

Exact fluctuations of nonequilibrium steady states from approximate auxiliary dynamics

Ushnish Ray and Garnet Kin-Lic Chan
*Division of Chemistry and Chemical Engineering,
California Institute of Technology, Pasadena, CA 91125*

David T. Limmer
*Department of Chemistry, University of California, Berkeley, CA 94609
Kavli Energy NanoScience Institute, Berkeley, CA 94609 and
Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94609*
(Dated: September 11, 2017)

We describe a framework to significantly reduce the computational effort to evaluate large deviation functions of time integrated observables within nonequilibrium steady states. We do this by incorporating an auxiliary dynamics into trajectory based Monte Carlo calculations, through a transformation of the system's propagator using an approximate guiding function. This procedure importance samples the trajectories that most contribute to the large deviation function, mitigating the exponentially complexity of such calculations. We illustrate the method by studying driven diffusions and interacting lattice models in one and two dimensions. Our work offers an avenue to calculate large deviation functions for high dimensional systems driven far from equilibrium.

Much like their equilibrium counterparts, fluctuations about nonequilibrium steady states encode physical information about a system. This is illustrated by the discovery of the fluctuation theorems [1–4], the thermodynamic uncertainty relations [5, 6], and the extensions of the fluctuation-dissipation theorem to systems far-from-equilibrium [7–9]. Large deviation rate functions characterize fluctuations out of equilibrium [10], and their evaluation has underpinned much of the recent progress in understanding driven systems. While this progress affords the potential for new insight into the many real and synthetic systems that operate irreversibly near thermal energy scales, the application of large deviation theory has been generally limited to idealized model systems. Applying large deviation theory to complex systems is made difficult because their numerical evaluation is restricted by the low statistical efficiency of standard Monte Carlo algorithms, which in general becomes exponentially small with the exponentially rare fluctuations that need to be sampled to converge them [11]. This makes the computational effort to compute large deviation functions intractable for many systems.

In principle, such difficulty can be mitigated by developing robust importance sampling schemes. Recent work by Nemoto et al. [12, 13], and Klymko et al. [14] suggests that significant improvements can be made in sampling efficiently by applying relatively simple forms of importance sampling. Here, we describe a general way to importance sample rare fluctuations for interacting systems that can be applied to dynamics that do not preserve detailed balance and need not be representable by an added potential [15]. The work builds upon existing diffusion Monte Carlo algorithms [16, 17] by introducing guiding functions that importance sample the trajectory space. We demonstrate our algorithm through the numerical evaluation of large deviation functions of time integrated currents within nonequilibrium

steady-states of a driven diffusion and interacting lattice models. This framework raises the possibility of efficiently computing large deviation function for higher dimensional systems, and to probe rarer fluctuations than possible with existing methods.

We consider steady states generated by a Markovian dynamics. A Markovian process generically obeys a continuity equation of the form,

$$\partial_t p_t(\mathcal{C}) = \mathcal{W} p_t(\mathcal{C}) \quad (1)$$

where $p_t(\mathcal{C})$ is the probability of observing a configuration of the system, \mathcal{C} , at a time t and \mathcal{W} is a linear operator in that configurational space that propagates the system in time. Provided \mathcal{W} is irreducible, Eq. 1 generates a unique steady state in the long time limit that in general produces non-vanishing currents and whose configurations do not necessarily follow a Boltzmann distribution. Here we consider the fluctuations of observables of the form $\mathcal{O} = \sum_{t=1}^{t_N} o(\mathcal{C}_{t+}, \mathcal{C}_{t-})$, where o is an arbitrary function of configurations at adjacent times, $t+$ and $t-$. Within the steady-state, the fluctuations of a time integrated observable can be characterized by a generating function of the form,

$$e^{\psi(\lambda)t_N} = \langle e^{-\lambda\mathcal{O}} \rangle = \sum_{\mathcal{C}(t_N)} P[\mathcal{C}(t_N)] e^{-\lambda\mathcal{O}} \quad (2)$$

where $\psi(\lambda)$ is known as the large deviation rate function, and λ is a counting field conjugate to \mathcal{O} . A member of the trajectory ensemble is denoted, $\mathcal{C}(t_N) = \{\mathcal{C}_0, \mathcal{C}_1, \dots, \mathcal{C}_{t_N}\}$, which is a vector of all configurations that a system has visited over an observation time t_N , and its likelihood is given by $P[\mathcal{C}(t_N)]$. The large deviation function is a scaled cumulant generating function, and its derivatives with respect to λ yield the time-intensive cumulants of \mathcal{O} .

In principle, $\psi(\lambda)$ is the largest eigenvalue of a tilted operator, \mathbb{W}_λ , which is constructed by taking the time derivative of the generating function in Eq. 2. This operator satisfies an eigenvalue equation,

$$\mathbb{W}_\lambda|\Xi\rangle = \psi(\lambda)|\Xi\rangle \quad (3)$$

where the large deviation function, $\psi(\lambda)$ is the largest eigenvalue and $|\Xi\rangle$ is the corresponding dominant right eigenvector [10]. For a discrete process \mathbb{W}_λ has elements

$$\mathbb{W}_\lambda(\mathcal{C}, \mathcal{C}') = \mathcal{W}(\mathcal{C}, \mathcal{C}')e^{-\lambda o(\mathcal{C}, \mathcal{C}')} (1 - \delta_{\mathcal{C}, \mathcal{C}'} - R(\mathcal{C})\delta_{\mathcal{C}, \mathcal{C}'} \quad (4)$$

where $R(\mathcal{C}) = \sum_{\mathcal{C}' \neq \mathcal{C}} \mathcal{W}(\mathcal{C}, \mathcal{C}')$ is the exit rate. For $\lambda = 0$, the tilted operator is Markovian and $\psi(0) = 0$ due to normalization. However, in general \mathbb{W}_λ does not conserve probability, and thus in a Markov chain Monte Carlo, it is necessary to keep track of the normalization through additional weight factors [18]. For the calculation of $\psi(\lambda)$, the normalization grows exponentially with λ and in a Monte Carlo algorithm, the weights used to represent the normalization have an exponentially growing variance.

