

$$\rho_s(a, K) \equiv \sum_{\mathbf{x}} P_s[\mathbf{x}] \delta(A[\mathbf{x}] - Ka) \quad (6)$$

of the intensive variable  $a = A/K$ , for a particular model  $s$ , over trajectories  $\mathbf{x} = \{C_0, C_1, \dots, C_K\}$  of  $K$  steps of the dynamics. (Other approaches to trajectory sampling calculate the cumulant-generating functions of dynamic observables [5, 27, 32].) In (6), the quantity

$$P_s[\mathbf{x}] = \rho_s(C_0) \prod_{k=0}^{K-1} p_s(C_k \rightarrow C_{k+1}) \quad (7)$$

is the probability of generating a trajectory  $\mathbf{x}$  within the model  $s$ , and

$$A[\mathbf{x}] = \sum_{k=0}^{K-1} \alpha(C_k \rightarrow C_{k+1}) \quad (8)$$

is the extensive dynamic observable for the trajectory  $\mathbf{x}$ . Direct simulation of the model leads to good sampling of  $\rho_s(a, K)$  for  $a \approx a_s$ , the value of  $a$  typical of the model. For models with a well-defined stationary measure

$$\pi_s(C) = \sum_{C'} \pi_s(C') p_s(C' \rightarrow C), \quad (9)$$

this value is given by

$$\begin{aligned} a_s &= \sum_C \pi_s(C) \sum_{C'} p_s(C \rightarrow C') \alpha(C \rightarrow C') \\ &= - \sum_C \pi_s(C) \partial_s \ln R_s(C). \end{aligned} \quad (10)$$

(The method also works for models that do not have a well-defined stationary measure [16].)

Trajectories of the model will concentrate on  $a_s$  in the large- $K$  limit, leading to good sampling there but poor sampling elsewhere. A standard way to overcome this problem is to use importance sampling [3–5, 7, 8]. Consider a second model  $s'$ , called the reference model, whose trajectories concentrate in the large- $K$  limit on  $a_{s'}$ . Direct simulation of the reference model will lead to good sampling in the vicinity of  $a_{s'}$ , and this statistics can be used to reproduce the piece  $\rho_s(a_{s'}, K)$  of the original model's distribution. To see how, note that the relative likelihood with which a trajectory  $\mathbf{x}$  is generated by reference and original models,  $w_{ss'}[\mathbf{x}] = P_s[\mathbf{x}]/P_{s'}[\mathbf{x}]$ , is

$$w_{ss'}[\mathbf{x}] = e^{(s'-s)A[\mathbf{x}] + Kq_{ss'}[\mathbf{x}]}, \quad (11)$$

where

$$q_{ss'}[\mathbf{x}] \equiv K^{-1} \sum_{k=0}^{K-1} \ln \frac{R_{s'}(C_k)}{R_s(C_k)}. \quad (12)$$

The quantity we want, the probability distribution of  $a = A/K$  for trajectories of length  $K$  of the original model, is

$$\begin{aligned} \rho_s(a, K) &\equiv \sum_{\mathbf{x}} P_{s'}[\mathbf{x}] w_{ss'}[\mathbf{x}] \delta(A[\mathbf{x}] - Ka) \\ &= e^{(s'-s)aK} \sum_{\mathbf{x}} P_{s'}[\mathbf{x}] e^{Kq_{ss'}[\mathbf{x}]} \delta(A[\mathbf{x}] - Ka) \\ &\equiv \rho_{s'}(a, K) e^{(s'-s)aK} \langle e^{Kq_{ss'}[\mathbf{x}]} \rangle_{s'}, \end{aligned} \quad (13)$$

where

$$\langle \cdot \rangle_{s'}^{aK} \equiv \frac{\sum_{\mathbf{x}} P_{s'}[\mathbf{x}] (\cdot) \delta(A[\mathbf{x}] - Ka)}{\sum_{\mathbf{x}} P_{s'}[\mathbf{x}] \delta(A[\mathbf{x}] - Ka)} \quad (14)$$

is an average over reference-model trajectories that possess  $A[\mathbf{x}] = Ka$ . Eq. (13) can be written

$$\frac{\rho_s(a, K)}{\rho_{s'}(a, K)} = e^{(s'-s)aK} \int dq_{ss'} e^{Kq_{ss'}} P_{s'}(q_{ss'}|a), \quad (15)$$

where  $P_{s'}(q_{ss'}|a)$  is the probability distribution of the fluctuating piece of the path weight,  $q_{ss'}[\mathbf{x}]$ , for trajectories of the reference model  $s'$  that possess  $A[\mathbf{x}] = Ka$ . Normalization is such that  $\int dq_{ss'} P_{s'}(q_{ss'}|a) = 1$ . For models for which the stationary measure  $\pi_s(C)$  exists, values of  $q_{ss'}[\mathbf{x}]$  fluctuate, from trajectory to trajectory, about the typical value

$$q_{ss'} = \sum_C \pi_{s'}(C) \ln \frac{R_{s'}(C)}{R_s(C)}. \quad (16)$$

Eq. (15) involves no approximations: it is a re-writing of (6), and is valid for arbitrary path length  $K$ . The expression makes clear that the trajectory-ensemble distributions of  $a$  for two models,  $s$  and  $s'$ , are related by an extended fluctuation relation [4].