It is possible to construct a Markovian process that samples the typical trajectories corresponding to the rare fluctuations that most contribute to $\psi(\lambda)$ and so avoid the exponentially growing variance. This is done through Doob's h-transformation [19]. In terms of the tilted operator and its dominant eigenfunction

$$\mathcal{W}_\lambda(\mathcal{C}, \mathcal{C}') = \Xi(\mathcal{C})\mathbb{W}_\lambda(\mathcal{C}, \mathcal{C}')\Xi^{-1}(\mathcal{C}') - \psi(\lambda), \quad (5)$$

is a canonical transformation that restores normalization. Here $\Xi(\mathcal{C}) = \langle \Xi | \mathcal{C} \rangle$ is the projection of $|\Xi\rangle$ onto the basis $\{\mathcal{C}\}$, and $\langle \Xi |$ is the dominant left eigenvector of \mathbb{W}_λ . Note that $\sum_{\mathcal{C}'} \mathcal{W}_\lambda(\mathcal{C}, \mathcal{C}') = 0$, since $\sum_{\mathcal{C}} \Xi(\mathcal{C})\mathbb{W}_\lambda(\mathcal{C}, \mathcal{C}') = \psi(\lambda)\Xi(\mathcal{C}')$. Naturally then, this auxiliary stochastic process is an optimal dynamics, since the normalization of \mathcal{W}_λ is completely independent of configuration. Indeed, if Doob's transform could be constructed, one would use this auxiliary dynamics to directly sample exponentially rare fluctuations within the nonequilibrium steady-state, as it produces the same distribution of trajectories in the long time limit as the biased path ensemble [20]. However, the transform requires that the dominant eigenvectors of the system's tilted operator be known explicitly, which is generally impossible for an interacting system.

We can, however, approximate the dominant eigenvectors of \mathbb{W}_λ . These approximate eigenvectors can then be used in an importance sampling of trajectories when evaluating large deviation functions with Monte Carlo methods. This is the basic idea of this work. Such importance sampling methodology can be incorporated into both transition path sampling [21] and the cloning algorithm [16]. Here we will consider only the cloning algorithm, which is a variant of diffusion Monte Carlo (DMC), details of which can be found in Ref. 17. In this case, incorporating importance sampling from an approximate auxiliary dynamics is directly analogous to the use

of guiding wavefunctions in quantum DMC [18, 22]. The key technical difference is that unlike the Hamiltonians in quantum DMC, \mathbb{W}_λ is not, in general, Hermitian.

Using DMC, the large deviation function can be computed with the estimator

$$\psi(\lambda) \sim \frac{1}{t_N} \ln \langle \mathbb{1} | e^{t_N \mathbb{W}_\lambda} | p_0 \rangle, \quad (6)$$

where $\langle \mathbb{1} | = \sum_{\mathcal{C}} \langle \mathcal{C} |$ is the uniform left vector, $|p_0\rangle$ is an arbitrary initial state and \sim denotes a long time limit. Since the full propagator matrix is not known explicitly, DMC Trotterizes it into short-time pieces which can be explicitly sampled [18]. The distribution p_t is represented by an ensemble of walkers, and the propagation $|p_{t+\Delta t}\rangle = \exp[\Delta t \mathbb{W}_\lambda] |p_t\rangle$ is obtained via Monte Carlo sampling. Since \mathbb{W}_λ is not normalized, the sampling procedure requires accounting for weights on the walkers. In the cloning algorithm, weights are redistributed via a branching step, where walkers are replicated or deleted in proportion to their relative weights such that the total number of walkers is kept constant.

We construct an auxiliary dynamics by approximating the dominant eigenvector, which we call the guiding distribution function (GDF), $\langle \tilde{\Xi} | = \sum_{\mathcal{C}} \tilde{\Xi}(\mathcal{C}) \langle \mathcal{C} |$. With the GDF, we can transform Eq. 6 by forming a diagonal matrix from $\tilde{\Xi}$ and

$$\psi(\lambda) \sim \frac{1}{t_N} \ln \langle \mathbb{1} | \tilde{\Xi}^{-1} [\tilde{\Xi} e^{t_N \mathbb{W}_\lambda} \tilde{\Xi}^{-1}] \tilde{\Xi} | p_0 \rangle. \quad (7)$$

The resulting transformed propagator, $\tilde{\mathbb{W}}_\lambda(\mathcal{C}, \mathcal{C}') = \tilde{\Xi}(\mathcal{C})\mathbb{W}_\lambda(\mathcal{C}, \mathcal{C}')\tilde{\Xi}^{-1}(\mathcal{C}')$, generates the importance sampled auxiliary dynamics, and has an associated dominant right eigenvector $\tilde{\Xi} | p_0 \rangle$. Physically, this similarity transform puts more weight into those transitions that will produce larger overlap with the GDF. Note that $\tilde{\mathbb{W}}_\lambda$ is only Markovian if $\langle \tilde{\Xi} | = \langle \Xi |$, which is generally not the case, and thus the problem of normalization persists. However, if $\langle \tilde{\Xi} |$ strongly overlaps with $\langle \Xi |$, the problem of normalization it is greatly mitigated, and the corresponding exponential growth of the variance is diminished. The key to efficient sampling is thus reduced to determining appropriate approximate $\langle \tilde{\Xi} |$ for specific problems.

The efficacy of the auxiliary dynamics strategy is best quantified in terms of the statistical efficiency of sampling, since this is the main source of error as systematic errors are well understood. For the cloning algorithm, the measure of interest is the number of correlated walkers N_c [12]. Correlation between walkers develops in DMC in the branching step, where the replication (or deletion) of walkers correlates their histories. In the case of perfect sampling, using the exact auxiliary dynamics, N_c is equal to 1. In the other limit, if all walkers are correlated, $N_c = N_w$, the number of walkers used in the simulation.

We first consider fluctuations of the entropy production of a driven brownian particle in a periodic potential. This is a paradigmatic model in nonequilibrium statistical mechanics [23] whose equation of motion for a position on a ring, θ , is $\partial_t \theta = F(\theta) + \eta$,

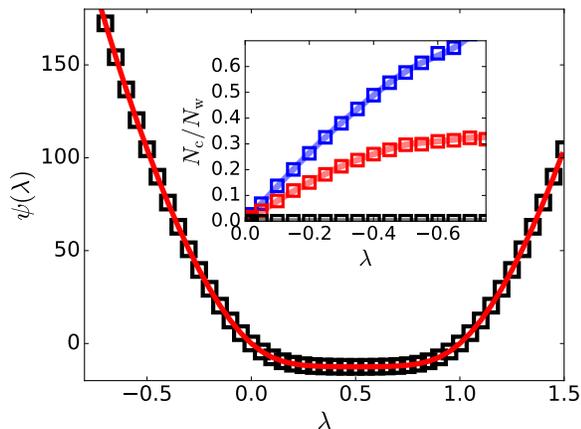


FIG. 1. Large deviation function for the entropy production of a driven brownian particle on a periodic potential with $v_o = 2$, $f = 12.5$. The main figure shows the functions computed with exact diagonalization (red) and DMC (black). The inset shows the fraction of correlated walkers without GDF (blue), or with GDF from an instanton approximation to the auxiliary process (red) or the exact auxiliary process (black).

with $F(\theta) = f - \partial_\theta V(\theta)$, where f is a constant, non-conservative force, and $V(\theta) = v_o \cos(\theta)$ is a periodic potential. The random force, η satisfies $\langle \eta(t) \rangle = 0$ and $\langle \eta(t)\eta(t') \rangle = 2\delta(t-t')$. The entropy production can be computed from $\sigma(t)t_N = \int_0^{t_N} f d\theta(t)$, which is linearly proportional to the current around the ring [24].