This relation can be used as a tool for sampling the distribution of a model  $s$  using a reference model  $s'$ . Consider the case of large  $K$ . For many models, for sufficiently large  $K$ , the distribution  $\rho_s(a, K)$  adopts the large-deviation form  $\rho_s(a, K) \sim e^{-KI_s(a)}$ , where  $I_s(a)$  is the large-deviation rate function. To use (15) as a method of reconstructing  $I_s(a)$ , note that the equation can be evaluated at the point  $a = a_{s'}$  where trajectories of the reference model concentrate, i.e. where  $I_{s'}(a_{s'}) = 0$ . Then (15) can be rewritten

$$\begin{aligned} I_s(a_{s'}) &= (s - s')a_{s'} - \bar{q}_{ss'} \\ &\quad - K^{-1} \ln \int dq_{ss'} e^{K(q_{ss'} - \bar{q}_{ss'})} P_{s'}(q_{ss'}|a_{s'}), \end{aligned} \quad (17)$$

whose evaluation gives one point on the curve  $I_s(a) = -K^{-1} \ln \rho(a, K)$ . (We omit the limit  $K \rightarrow \infty$  in the definition of the large-deviation rate function, in order to emphasize that we calculate quantities for large but finite values of path length. We have verified that the precise value of  $K$  makes no difference to the plots in the paper, provided that it is large enough.) Here  $\bar{q}_{ss'}$  is the mean of the quantity  $q_{ss'}[\mathbf{x}]$ , for trajectories of the

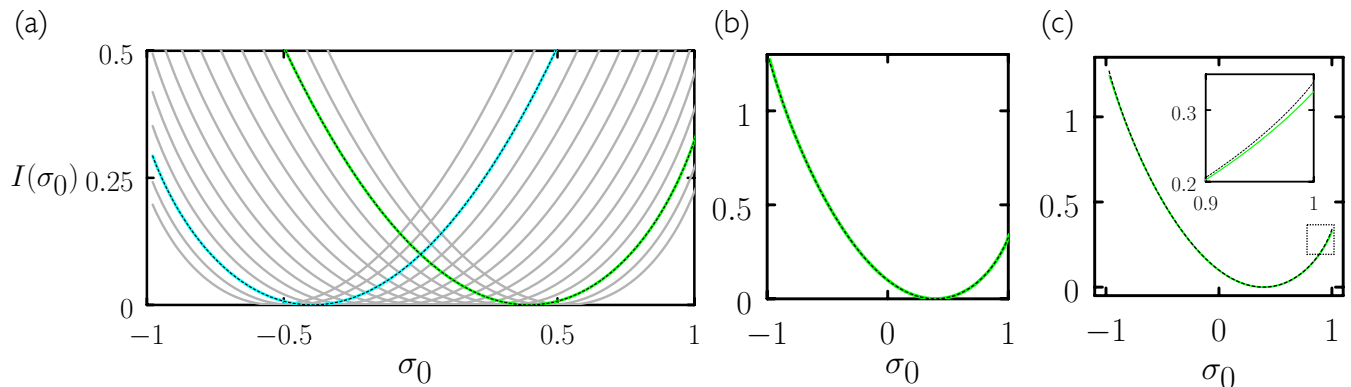


FIG. 2. (a) Large-deviation rate functions for entropy production,  $I(\sigma_0) = -K^{-1} \ln \rho(\sigma_0, K)$ , for 15 models  $s = \{-1.2, 1.1, \dots, 0.1, 0.2\}$  (left to right) in the family (5). The ‘original’ model ( $s = 0$ ) is the the 4-state model of Ref. [33]. Results were calculated using the multi-point nonequilibrium umbrella sampling scheme described in the text. The green and cyan curves denote the original ( $s = 0$ ) and reverse ( $s = 1$ ) models; the black dashed lines, which denote  $I_0(\sigma_0) + \sigma_0$  on the left and  $I_1(\sigma_0) - \sigma_0$  on the right, verify that these two models obey the fluctuation relation (26). Note that entropy production  $\sigma_0$  is defined with respect to the model  $s = 0$ . (b) Rate function for the model  $s = 0$  computed using the present method (green) and by diagonalization of the tilted transition matrix (black). (c) Comparison of the exact result (green) and the rate-function upper bound, Eq. (19) (black). Inset: zoom-in of the dotted area.

reference model  $s'$  that have  $A[\mathbf{x}] = Ka_{s'}$ . The variance of the integral in (17) diverges (in general) exponentially with  $K$ , but its logarithm can, in the cases we have encountered [16, 17], be evaluated without excessive numerical effort. Upon simulating the reference model, we can measure  $P_{s'}(q_{ss'}|a_{s'})$  (provided that this quantity does not have long tails that cannot be sampled efficiently by direct simulation). If this distribution happens to be Gaussian in  $q_{ss'}$ , then (17) can be evaluated analytically to give

$$I_s^{(1)}(a_{s'}) = (s - s')a_{s'} - \bar{q}_{ss'} - K\sigma_{ss'}^2/2. \quad (18)$$

For non-Gaussian weight fluctuations we can regard Eq. (18) as an approximation of the rate function. Here  $\sigma_{ss'}^2$  is the variance of the quantity  $q_{ss'}[\mathbf{x}]$ , for trajectories of the reference model  $s'$  that have  $A[\mathbf{x}] = Ka_{s'}$ . In this case the numerical effort required to evaluate the point  $I_s(a_{s'})$  is (only) the effort required to compute the variance (from trajectory to trajectory)  $\sigma_{ss'}^2 \propto 1/K$  of the intensive quantity  $q_{ss'}[\mathbf{x}]$ . If  $P_{s'}$  is non-Gaussian in  $q_{ss'}$  then the integral in (17) can be evaluated via calculation of higher-order cumulants, with the associated additional computational cost [16]. The fact that the integral in (17) is exponentially large in  $Kq[\mathbf{x}]$  does not automatically defeat the re-weighting procedure: if the distribution  $P_{s'}(q_{ss'}|a_{s'})$  can be sampled efficiently by the reference model dynamics (e.g. if it is approximately Gaussian) then (18) can be evaluated with reasonable accuracy.

The first two terms on the right-hand side of Eq. (17) provide, by Jensen’s inequality, an upper bound  $I_s^{(0)}(a_{s'})$  on the rate function piece  $I_s(a_{s'})$  (whether or not  $P_{s'}$  is Gaussian),

$$I_s^{(0)}(a_{s'}) = (s - s')a_{s'} - \bar{q}_{ss'}. \quad (19)$$

This bound, like the rate function itself, is not required to be quadratic about its minimum or even to be convex [16, 17]. It can be calculated using a single reference-model trajectory, or analytically if  $\pi_s(C)$  can be calculated analytically [17].

In Refs. [16, 17, 34] we used this method to calculate large-deviation rate functions for a single model  $s = 0$  using a set of reference models  $s'$ , a procedure illustrated in Fig. 1(a). The rate functions of the reference models serve as nonequilibrium umbrella potentials, concentrating sampling at certain points in parameter space. In this paper we point out that Equations (15) and (17), evaluated using a reference model  $s'$ , can be used to produce large-deviation rate functions for any number of ‘original’ models  $\{s\}$ , without doing additional simulations. The quantity  $q_{ss'}[\mathbf{x}]$ , given by Eq. (12), depends on the identity of the original model only through the label  $s$ ; the choice of microstates visited is determined solely by the reference model  $s'$ . Thus by keeping track, in a reference-model simulation, of the set of  $N$  values  $q_{\{s\}s'}[\mathbf{x}]$ , where  $\{s\} = \{s_1, \dots, s_N\}$ , we can recover  $N$  pieces  $I_{\{s\}}(a_{s'})$  of  $N$  distinct ‘original’ models. This idea is illustrated in Fig. 1(b). Then, by scanning  $s'$ , we can simultaneously reconstruct  $N$  different rate functions  $I_{\{s\}}(a)$ .

### III. SAMPLING ENTROPY PRODUCTION

To illustrate this procedure we choose to sample entropy production. Let the bias appearing in Eq. (5) be

$$\alpha(C \rightarrow C') = \ln \frac{\pi_0(C)p_0(C \rightarrow C')}{\pi_0(C')p_0(C' \rightarrow C)}, \quad (20)$$

the entropy produced by the model  $s = 0$  upon moving from  $C \rightarrow C'$ ; here  $\pi_0(C)$  is the stationary measure of the model  $s = 0$  [see Eq. (9)]. The extensive dynamical order parameter is then the total entropy produced by a trajectory  $\mathbf{x}$  of  $K$  steps,  $A[\mathbf{x}] = \sum_k \alpha(C_k \rightarrow C_{k+1}) \equiv \Sigma_0[\mathbf{x}]$ . The rate of entropy production is the intensive version of this parameter,  $\sigma_0[\mathbf{x}] = \Sigma_0[\mathbf{x}]/K$ . In this case the family of models (5) is

$$W_s(C \rightarrow C') = \left( \frac{\pi_0(C')p_0(C' \rightarrow C)}{\pi_0(C)p_0(C \rightarrow C')} \right)^s W_0(C \rightarrow C'). \quad (21)$$

The parameter  $s$  influences the rate of entropy production (here defined with respect to the model  $s = 0$ ), distributions of which obey the expression (15).

The distributions of certain members of this family are related in a simple way. In particular, the model  $s = 1$  is the reverse of the model  $s = 0$ . From (21) we have

$$W_1(C \rightarrow C') = \frac{\pi_0(C')}{\pi_0(C)} R_0(C) p_0(C' \rightarrow C). \quad (22)$$

Then

$$R_1(C) \equiv \sum_{C'} W_1(C \rightarrow C') = R_0(C), \quad (23)$$

using (9), and so

$$p_1(C \rightarrow C') = \frac{\pi_0(C')}{\pi_0(C)} p_0(C' \rightarrow C), \quad (24)$$

which is the time-reversed Markov chain [28]. The probability distributions of entropy produced by the models  $s = 0$  and  $s = 1$  then obey the expected fluctuation relation. Setting  $s = 0$  and  $s' = 1$ , the terms of the fluctuating piece  $q_{ss'}[\mathbf{x}']$  of the path weight, Eq. (12), are

$$\ln \frac{R_1(C)}{R_0(C)} = 0, \quad (25)$$

using (23), giving  $q_{01}[\mathbf{x}] = 0$  and  $P_1(q_{01}|a) = \delta(q_{01})$ . Hence for  $s = 0$  and  $s' = 1$ , Eq. (15) reduces to

$$\frac{\rho_0(\sigma_0, K)}{\rho_1(\sigma_0, K)} = e^{\sigma_0 K}. \quad (26)$$

Given that the entropy production  $\sigma$  changes sign if we switch from the forward to the reverse model, i.e.  $\sigma_0 = -\sigma_1$ , we can write (26) in the form

$$\frac{\rho(\Sigma)}{\rho_{\text{rev}}(-\Sigma)} = e^\Sigma, \quad (27)$$

which is the fluctuation theorem for nonequilibrium steady states [24–27, 29]. Here ‘rev’ denotes the reverse model, and  $\Sigma = \sigma K$ . Thus, the general expression (15) relates the distribution of an arbitrary dynamic observable  $A$  for any two members of the the family of models  $s$ . If  $A$  is chosen to be entropy production  $\Sigma$ , then, for

$0 < s < 1$ , these models interpolate between the forward and reverse versions of the model whose rates are  $W_0(C \rightarrow C')$ .

We can use these fluctuation relations, in the form (17), as a method for reconstructing the large-deviation rate functions for entropy production for all models  $\{s\}$  in the family (5). To illustrate this procedure we consider the 4-state model of Ref. [33]. This is a fully connected model with states  $C \in \{1, 2, 3, 4\}$ . We take the rates  $W_0(C \rightarrow C')$  to be the quantities  $r(C, C')$  in the caption of Fig. 2 of that paper:

$$\begin{aligned} r(1, 2) &= 3, & r(1, 3) &= 10, & r(1, 4) &= 9, \\ r(2, 1) &= 10, & r(2, 3) &= 1, & r(2, 4) &= 2, \\ r(3, 1) &= 6, & r(3, 2) &= 4, & r(3, 4) &= 1, \\ r(4, 1) &= 7, & r(4, 2) &= 9, & r(4, 3) &= 5, \end{aligned}$$

with  $r(C, C) = 0$ . We carried out a series of simulations of 100 different reference models, whose values of  $s'$  run from  $-2$  to  $2.5$  in evenly-spaced intervals. For each reference model  $s'$  we generated  $8 \times 10^3$  independent trajectories of length  $K = 10^6$ . From this ensemble we calculated the typical value  $a_{s'}$  of the dynamical observable, and identified those trajectories whose values of  $a$  were within an interval  $\pm 10^{-3}$  of  $a_{s'}$ . From this restricted ensemble we collected values of the fluctuating path weights  $q_{\{s\}, s'}[\mathbf{x}]$ , for 15 different values  $\{s\}$ , evenly spaced between  $-0.2$  and  $1.2$ .