The tilted operator for this model is obtained by absorbing the biasing term $\exp[-\lambda t_N \sigma(t)]$ into the bare Fokker-Planck propagator, $\mathcal{W} = \partial_\theta^2 - \partial_\theta F(\theta)$, and is given by,

$$\mathbb{W}_\lambda = \mathcal{W} + 2f\lambda\partial_\theta + f\lambda(f\lambda - F(\theta)). \quad (8)$$

The last term breaks normalization and must be accommodated through branching. The first two terms represent a drift-diffusion process, states for which can be generated via an associated Langevin equation,

$$\partial_t \theta = F(\theta) - 2f\lambda + \eta. \quad (9)$$

Importance sampling this system with a GDF, $\tilde{\Xi}(\theta)$, produces the transformed propagator

$$\tilde{W}_\lambda = \mathcal{W} + 2\partial_\theta [f\lambda - \partial_\theta \ln \tilde{\Xi}(\theta)] + \tilde{\Xi}^{-1}(\theta) \mathbb{W}_\lambda^\dagger \tilde{\Xi}(\theta), \quad (10)$$

where the adjoint operator is defined as $\mathbb{W}_\lambda^\dagger = F(\theta)(\partial_\theta - f\lambda) + (\partial_\theta - f\lambda)^2$ [24]. Trajectories for \tilde{W}_λ can be generated via a Langevin dynamics similar to Eq. 9, but with an additional force $2\partial_\theta \ln \tilde{\Xi}(\theta)$, and branching weight $\tilde{\Xi}^{-1}(\theta) \mathbb{W}_\lambda^\dagger \tilde{\Xi}(\theta)$. Note that since the left eigenvector of \mathbb{W}_λ is the right eigenvector of its adjoint $\mathbb{W}_\lambda^\dagger$, if $\tilde{\Xi}(\theta) = \Xi(\theta)$ the branching term is constant and equal to $\psi(\lambda)$.

For this simple one particle system, \mathbb{W}_λ can be diagonalized by projecting it into a basis of plane waves, so

that the dominant right eigenvector is given by $\Xi(\theta) = \sum_{n \in \mathbb{Z}} c_n \exp[in\theta]$. This equation can be solved to obtain the coefficients c_n in a straightforward way [25], and the exact eigenvectors can be used to optimally guide the trajectories in DMC. To show what happens when it is not possible to obtain the exact guiding distribution, we also construct an approximate auxiliary process corresponding to an instantonic solution to the eigenvalue equation, which captures the correct limiting behavior of $\Xi(\theta)$ at large λ , where $\Xi(\theta)$ is just a constant [25].

Shown in Fig. 1 is the large deviation function computed from exact diagonalization and the DMC calculations without a GDF, with the optimal GDF, and with the instantonic GDF. All methods are able to converge $\psi(\lambda)$ to good accuracy over the range of λ , and illustrate the fluctuation theorem symmetry $\psi(\lambda) = \psi(1-\lambda)$. However, the statistical effort required to converge the different Monte Carlo calculations varies significantly. This is summarized in the inset of Fig. 1, which shows N_c as a function of λ . The number of correlated walkers increases exponentially without a guiding function, but plateaus if the instantonic guiding function is used – as for a finite t_N , this guiding function captures the limiting dynamics for large λ . Using an auxiliary process constructed from the Doob transform results in walkers that maintain equal weights and stay completely independent, with $N_c = 1$ for all times and all λ 's.

To explore our framework in an interacting system, we consider current fluctuations of a simple exclusion process (SEP) [26]. SEP is a model of transport on a lattice with L sites, defined by a set of occupation numbers, $n_i = \{0, 1\}$, e.g. $\mathcal{C} = \{0, 1, \dots, 1, 1\}$. The tilted propa-

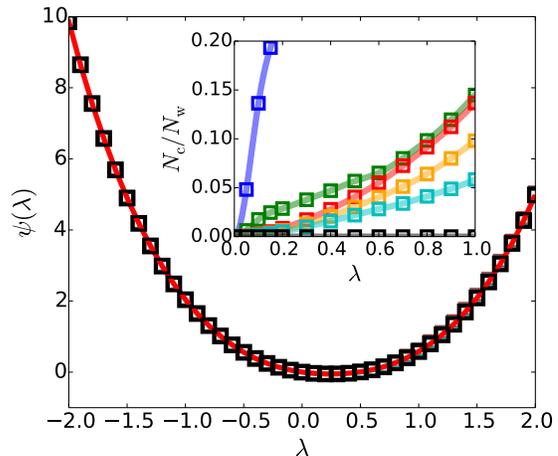


FIG. 2. Large deviation function for the mass current of an open simple exclusion process. The main figure shows the functions computed with exact diagonalization (red) and DMC (black). The inset shows the fraction of correlated walkers without guiding functions (blue), or with GDF from approximations to the auxiliary process using a uniform GDF (green) or clusters of 1 (red), 2 (orange), 4 (cyan) and 8 (black) sites.

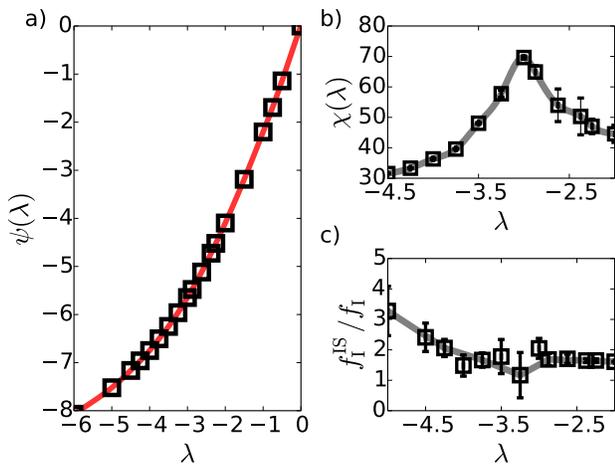


FIG. 3. Large deviation function for the mass current of a closed 2d asymmetric exclusion process. a) Large deviation function computed from DMC with importance sampling. b) Susceptibility for current fluctuations as a function of λ . c) Ratio of the fraction of independent walkers with importance sampling, f_I^{IS} , and without importance sampling, f_I . Continuous lines are guides to the eye.

gator, \mathbb{W}_λ , has elements of the form of Eq. 4, where $W(\mathcal{C}, \mathcal{C}')$ includes rates to insert and remove particles at the boundaries if the model is open, with insertion rates α and γ , and removal rates β and ν . Within the bulk of the lattice, particles move to the right with rate p and to the left with rate q , subject to the constraint of single site occupancy. The hard core constraint results in correlations between particles moving on the lattice. We aim to compute the large deviation function for mass currents, $Q(t_N)$, equal to the number of particle hops to the left minus the number of hops to the right,

$$Q(t_N) = \sum_{t=0}^{t_N} \sum_{i=0}^{L-1} \delta_{i+1}(t+1)\delta_i(t) - \delta_i(t+1)\delta_{i+1}(t) \quad (11)$$

where δ_i is the Kronecker delta function and the sum runs over the lattice site and t_N . The propagator is thus dressed by a factor of $\exp[-\lambda Q(t_N)]$, which enters into \mathbb{W}_λ in Eq. 4 with $o(\mathcal{C}, \mathcal{C}')$ being the current produced from a change in configuration \mathcal{C}' to configuration \mathcal{C} , which for single particle move is $0, \pm 1$.

For all but the smallest lattices, direct diagonalization of \mathbb{W}_λ is impossible, as the size of the matrix scales exponentially with L . However, we can find an approximate set of eigenvectors using a cluster based mean-field approximation [24]. For example, we can write $|\tilde{\Xi}\rangle$ as product state of single sites expanded in a basis of single particle states, $|\tilde{\Xi}\rangle = \prod_{i=1}^L \sum_{n=0,1} \xi_i(n) |n_i\rangle$ where $\xi_i(n)$ are the site expansion coefficients. Clusters of larger size can also be used, as well as other expansions of low rank, such as matrix product states [27]. Figure 2 shows the results of using the cluster mean-field ansatz as the GDF. We find that for an $L = 8$ site, symmetric

SEP model[28], all DMC calculations agree with the numerically exact result, and again illustrate fluctuation theorem symmetry, though this time ψ is symmetric about the current's affinity [3]. The statistical effort needed to converge each calculation is decreased by several orders of magnitude when a GDF is used. As shown in the inset of Fig. 2, even with an auxiliary dynamics computed from the single-site mean field theory, the fraction of independent walkers, $f_I = 1 - N_c/N_w$, is increased by a factor of 40, and this efficiency is systematically improved with auxiliary dynamics computed from larger cluster states. As before, if the exact auxiliary process is used, corresponding to a cluster of 8 sites, $N_c = 1$ for all times and all λ 's.

As a final illustration of the auxiliary dynamics framework we have introduced, we consider a 2D generalization of the closed SEP model in the presence of a weak external field that biases transport in one direction [24]. This system has been considered recently [29], where it was found that its large deviation function, for current fluctuations in the direction of the driving, exhibits a dynamical phase transition. For small λ , the system is in a homogenous phase, while for large negative λ the system phase separates, forming a traveling wave state in the direction of the biased current. We find critical behavior for a 12×12 lattice as illustrated in Figs. 3a,b, where for $\lambda \approx -3$ the fluctuations in the current, $\chi(\lambda) = d^2\psi/d\lambda^2$, are maximized and presumed to diverge in the infinite system limit [30]. Beyond this critical value, the state of the system is fluctuation dominated, and as such serves as a good test of our importance sampling methodology. Shown in Fig. 3c is the ratio of the fraction of independent walkers, f_I , computed using a 4×2 per site cluster mean field GDF and without importance sampling, as a function of λ . For small $|\lambda|$ the bare dynamics is capable of sampling the biased distribution and the enhancement from importance sampling is ~ 2 . However, even for the traveling wave state, where the system is not well described by mean field theory, we find an increased sampling efficiency by a factor of 2-4 over bare sampling [24].

In conclusion, the application of importance sampling to nonequilibrium steady states makes computing large deviation functions possible even in complex systems. Our framework enables the integration of approximate solutions to systematically improve computations. In future this opens the possibility to study ever larger systems, for long times, with increased molecular resolution.

Acknowledgements: D.T.L was supported by UC Berkeley College of Chemistry. U. R. was supported by the Simons Collaboration on the Many-Electron Problem and the California Institute of Technology. G. K.-L. C. is a Simons Investigator in Theoretical Physics and was supported by the California Institute of Technology and the US Department of Energy, Office of Science via de-sc0018140. These calculations were preformed with CANSS, available at <https://github.com/ushnishray/CANSS>.