These quantities allow the approximation, via Eq. (18), of the large-deviation rate function  $I_s(\sigma_0) = -K^{-1} \ln \rho_s(\sigma_0, K)$  for 15 members of the family (5). Results are shown in Fig. 2(a). Recall that the variable  $\sigma_0$  is the entropy production rate defined with respect to the model  $s = 0$ . The green and cyan curves correspond to the models  $s = 0$  and  $s = 1$  respectively, which are the forward and reverse versions of the model with rates  $W_0(C \rightarrow C')$ . The black dotted lines verify that these models obey the fluctuation relation (26). The left-hand black dotted line is  $I_0(\sigma_0) + \sigma_0$ , which is equal to  $I_1(\sigma_0)$  (cyan line). The right-hand black dotted line is  $I_1(\sigma_0) - \sigma_0$ , which is equal to  $I_0(\sigma_0)$  (green). This comparison provides a nontrivial check on the numerics, because it involves the superposition of the centers and tails of distinct distributions. As a further check on our numerics we show in Fig. 2(b) a comparison between the rate function calculated by the procedure described here, and that obtained by Legendre transform of the logarithm of the largest eigenvalue of the tilted transition-probability matrix [4, 35]. More generally, the distributions shown satisfy  $I_s(\sigma_0) = I_{1-s}(-\sigma_0)$ .

Recent work has identified rate-function bounds for generalized currents for discrete-time and continuous-time Markov chains [33, 35–38]. The rate-function upper bound obtained from the present method, Eq. (19), can be evaluated analytically for simple models or, in general, using a single reference-model trajectory [17]. The bound is not necessarily quadratic or convex. Combined with the multi-point sampling method, the bound

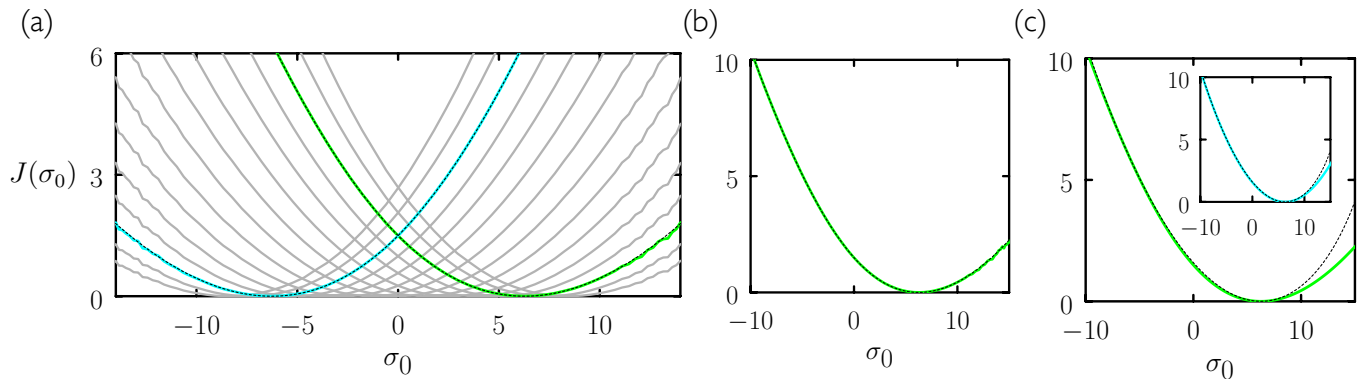


FIG. 3. As Fig. 2, but for trajectories of constant time rather than constant event number; here  $J(\sigma_0) = -T^{-1} \ln \rho(\sigma_0, T)$ . (a) The black dashed lines,  $J_0(\sigma_0) + \sigma_0$  on the left and  $J_1(\sigma_0) - \sigma_0$  on the right, verify that the fluctuation relation (27) holds. (b) Numerically computed rate function for the model  $s = 0$  (green), versus the exact solution (black) calculated via matrix diagonalization. (c) Numerically computed rate-function upper bound, Eq. (41) (black), versus the exact solution (green). Inset: bound from the present method (black), compared with the thermodynamic uncertainty relation from Ref. [33] (cyan; this is the upper line from the main panel of Fig. 2 of that reference).

provides a cheap way of estimating the large-deviation rate function for a range of models and conditions. For the present model and order parameter, the bound is relatively tight: see Fig. 2(c).

Here we have demonstrated that multiple rate functions for path-extensive observables of Markov chains of a fixed number of configuration changes can be calculated in a simple way using a set of stochastic simulations. The rate functions of more complex models, e.g. lattice models, can be obtained in a similar way [16, 17, 34].

#### IV. SAMPLING TRAJECTORIES OF FIXED TIME

##### A. General considerations

The procedure described thus far applies to Markov chains in which the total number of configuration changes  $K$  is fixed. Here we present one way to adapt this method to trajectories of fixed time  $T$ . There are more sources of fluctuation in this ensemble, because upon making a move we select both a new microstate and a time increment, the latter being an exponentially-distributed random variable [39]. In general we must impose a bias on both choices in order to sample dynamic observables.

Consider a set of models with rates

$$\hat{W}_s^\lambda(C \rightarrow C') = \Gamma_s^\lambda(C) W_s(C \rightarrow C'). \quad (28)$$

Here  $W_s(C \rightarrow C')$  is given by (5), and biases the choice of microstate  $C'$ . The factor  $\Gamma_s^\lambda(C)$  is chosen to bias the jump time from microstate  $C$ . A convenient choice is to take

$$\Gamma_s^\lambda(C) = \frac{R_0(C) + \lambda}{R_s(C)}, \quad (29)$$

which makes the escape rate of the new model to be a constant shift different to that of the original model, i.e.  $\hat{R}_s^\lambda(C) \equiv \sum_{C'} \hat{W}_s^\lambda(C \rightarrow C') = \Gamma_s^\lambda(C) R_s(C) = R_0(C) + \lambda$ . Here  $\lambda > -\min_C R_0(C)$  is a constant that serves to make the escape times from microstate  $C$  unusually large or small (by the reckoning of the original model). It has no effect upon the choice of microstate, i.e.

$$\hat{p}_s(C \rightarrow C') = \frac{\hat{W}_s^\lambda(C \rightarrow C')}{\hat{R}_s^\lambda(C)} = p_s(C \rightarrow C'), \quad (30)$$

as Eq. (4).