-
- [1] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
- [2] J. Kurchan, Journal of Physics A: Mathematical and General **31**, 3719 (1998).
- [3] J. L. Lebowitz and H. Spohn, Journal of Statistical Physics **95**, 333 (1999).
- [4] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995).
- [5] A. C. Barato and U. Seifert, Physical review letters **114**, 158101 (2015).
- [6] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Physical review letters **116**, 120601 (2016).
- [7] T. Harada and S.-i. Sasa, Physical review letters **95**, 130602 (2005).
- [8] T. Speck and U. Seifert, EPL (Europhysics Letters) **74**, 391 (2006).
- [9] M. Baiesi, C. Maes, and B. Wynants, Physical review letters **103**, 010602 (2009).
- [10] H. Touchette, Physics Reports **478**, 1 (2009).
- [11] U. Ray, G. K. Chan, and D. T. Limmer, arXiv:1708.00459 (2017).
- [12] T. Nemoto, F. Bouchet, R. L. Jack, and V. Lecomte, Phys. Rev. E **93**, 062123 (2016).
- [13] T. Nemoto, R. L. Jack, and V. Lecomte, Physical Review Letters **118**, 115702 (2017).
- [14] K. Klymko, P. L. Geissler, J. P. Garrahan, and S. Whitlam, arXiv:1707.00767 (2017).
- [15] R. L. Jack and P. Sollich, The European Physical Journal Special Topics **224**, 2351 (2015).
- [16] C. Giardinà, J. Kurchan, and L. Peliti, Phys. Rev. Lett. **96**, 120603 (2006).
- [17] C. Giardinà, J. Kurchan, V. Lecomte, and J. Tailleur, Journal of statistical physics **145**, 787 (2011).
- [18] W. Foulkes, L. Mitas, R. Needs, and G. Rajagopal, Reviews of Modern Physics **73**, 33 (2001).
- [19] R. Chetrite and H. Touchette, in *Annales Henri Poincaré*, Vol. 16 (Springer, 2015) pp. 2005–2057.
- [20] R. Chetrite and H. Touchette, Physical review letters **111**, 120601 (2013).
- [21] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Annual review of physical chemistry **53**, 291 (2002).
- [22] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
- [23] U. Seifert, Reports on Progress in Physics **75**, 126001 (2012).
- [24] Supplemental Material.
- [25] P. Tsobgni Nyawo and H. Touchette, Phys. Rev. E **94**, 032101 (2016).
- [26] B. Schmittmann and R. K. Zia, Phase transitions and critical phenomena **17**, 3 (1995).
- [27] M. Gorissen, J. Hooyberghs, and C. Vanderzande, Physical Review E **79**, 020101 (2009).
- [28] The model we study has $p = q = 0.5$, $\alpha = \beta = 0.9$ and $\gamma = \nu = 0.1$.
- [29] N. Tizón-Escamilla, C. Pérez-Espigares, P. L. Garrido, and P. I. Hurtado, Physical Review Letters **119**, 090602 (2017).
- [30] For the 2D SEP model, we study a 12×12 lattice, with hopping rates in the x and y direction that are now vectorial with components $\mathbf{p} = \{0.22, 0.5\}$ and $\mathbf{q} = \{1.15, 0.5\}$ and bias on the current in just the x direction as in Eq. 11.

Exact fluctuations of nonequilibrium steady states from approximate auxiliary dynamics: Supplementary Information

Ushnish Ray and Garnet Kin-Lic Chan
*Division of Chemistry and Chemical Engineering,
California Institute of Technology, Pasadena, CA 91125*

David T. Limmer
*Department of Chemistry, University of California, Berkeley, CA 94609
Kavli Energy NanoScience Institute, Berkeley, CA 94609 and
Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94609*
(Dated: September 11, 2017)

arXiv:1708.09482v2 [cond-mat.stat-mech] 8 Sep 2017

GUIDING DISTRIBUTION FUNCTION IN THE CONTINUUM

Here we derive the tilted propagator used for importance sampling the driven brownian walker in the main text and discuss the adjoint operator for generating the Langevin dynamics. The Langevin equation that governs the overdamped motion of a particle in the presence of an external force, $F(\theta)$, is given by,

$$\partial_t \theta = F(\theta) + \eta \quad (\text{S1})$$

As in the main text, the associated Fokker-Planck equation corresponding to this Langevin equation is

$$\partial_t p_t(\theta) = \mathcal{W} p_t(\theta) \quad (\text{S2})$$

where $p_t(\theta)$ is the probability of observing the particle at θ at a time t and

$$\mathcal{W} = -\partial_\theta(F - \partial_\theta) \quad (\text{S3})$$

is the Fokker-Planck operator. We are interested in the large deviation function of the entropy production, $s(t_N) = \sigma(t_N)t_N$, which is proportional to the current around the ring, unwrapped so that winding numbers are included,

$$s(t_N) = \int_0^{t_N} f d\theta(t) = f\Delta\theta(t_N) \equiv fx \quad (\text{S4})$$

The generating function is related to the probability $p(\theta, s, t_N)$ of finding the particle at θ with entropy produced s at a time t_N by the Laplace transform,

$$\rho(\theta, \lambda, t) = \int ds e^{-\lambda s} p(\theta, s, t_N). \quad (\text{S5})$$

Since we are interested in the behavior of probability distribution conditioned on x and θ and since both of them share the same noise source η , we expand the Fokker-Planck operator via $\partial_\theta \rightarrow \partial_\theta + \partial_x$ to obtain:

$$\tilde{\mathcal{W}} = \mathcal{W} + (2\partial_\theta - F(\theta))\partial_x + \partial_x^2 \quad (\text{S6})$$

and the corresponding modified Fokker-Planck equation is

$$\partial_t p(\theta, s, t) = \tilde{\mathcal{W}} p(\theta, s, t). \quad (\text{S7})$$

By differentiating Eq. S5 with respect to t and inserting Eq. S7 we get

$$\partial_t \rho(\theta, \lambda, t) = \int ds e^{-\lambda s} \tilde{\mathcal{W}} p(\theta, s, t). \quad (\text{S8})$$

Now,

$$\begin{aligned} \int ds e^{-\lambda s} \tilde{\mathcal{W}} p(\theta, s, t) &= \int ds e^{-\lambda s} \mathcal{W} p(\theta, s, t) \\ &+ \int ds e^{-\lambda s} (2\partial_\theta - F(\theta))\partial_x p(\theta, s, t) \\ &+ \int ds e^{-\lambda s} \partial_x^2 p(\theta, s, t) \end{aligned} \quad (\text{S9})$$

Performing integration by parts we get:

$$\begin{aligned} \partial_t \rho(\theta, \lambda, t) &= \mathcal{W} \int ds e^{-\lambda s} p(\theta, s, t) \\ &+ (2\partial_\theta - F(\theta))(f\lambda) \int ds e^{-\lambda s} p(\theta, s, t) \\ &+ (f\lambda)^2 \int ds e^{-\lambda s} p(\theta, s, t) \\ &= \mathbb{W}_\lambda \int ds e^{-\lambda s} p(\theta, s, t) = \mathbb{W}_\lambda \rho(\theta, \lambda, t), \end{aligned} \quad (\text{S10})$$

where,

$$\mathbb{W}_\lambda = \mathcal{W} + 2f\lambda\partial_\theta + f\lambda(f\lambda - F(\theta)), \quad (\text{S11})$$

is the tilted operator given by Eq. 8 in the main text used to obtain the modified or tilted dynamics [1]. The mapping of the second-order differential operator onto a stochastic diffusion process then follows the well-known Feynman-Kac theorem [2].

Following the general derivation of importance sampling in the main text, we use the specific form of \mathbb{W}_λ to get

$$\tilde{\Xi} \mathbb{W}_\lambda (\tilde{\Xi}^{-1} \rho) = \rho [\tilde{\Xi} \mathbb{W}_\lambda \tilde{\Xi}^{-1} + \partial_\theta \zeta] - \partial_\theta (\zeta - \partial_\theta) \rho, \quad (\text{S12})$$

with $\zeta = F - 2\lambda + 2\partial_\theta \ln \tilde{\Xi}$. The last term corresponds to a Fokker-Planck type of operator with a force term equal to ζ and can be used to generate trajectories. The first term changes normalization and is, therefore, used for branching. The equation of motion for θ is obtained from the adjoint operator $\mathbb{W}_\lambda^\dagger$, where,

$$\tilde{\Xi}^{-1} \mathbb{W}_\lambda^\dagger \tilde{\Xi} = \tilde{\Xi} \mathbb{W}_\lambda \tilde{\Xi}^{-1} + \partial_\theta \zeta \quad (\text{S13})$$

which follows from integration by parts. The specific form of the branching term following the operators in the main text is given by

$$\tilde{\Xi}^{-1} \mathbb{W}_\lambda^\dagger \tilde{\Xi} = \frac{1}{\tilde{\Xi}} \frac{d^2 \tilde{\Xi}}{d\theta^2} + (F - 2\lambda f) \frac{d \ln \tilde{\Xi}}{d\theta} + f\lambda(f\lambda - F(\theta)), \quad (\text{S14})$$

where $F(\theta) = f - \partial_\theta V(\theta)$, with f a constant, nonconservative force, and $V(\theta) = v_o \cos(\theta)$ is a periodic potential.

GUIDING DISTRIBUTION FUNCTIONS FROM DISCRETE MEAN-FIELD SOLUTIONS

In this section we outline the procedure needed to generate mean-field (MF) and cluster solutions that form

the guiding distribution functions in our discrete models. We will illustrate the procedure for the SEP but it can be easily generalized to other models.

For the open boundary SEP the tilted matrix may be written as:

$$\mathbb{W}_\lambda = w_L + \sum_{i=1}^L w_i + w_R \quad (\text{S15})$$

where the single particle transition matrices $\{w_L, w_i, w_R\}$ and their operator representations are given by

$$w_L = \begin{bmatrix} -\alpha & \gamma e^\lambda \\ \alpha e^{-\lambda} & -\gamma \end{bmatrix} \quad w_R = \begin{bmatrix} -\nu & \beta e^{-\lambda} \\ \nu e^\lambda & -\beta \end{bmatrix} \quad (\text{S16})$$

and

$$w_i = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -q & p e^{-\lambda} & 0 \\ 0 & q e^\lambda & -p & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{S17})$$

These matrices can be combined to form a many-body matrix that in operator notation is written as

$$\begin{aligned} \mathbb{W}_\lambda = & -\alpha(\mathbb{1} - \hat{n}_1) + \alpha e^{-\lambda} \hat{c}_1^\dagger + \gamma e^\lambda \hat{c}_1 - \gamma \hat{n}_1 + p e^{-\lambda} \hat{c}_2^\dagger \hat{c}_1 - p \hat{n}_1 (\mathbb{1} - \hat{n}_2) \\ & + \sum_{i=2}^{L-1} p e^{-\lambda} \hat{c}_{i+1}^\dagger \hat{c}_i - p \hat{n}_i (\mathbb{1} - \hat{n}_{i+1}) + q e^\lambda \hat{c}_{i-1}^\dagger \hat{c}_i - q \hat{n}_i (\mathbb{1} - \hat{n}_{i-1}) \\ & - \nu (\mathbb{1} - \hat{n}_L) + \nu e^\lambda \hat{c}_L^\dagger + \beta e^{-\lambda} \hat{c}_L - \beta \hat{n}_L + q e^\lambda \hat{c}_{L-1}^\dagger \hat{c}_L - q \hat{n}_L (\mathbb{1} - \hat{n}_{L-1}). \end{aligned} \quad (\text{S18})$$

Here \hat{c}_i^\dagger creates a particle at site i , \hat{c}_j destroys a particle at site j , and \hat{n}_i counts the number of particles at site i . The combined operators $\hat{c}_i^\dagger \hat{c}_j$ correspond to kinetic-like terms that move a particle from site j to i , and $\hat{n}_i (\mathbb{1} - \hat{n}_j)$ represents a hard-core interaction that prevents double occupation of sites. The exact solution is the usual eigenvalue problem $\mathbb{W}_\lambda |\Xi\rangle = \varepsilon |\Xi\rangle$ where $(\varepsilon, |\Xi\rangle)$ is a particular eigenpair (the inverse eigenvectors $\{|\Xi\rangle\}$ can be constructed from $\{|\Xi\rangle\}$).

An obvious route to explore in constructing approximate GDF is to use a mean-field (MF) solution. The MF approach is to approximate the many-body state as a product state. Starting with a product of single site states, we can systematically improve our results by moving to products of cluster states, where interactions in the cluster are treated explicitly, while the links between clusters are treated at the MF level. We first illustrate the procedure for single site clusters and then show how

to generalize to multi-site clusters.

Site-Decoupled Mean-Field

For the single site clusters we approximate the solution by the form $|\tilde{\Xi}\rangle = \prod_{i=1}^L |\xi_i\rangle$, where the single site state $|\xi_i\rangle = \sum_{n=0}^1 \xi_i(n) |n_i\rangle$ is written in a basis of occupation numbers $\{|n\rangle\}$. Here $\xi_i(n)$ is a scalar function dependent on the occupation number n . Additionally, the left states $\langle \xi_i | = \sum_{n=0}^1 \xi_i(n) \langle n_i |$ are biorthogonal to the right states $|\xi_i\rangle$, i.e. $\langle \xi_i | \xi_j \rangle = \delta_{ij}$.

To determine the coefficients $\xi_i(n)$, we will use a variational procedure to extremize $\langle \tilde{\Xi} | \mathbb{W}_\lambda | \tilde{\Xi} \rangle$ subject to normalization. For small variations, we see that

$$|\delta \tilde{\Xi}\rangle = \sum_{i=1}^L \left[\sum_{n_i=0}^1 \delta \xi_i(n) |n_i\rangle \right] \prod_{j \neq i} \left[\sum_{n_j=0}^1 \xi_j(n) |n_j\rangle \right] \quad (\text{S19})$$

and, independently, this form also holds for the inverse $\langle \delta \tilde{\Xi} |$. We may treat the variation of $|\tilde{\Xi}\rangle$ and its inverse independently.