We wish to calculate [40]

$$\rho_s^\lambda(a, T) \equiv \sum_{\mathbf{x}} \hat{P}_s^\lambda[\mathbf{x}] \delta(A[\mathbf{x}] - Ta) \delta(T[\mathbf{x}] - T), \quad (31)$$

where now

$$\sum_{\mathbf{x}} = \sum_{K>0} \sum_{C_0 \dots C_K} \prod_{k=0}^{K-1} \int_0^\infty dt_k \quad (32)$$

denotes a sum over trajectories  $\mathbf{x}$  that can have variable numbers of configuration changes and jump times [8, 41]. [Using this notation, Eq. (6) would be written

$$\rho_s(a, K) \equiv \sum_{\mathbf{x}} P_s[\mathbf{x}] \delta(A[\mathbf{x}] - Ka) \delta(K[\mathbf{x}] - K).] \quad (33)$$

The quantity  $T[\mathbf{x}] = \sum_{k=0}^{K[\mathbf{x}]-1} t_k$  is the total time of trajectory  $\mathbf{x}$ ;  $t_k$  is the time increment of configuration change  $k+1$ , and is  $-\ln \eta / (R_0(C_k) + \lambda)$ , where  $\eta$  is a random number uniformly drawn from the interval  $(0, 1]$  [39].

The (bulk) path weight  $\hat{P}_s^\lambda[\mathbf{x}]$  for the models  $\hat{W}_s^\lambda$  is

$$\hat{P}_s^\lambda[\mathbf{x}] = \prod_{k=0}^{K[\mathbf{x}]-1} \hat{W}_s^\lambda(C_k \rightarrow C_{k+1}) e^{-[R_0(C_k) + \lambda] t_k}. \quad (34)$$

The relative weight  $\hat{w}_{ss'}^{\lambda\lambda'}[\mathbf{x}] = \hat{P}_s^\lambda[\mathbf{x}]/\hat{P}_{s'}^{\lambda'}[\mathbf{x}]$  is

$$\hat{w}_{ss'}^{\lambda\lambda'}[\mathbf{x}] = e^{(s'-s)A[\mathbf{x}] + (\lambda' - \lambda)T[\mathbf{x}] + T\phi_{ss'}^{\lambda\lambda'}[\mathbf{x}]}, \quad (35)$$

where  $A[\mathbf{x}] = \sum_{k=0}^{K[\mathbf{x}]-1} \alpha(C_k \rightarrow C_{k+1})$  is the extensive dynamic order parameter of trajectory  $\mathbf{x}$ , and

$$\phi_{ss'}^{\lambda\lambda'}[\mathbf{x}] \equiv T^{-1} \sum_{k=0}^{K[\mathbf{x}]-1} \ln \left( \frac{R_{s'}(C_k) R_0(C) + \lambda}{R_s(C_k) R_0(C) + \lambda'} \right) \quad (36)$$

is the analog of (12), with  $T$  replacing  $K$  as the measure of path length. Note that for fixed  $A[\mathbf{x}] = A$  and  $T[\mathbf{x}] = T$  the path weight (35) depends, as before, only on configurations visited. It does not depend upon the time of jumps, on account of all models  $\hat{W}_s^\lambda$  possessing identically shifted escape rates from a given configuration. Biasing jump times allows us to overcome the problem of sampling exponential tails of jump-time distributions using direct simulation (see Appendix B of Ref. [16]).

The quantity we want, the probability distribution of  $a = A/K$  for trajectories of the original model of length  $T$ , is

$$\begin{aligned} \rho_s^\lambda(a, T) &\equiv \sum_{\mathbf{x}} \hat{P}_{s'}^{\lambda'}[\mathbf{x}] \hat{w}_{ss'}^{\lambda\lambda'}[\mathbf{x}] \delta(A[\mathbf{x}] - Ta) \delta(T[\mathbf{x}] - T) \\ &= \rho_{s'}^{\lambda'}(a, T) e^{[(s'-s)a + (\lambda' - \lambda)]T} \langle e^{T\phi_{ss'}^{\lambda\lambda'}[\mathbf{x}]} \rangle_{s'\lambda'}^{aT}, \end{aligned} \quad (37)$$

where

$$\langle \cdot \rangle_{s'\lambda'}^{aT} \equiv \frac{\sum_{\mathbf{x}} \hat{P}_{s'}^{\lambda'}[\mathbf{x}] (\cdot) \delta(A[\mathbf{x}] - Ta) \delta(T[\mathbf{x}] - T)}{\sum_{\mathbf{x}} \hat{P}_{s'}^{\lambda'}[\mathbf{x}] \delta(A[\mathbf{x}] - Ta) \delta(T[\mathbf{x}] - T)} \quad (38)$$

is an average over reference-model trajectories of total time  $T$  that possess  $A[\mathbf{x}] = Ta$ . Eq. (37) can be written

$$\begin{aligned} \frac{\rho_s^\lambda(a, T)}{\rho_{s'}^{\lambda'}(a, T)} &= e^{[(s'-s)a + (\lambda' - \lambda)]T} \\ &\times \int d\phi_{ss'}^{\lambda\lambda'} e^{T\phi_{ss'}^{\lambda\lambda'}} P_{s'}(\phi_{ss'}^{\lambda\lambda'} | a), \end{aligned} \quad (39)$$

where  $P_{s'}^{\lambda'}(\phi_{ss'}^{\lambda\lambda'} | a)$  is the probability distribution of the fluctuating piece of the path weight,  $\phi_{ss'}^{\lambda\lambda'}[\mathbf{x}]$ , for trajectories of the reference model  $(s', \lambda')$  that possess  $A[\mathbf{x}] = Ta$ . Normalization is such that  $\int d\phi_{ss'}^{\lambda\lambda'} P_{s'}(\phi_{ss'}^{\lambda\lambda'} | a) = 1$ .