Consider variations with respect to the right state (the same procedure can be applied to the left state):

$$\langle \tilde{\Xi} | \mathbb{W}_\lambda | \delta \tilde{\Xi} \rangle - \varepsilon \langle \tilde{\Xi} | \delta \tilde{\Xi} \rangle = 0 \quad (\text{S20})$$

where ε is the Lagrange multiplier for the normalization. We can, then, write $\mathbb{W}_\lambda = (\mathcal{W}_\lambda - V_i) + V_i$ where V_i represents all terms in \mathcal{W}_λ that contains terms involving site i . Then we obtain the following equation for each i :

$$\sum_i \left\{ \langle \tilde{\Xi}_i | V_i | \delta \tilde{\Xi}_i \rangle - \varepsilon_i \langle \tilde{\Xi}_i | \delta \tilde{\Xi}_i \rangle \right\} = 0, \quad (\text{S21})$$

which is equivalent to an eigenvalue equation with eigenvalue ε_i , where we have defined

$$\varepsilon_i = \varepsilon - \left[\prod_{j \neq i} \langle \tilde{\Xi}_j | \right] (\mathbb{W}_\lambda - V_i) \left[\prod_{k \neq i} | \tilde{\Xi}_k \rangle \right] \quad (\text{S22})$$

As we are interested in the maximal eigenvalue/eigenvector pair of \mathbb{W}_λ , we are interested in the maximal eigenvalue/eigenvector pair for each site i .

For the SEP \mathbb{W}_λ , Eq. S21 implies solving the following decoupled eigenvalue equation for $i = 1$

$$\begin{aligned} & [(-\alpha - q \langle \hat{n}_2 \rangle)(\mathbb{1} - \hat{n}_1) + (pe^{-\lambda} \langle \hat{c}_2^\dagger \rangle + \gamma e^\lambda) \hat{c}_1 \\ & + (\alpha e^{-\lambda} + qe^\lambda \langle \hat{c}_2 \rangle) \hat{c}_1^\dagger + (-\gamma - p(1 - \langle \hat{n}_2 \rangle)) \hat{n}_1] |\xi_1\rangle \\ & = \varepsilon_1 |\xi_1\rangle \end{aligned} \quad (\text{S23})$$

For $i = L$, the eigenvalue equation is

$$\begin{aligned} & [(-\nu - p \langle \hat{n}_{L-1} \rangle)(\mathbb{1} - \hat{n}_L) + (qe^\lambda \langle \hat{c}_{L-1}^\dagger \rangle + \beta e^{-\lambda}) \hat{c}_{L-1} \\ & + (\nu e^\lambda + pe^{-\lambda} \langle \hat{c}_{L-1} \rangle) \hat{c}_L^\dagger + (-\beta - q(1 - \langle \hat{n}_{L-1} \rangle)) \hat{n}_L] |\xi_L\rangle \\ & = \varepsilon_L |\xi_L\rangle \end{aligned} \quad (\text{S24})$$

Finally for all other i , the eigenvalue equation is

$$\begin{aligned} & [(-p \langle \hat{n}_{i-1} \rangle - q \langle \hat{n}_{i+1} \rangle)(\mathbb{1} - \hat{n}_i) \\ & + (pe^{-\lambda} \langle \hat{c}_{i+1}^\dagger \rangle + qe^\lambda \langle \hat{c}_{i-1}^\dagger \rangle) \hat{c}_i \\ & + (qe^\lambda \langle \hat{c}_{i+1} \rangle + pe^{-\lambda} \langle \hat{c}_{i-1} \rangle) \hat{c}_i^\dagger \\ & + (-p(1 - \langle \hat{n}_{i+1} \rangle) - q(1 - \langle \hat{n}_{i-1} \rangle)) \hat{n}_i] |\xi_i\rangle \\ & = \varepsilon_i |\xi_i\rangle. \end{aligned} \quad (\text{S25})$$

The maximal eigenvalue/eigenvector pair approximation can be obtained by solving the above equations for each site i and choosing the set of states $\{|\xi_i\rangle\}$ corresponding to the largest eigenvalues $\{\varepsilon_i\}$. Notice that even when collecting terms involving site i , there will inevitably be terms involving other sites due to the two-body interaction present in the matrix (S18). The natural way to solve this system, thus, requires the use

of a self-consistent procedure. We start with an initial set of guess values for $\langle \hat{n}_i \rangle$ and $\langle \hat{c}_i \rangle$ and proceed to solve the individual eigenvalue problems for each site. At the end of each iteration we use the states $\{|\xi_i\rangle\}$ to recompute $\langle \hat{n}_i \rangle$, $\langle \hat{c}_i \rangle$, and $\langle \hat{c}_i^\dagger \rangle$ and use them for the next iteration. This is continued until $\{\langle \hat{n}_i \rangle, \langle \hat{c}_i \rangle, \langle \hat{c}_i^\dagger \rangle\}$ do not change.

Once the solutions have converged it is straightforward to obtain the MF estimate of the LDF, $\epsilon(\lambda) = \langle \tilde{\Xi} | \mathbb{W}_\lambda | \tilde{\Xi} \rangle$. The MF approximation to the state $|\tilde{\Xi}\rangle$ is precisely the GDF that we need to construct the auxiliary process that will importance sample the LDF. The effective matrices incorporating the importance sampling are:

$$\tilde{w}_L = \begin{bmatrix} -\alpha & \gamma e^\lambda \frac{\xi_1(0)}{\xi_1(1)} \\ \alpha e^{-\lambda} \frac{\xi_1(1)}{\xi_1(0)} & -\gamma \end{bmatrix} \quad (\text{S26})$$

$$\tilde{w}_i = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -q & pe^{-\lambda} \frac{\xi_i(0)\xi_{i+1}(1)}{\xi_i(1)\xi_{i+1}(0)} & 0 \\ 0 & qe^\lambda \frac{\xi_{i-1}(1)\xi_i(0)}{\xi_{i-1}(0)\xi_i(1)} & -p & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{S27})$$

$$\tilde{w}_R = \begin{bmatrix} -\nu & \beta e^{-\lambda} \frac{\xi_L(0)}{\xi_L(1)} \\ \nu e^\lambda \frac{\xi_L(1)}{\xi_L(0)} & -\beta \end{bmatrix} \quad (\text{S28})$$