For models and observables for which, for long times, the large-deviation rate function  $J_{s\lambda}(a) = -T^{-1} \ln \rho_s^\lambda(a, T)$  exists, we can evaluate (39) at the concentration point  $a = a_{s'\lambda'}$  of the reference model, where  $J_{s'}^{\lambda'}(a_{s'\lambda'}) = 0$ , to give

$$\begin{aligned} J_{s\lambda}(a_{s'\lambda'}) &= (s - s')a_{s'\lambda'} + (\lambda - \lambda') - \bar{\phi}_{ss'}^{\lambda\lambda'} \\ &- T^{-1} \ln \int d\phi_{ss'}^{\lambda\lambda'} e^{T(\phi_{ss'}^{\lambda\lambda'} - \bar{\phi}_{ss'}^{\lambda\lambda'})} P_{s'}^{\lambda'}(\phi_{ss'}^{\lambda\lambda'} | a_{s'\lambda'}). \end{aligned} \quad (40)$$

Here  $\bar{\phi}_{ss'}^{\lambda\lambda'}$  is the mean of the quantity  $\phi_{ss'}^{\lambda\lambda'}[\mathbf{x}]$ , for trajectories of the reference model  $(s', \lambda')$  of length  $T$  that

have  $A[\mathbf{x}] = Ta_{s'\lambda'}$ . By Jensen's inequality, the first line on the right-hand side of (40) provides an upper bound

$$J_{s\lambda}^{(0)}(a_{s'\lambda'}) = (s - s')a_{s'\lambda'} + (\lambda - \lambda') - \bar{\phi}_{ss'}^{\lambda\lambda'} \quad (41)$$

on the rate function. A refinement to this estimate can be computed by assuming  $P_{s'}^{\lambda'}(\phi_{ss'}^{\lambda\lambda'} | a_{s'})$  to be Gaussian, in which case we get

$$J_{s\lambda}^{(1)}(a_{s'\lambda'}) = J_{s\lambda}^{(0)}(a_{s'\lambda'}) - T\sigma_{ss'}^{2\lambda\lambda'}/2. \quad (42)$$

Here  $\sigma_{ss'}^{2\lambda\lambda'}$  is the variance of the quantity  $\phi_{ss'}^{\lambda\lambda'}[\mathbf{x}]$ , for trajectories of the reference model  $(s', \lambda')$  of length  $T$  that have  $A[\mathbf{x}] = Ta_{s'\lambda'}$ . Simulations of the reference model  $(s', \lambda')$  for fixed time  $T$  therefore provide an estimate of one point  $a = a_{s'\lambda'}$  on the rate function  $J_{s\lambda}(a)$ .

The procedure described in this section is similar to that described in Section II, but with path length set by time  $T$  rather than number of configuration changes  $K$ . In both versions of the method the set of reference models is guided toward unlikely states of the original model with biased probabilities  $p_s(C \rightarrow C)$ . The additional consideration required for constant time is to draw random numbers  $\eta$  in order to compute jump times  $-\ln \eta / (R_0(C_k) + \lambda)$ , using  $\lambda$  as a way of sampling unlikely time changes. The choice of a constant shift  $\lambda$  removes the need to account for fluctuating jump times in the reweighting factor. The fluctuations of the remaining piece,  $\phi_{ss'}^{\lambda\lambda'}[\mathbf{x}]$ , Eq. (36), derive from fluctuations of the empirical measure of the reference model and from the fact that the number of terms in the sum,  $K[\mathbf{x}]$ , varies from trajectory to trajectory. What is required is to measure the statistics of this quantity for trajectories of fixed length  $T$  and fixed  $A[\mathbf{x}] = Ta$  (rather than for trajectories of fixed  $K$  and fixed  $A[\mathbf{x}] = Ka$ ).

## B. Entropy production

To test this method we calculated the large-deviation rate function for entropy production for the 4-state model of Ref. [33], for trajectories of constant time. In principle we should sample both rare states and rare jump times, but we found that the rate function for entropy production can be well approximated, over a large interval, by sampling only states. In this case we set  $\lambda' = 0$  and suppress  $\lambda$ -labels on quantities (the original model has  $s = \lambda = 0$ ). With the bias in (28) given by (20), the intensive dynamical order parameter  $a$  for a trajectory  $\mathbf{x}$  is  $\sigma_0[\mathbf{x}] = T^{-1} \sum_{k=0}^{K[\mathbf{x}]-1} \alpha(C_k \rightarrow C_{k+1})$ , the rate of entropy production. As described by the equations above, the procedure used to sample fluctuations of this variable for paths of fixed  $T$  is essentially identical to that used in Section III, except that now we run reference-model trajectories for a fixed time  $T = 10^6$  rather than for a fixed number of events. We generated  $5 \times 10^3$  reference-model trajectories for each value of  $s'$  (100 different values between  $-2$  and  $2.5$ ). The values of  $a$  of these trajectories fluctuate around a typical value  $a_{s'}$ . We calculated



the mean  $\bar{\phi}_{ss'}$  and variance  $\sigma_{ss'}$  using a series of narrow windows at and either side of the typical value  $a_{s'}$ , and constructed a weighted average of the values of  $\sigma_{ss'}$  calculated in these windows. This weighted average was used to compute the approximation (42). In this way we use most of the reference-model trajectories generated.

We show the results of this procedure in Fig. 3. Panel (a) shows that trajectories of constant time of the forward ( $s = 0$ ) and time-reversed ( $s = 1$ ) models obey the fluctuation relation (27), as expected [the passage from (39) to (27), for  $\lambda = \lambda' = 0$ , is the same as that from (15) to (27)], where now  $\Sigma = \sigma_0 T$ . Noise is visible in the tails of the numerical results – for constant  $T$  there are more sources of fluctuation than for constant  $K$  – but the rate functions can be reconstructed with reasonable accuracy within the interval shown. In panel (b) we compare the numerically computed approximation of the rate function, Eq. (42) (green), with the exact solution (black). The latter was calculated by Legendre transform of the largest eigenvalue of the tilted rate matrix [5, 27, 35]. In panel (c) we compare the exact solution (green) with the rate-function upper bound, Eq. (41) (black), which is less tight in general than its constant- $K$  analog [see Fig. 2(c)] but still close to the exact answer over a considerable interval. In the inset of the figure we compare the bound to the thermodynamic uncertainty relation [33]; the two bounds are not exactly the same, but are similar over a large interval. The thermodynamic uncertainty relation provides a ‘master’ upper bound for all currents. By contrast, the present approach provides a bound specific to the chosen observable, including observables that are not currents (see e.g. Fig. 4), as well as a scheme for improving upon the bound.

In the far wings of the rate functions (not shown) we found that calculation of the variance  $\sigma_{ss'}^2$  was unreliable; there, more trajectories are needed. In addition, we would expect, for extremely rare values of entropy production, that trajectories resulting from unusual jump times must be sampled (using the  $\lambda$ -bias) in order to accurately compute the rate function. One simple approach that could be used here is to carry out two-dimensional sampling in  $s'$  (biasing states) and  $\lambda$  (biasing jump times), similar to that done in Sec. V of [16]. Such sampling would produce a set of rate-function bounds at least as tight as the one shown in Fig. 3(c), and those bounds could be improved by evaluation of the fluctuation terms. However, given the close correspondence between numerics and the exact answer, such sampling does not appear to be necessary for the range of values of  $\sigma_0$  shown in the figure.

### C. Events

In studies of glass-like models, an important observable is the number of configuration changes (events)  $K$

occurring in time  $T$  [5]. In order to sample

$$\rho_0^0(K, T) \equiv \sum_{\mathbf{x}} \hat{P}_0^0[\mathbf{x}] \delta(K[\mathbf{x}] - K) \delta(T[\mathbf{x}] - T), \quad (43)$$

we used the umbrella procedure of Section IV A, choosing only to bias jump times (using  $\lambda$ ) and not the choice of state. In this case the reference model  $\hat{W}_0^\lambda(C \rightarrow C')$  selects states in the same way as the original model, with probabilities  $p_0(C \rightarrow C')$ , and selects jump times from exponential distributions with mean  $R_0(C) + \lambda$ . The path weight (36) simplifies to

$$\phi_{00}^{0\lambda'}[\mathbf{x}] \equiv T^{-1} \sum_{k=0}^{K[\mathbf{x}]-1} \ln \frac{R_0(C)}{R_0(C) + \lambda'}. \quad (44)$$

We carried out the usual procedure in order to estimate the large-deviation rate function for events  $k = K/T$  in a fixed time for the 4-state model of Ref. [33]. We used 70 reference models whose values of  $\lambda'$  were evenly spaced between  $\pm 10$ . For each value of  $\lambda'$  we generated  $2 \times 10^4$  trajectories of length  $T = 10^5$ , and calculated for each set the typical value  $a_{0\lambda'} = k_{0\lambda'}$ , and the values of  $\bar{\phi}_{00}^{0\lambda'}$  and  $\sigma_{00}^{0\lambda'}$  (as in the previous subsection we used a series of windows around the typical value  $a_{0\lambda'}$ ). We used these quantities to calculate the Gaussian weight-fluctuation approximation (42) to  $J(k) = -T^{-1} \ln \rho(k, T)$  [a single trajectory suffices to calculate the bound (41)]. Note that in those equations we have  $s = s' = 0$ .

Numerical results are shown in Fig. 4, compared with the exact result. The latter was obtained by Legendre

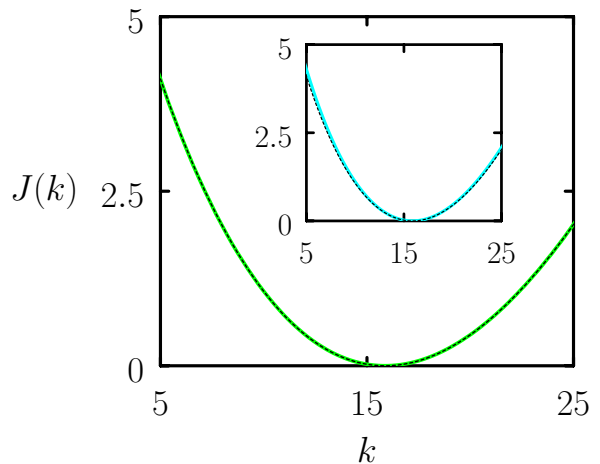


FIG. 4. Large-deviation rate function  $J(k) = -T^{-1} \ln \rho(k, T)$  for the number of configuration changes  $k = K/T$  per unit time in the 4-state model of Ref. [33]. The black line is the exact result, from matrix diagonalization. The green line results from numerical evaluation of Eq. (42), using the sampling procedure described in this paper. Inset: the cyan line is the upper bound (41), estimated using a single reference-model trajectory. This bound is close to the exact answer (black).

transform of the largest eigenvalue of the tilted rate matrix [5, 27, 35]. Both the bound and its refinement are close to the exact answer.

## V. CONCLUSIONS

We have described a simple method of umbrella sampling for Markov chains. The method uses a simple modification of a model  $W_0(C \rightarrow C')$ , namely (5) or (28), to estimate the large-deviation rate functions of a path-extensive observable for that model. In addition, with the same set of simulations one can reconstruct the large-deviation rate functions for any member of the family (5) or (28). The general relationship between probability distributions of members of this family is an extended fluctuation relation [Eqs. (15) or (39)]. When the observable is chosen to be the rate of entropy production, members of this family include the forward and reverse version of a certain model, whose probability distributions are related by a simple fluctuation relation, Eq. (26). The extension to multiple dynamic observables is possible, following Section V of [16] (see also section 2.3 of [27]).

The method described here applies to Markov chains of a fixed number of configuration changes or fixed time, with computation of the rate function in the latter case being (in general) numerically more demanding. However, in both cases it is straightforward to compute a

process-specific upper bound on the rate function, (19) or (41). The bound derives from the typical weight of a simply-modified version of the original process, and can be computed (for a particular value of the order parameter) using a single dynamic trajectory. The bound is not necessarily quadratic about its origin or convex. Combined with the multi-point sampling procedure, computation of the bound provides a simple and numerically cheap way to identify special points, such as dynamic phase transitions [16, 17], in a model's parameter space. The method also allows one to improve upon the bound. In the cases studied here this can be done with reasonable numerical effort, and results in estimates for large-deviation rate functions for path-extensive observables that agree with the answers obtained by matrix diagonalization.

## ACKNOWLEDGMENTS

I thank John Edison, Katherine Klymko, and Zdeněk Preisler for discussions, and thank Todd Gingrich for discussions and for sharing data from Ref. [33]. This work was performed at the Molecular Foundry, Lawrence Berkeley National Laboratory, supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

- 
- [1] D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications*, Vol. 1 (Academic Press, 2001).
  - [2] G. Torrie and J. Valleau, *Journal of Computational Physics* **23**, 187 (1977).
  - [3] J. A. Bucklew, *Large deviation techniques in decision, simulation, and estimation* (Wiley New York, 1990).
  - [4] H. Touchette, *Physics Reports* **478**, 1 (2009).
  - [5] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, *Journal of Physics A: Mathematical and Theoretical* **42**, 075007 (2009).
  - [6] C. Giardinà, J. Kurchan, and L. Peliti, *Physical Review Letters* **96**, 120603 (2006).
  - [7] C. Maes and K. Netočný, *EPL (Europhysics Letters)* **82**, 30003 (2008).
  - [8] C. Giardinà, J. Kurchan, V. Lecomte, and J. Tailleur, *Journal of Statistical Physics* **145**, 787 (2011).
  - [9] V. Lecomte and J. Tailleur, *Journal of Statistical Mechanics: Theory and Experiment* **2007**, P03004 (2007).
  - [10] T. Nemoto and S.-i. Sasa, *Physical Review Letters* **112**, 090602 (2014).
  - [11] C. M. Rohwer, F. Angeletti, and H. Touchette, *Physical Review E* **92**, 052104 (2015).
  - [12] T. Nemoto, F. Bouchet, R. L. Jack, and V. Lecomte, *Physical Review E* **93**, 062123 (2016).
  - [13] U. Ray, G. K. Chan, and D. T. Limmer, arXiv preprint arXiv:1708.00459 (2017).
  - [14] T. Oakes, S. Powell, C. Castelnovo, A. Lamacraft, and J. P. Garrahan, arXiv preprint arXiv:1802.09576 (2018).
  - [15] U. Ray, G. K. Chan, and D. T. Limmer, arXiv preprint arXiv:1708.09482 (2017).
  - [16] K. Klymko, P. L. Geissler, J. P. Garrahan, and S. Whitelam, *Phys. Rev. E* **97**, 032123 (2018).
  - [17] S. Whitelam, *Phys. Rev. E* **97**, 032122 (2018).
  - [18] R. Chetrite and H. Touchette, *Physical Review Letters* **111**, 120601 (2013).
  - [19] R. L. Jack and P. Sollich, *The European Physical Journal Special Topics* **224**, 2351 (2015).
  - [20] R. Chetrite and H. Touchette, *Journal of Statistical Mechanics: Theory and Experiment* **2015**, P12001 (2015).
  - [21] J. P. Garrahan, *Journal of Statistical Mechanics: Theory and Experiment* **2016**, 073208 (2016).
  - [22] G. Ferré and H. Touchette, arXiv preprint arXiv:1803.11117 (2018).
  - [23] S. S. Varadhan, in *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010) (In 4 Volumes) Vol. I: Plenary Lectures and Ceremonies Vols. II–IV: Invited Lectures* (World Scientific, 2010) pp. 622–639.
  - [24] G. Gallavotti and E. G. D. Cohen, *Physical Review Letters* **74**, 2694 (1995).
  - [25] J. Kurchan, *Journal of Physics A: Mathematical and General* **31**, 3719 (1998).
  - [26] C. Maes, *Journal of Statistical Physics* **95**, 367 (1999).
  - [27] J. L. Lebowitz and H. Spohn, *Journal of Statistical Physics* **95**, 333 (1999).
  - [28] G. E. Crooks, *Physical Review E* **61**, 2361 (2000).
  - [29] U. Seifert, *Physical Review Letters* **95**, 040602 (2005).

- [30] T. Speck, A. Engel, and U. Seifert, *Journal of Statistical Mechanics: Theory and Experiment* **2012**, P12001 (2012).
- [31] A. Dembo and O. Zeitouni, *Large deviations techniques and applications, volume 38 of Stochastic Modelling and Applied Probability* (Springer-Verlag, Berlin, 2010).
- [32] P. Pietzonka, A. C. Barato, and U. Seifert, *Physical Review E* **93**, 052145 (2016).
- [33] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, *Physical Review Letters* **116**, 120601 (2016).
- [34] S. Whitelam, K. Klymko, and D. Mandal, *The Journal of Chemical Physics* **148**, 154902 (2018).
- [35] D. Chiuchì and S. Pigolotti, *Phys. Rev. E* **97**, 032109 (2018).
- [36] K. Proesmans and C. Van den Broeck, *EPL (Europhysics Letters)* **119**, 20001 (2017).
- [37] P. Pietzonka, F. Ritort, and U. Seifert, *Physical Review E* **96**, 012101 (2017).
- [38] J. M. Horowitz and T. R. Gingrich, *Physical Review E* **96**, 020103 (2017).
- [39] D. Gillespie, *Journal of Computational Physics* **22**, 403 (2005).
- [40] For simplicity we omit the hat on  $\rho_s^\lambda(a, T)$ .
- [41] A. A. Budini, R. M. Turner, and J. P. Garrahan, *Journal of Statistical Mechanics: Theory and Experiment* **2014**, P03012 (2014).