Notice that these are not normalized and their renormalization determines the branching weights. In the text, we use a first order Trotterization on the short-time importance-sampled propagator $(\langle \tilde{\Xi} | e^{dt \mathbb{W}_\lambda} | \tilde{\Xi} \rangle)$ to obtain the Markovian transition probability matrix $\tilde{U}_\lambda \equiv \mathbb{1} + dt \langle \tilde{\Xi} | \mathbb{W}_\lambda | \tilde{\Xi} \rangle$. Therefore, \tilde{U}_λ follows directly from the transformed components of \mathbb{W}_λ , and is given by

$$\tilde{U}_\lambda = \tilde{u}_L + \sum_{i=1}^L \tilde{u}_i + \tilde{u}_R, \quad (\text{S29})$$

where,

$$\tilde{u}_L = \begin{bmatrix} 1 - \alpha dt & \gamma e^\lambda \frac{\xi_1(0)}{\xi_1(1)} \\ \alpha e^{-\lambda} \frac{\xi_1(1)}{\xi_1(0)} & 1 - \gamma dt \end{bmatrix} \quad (\text{S30})$$

$$\tilde{u}_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - qdt & pe^{-\lambda} \frac{\xi_i(0)\xi_{i+1}(1)}{\xi_i(1)\xi_{i+1}(0)} & 0 \\ 0 & qe^\lambda \frac{\xi_{i-1}(1)\xi_i(0)}{\xi_{i-1}(0)\xi_i(1)} & 1 - pdt & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{S31})$$

$$\tilde{u}_R = \begin{bmatrix} 1 - \nu dt & \beta e^{-\lambda} \frac{\xi_L(0)}{\xi_L(1)} \\ \nu e^\lambda \frac{\xi_L(1)}{\xi_L(0)} & 1 - \beta dt \end{bmatrix} \quad (\text{S32})$$

are the associated transition probabilities at the ends of the lattice, \tilde{u}_L and \tilde{u}_R , or in its interior, \tilde{u}_i . At every time step or Monte Carlo sweep, the current state $|\mathcal{C}\rangle = |n_1 n_2 \dots n_L\rangle$ that corresponds to a column of \tilde{U}_λ is used to propose moves such that the outgoing state $|\mathcal{C}'\rangle$ has the probability $\tilde{U}_\lambda(\mathcal{C}', \mathcal{C})/\mathcal{N}(\mathcal{C})$ of being accepted, where $\mathcal{N}(\mathcal{C}) \equiv \sum_{\mathcal{C}'} \tilde{U}_\lambda(\mathcal{C}', \mathcal{C})$ is the normalization factor. Over the course of the short trajectory generated in between branching steps, a walker's weight is accumulated as a product of these normalization factors.

Cluster Approach

The idea of a site-decoupled mean-field can be extended to multiple sites collected into clusters. The

MF ansatz is then $|\tilde{\Xi}\rangle = \prod_{i=1}^{L/c_L} |\xi_i\rangle$ where c_L is the number of sites that constitutes a cluster. $|\xi_i\rangle$ is expanded in the occupation basis of the cluster $|\xi_i\rangle = \sum_{\{n\}} \xi_i(n_{(i-1)c_L+1}, \dots, n_{ic_L}) |n_{(i-1)c_L+1}, \dots, n_{ic_L}\rangle$. For each cluster i , we can write down the analogous V_i operator which contains all terms in Eq. S18 that involve sites in the cluster. All terms involving only sites inside the given cluster are treated exactly (i.e. treated as operators) while terms that involve sites that are inside two different clusters are split up via the MF approximation $\hat{A}_i \hat{B}_j \sim \langle \hat{A}_i \rangle \hat{B}_i + \hat{A}_i \langle \hat{B}_i \rangle - \langle \hat{A}_i \rangle \langle \hat{B}_i \rangle$, where as before $\langle \hat{B}_j \rangle$ and $\langle \hat{A}_i \rangle$ are determined self-consistently.

More explicitly, the mean-field cluster V_i operator is

$$V_i = \hat{L}(j = c_L(i-1) + 1) + \hat{R}(j = c_L i - 1) + \sum_{j=c_L(i-1)+1}^{c_L i - 1} pe^{-\lambda} \hat{c}_{j+1}^\dagger \hat{c}_j - p \hat{n}_j (\mathbb{1} - \hat{n}_{j+1}) + qe^\lambda \hat{c}_j^\dagger \hat{c}_{j+1} - q \hat{n}_{j+1} (\mathbb{1} - \hat{n}_j) \quad (\text{S33})$$

where

$$\begin{aligned} \hat{L}(j) &= -p \langle \hat{n}_{j-1} \rangle (\mathbb{1} - \hat{n}_j) + qe^\lambda \langle \hat{c}_{j-1}^\dagger \rangle \hat{c}_j + pe^{-\lambda} \langle \hat{c}_{j-1} \rangle \hat{c}_j^\dagger - q \hat{n}_j (1 - \langle \hat{n}_{j-1} \rangle) & \text{for } i > 1 \\ \hat{L}(j) &= -\alpha (\mathbb{1} - \hat{n}_1) + \alpha e^{-\lambda} \hat{c}_1^\dagger + \gamma e^\lambda \hat{c}_1 - \gamma \hat{n}_1 & \text{for } i = 1 \\ \hat{R}(j) &= -q \langle \hat{n}_{j+1} \rangle (\mathbb{1} - \hat{n}_j) + pe^{-\lambda} \langle \hat{c}_{j+1}^\dagger \rangle \hat{c}_j + qe^\lambda \langle \hat{c}_{j+1} \rangle \hat{c}_j^\dagger - p \hat{n}_j (1 - \langle \hat{n}_{j+1} \rangle) & \text{for } i < L/c_L \\ \hat{R}(j) &= -\nu (\mathbb{1} - \hat{n}_L) + \nu e^\lambda \hat{c}_L^\dagger + \beta e^{-\lambda} \hat{c}_L - \beta \hat{n}_L & \text{for } i = L/c_L \end{aligned}$$

This treatment ensures that all cluster based eigenvalue problems can be solved separately using only expectation values to estimate the couplings between clusters. The latter couplings are calculated separately at the end of each self-consistent step. This is continued until the averages do not change. Once the MF calculations converge we will obtain the required GDF $\langle \tilde{\Xi} |$ needed to construct the generator of auxiliary dynamics, similar to the case for the single site MF. However, now the ξ_i involve multiple sites and thus the proposal matrix must be updated accordingly to distinguish between inter- and intra- cluster states.

CALCULATION DETAILS FOR RESULTS IN THE MAIN TEXT.

Driven brownian motion: All of the calculations on the driven brownian walker were accomplished with a second order stochastic Runge-Kutta integrator with a timestep of 0.01. Observation times of 20.0 were

needed to converge the calculations and branching steps attempted every 0.05 unit of time.

1D SEP: For the 1D system consisting of $L = 8$ sites, clusters were made using $c_L = 1, 2, 4$ sites. All calculations were done using $N_w = 20^4$ walkers, with a time step $dt = 0.001$, observation time $t_N = 1$ with branching occurring at intervals of time $t_{\text{int}} = 0.01$. In the main text, we have used calculations done with cluster size of 0 (“zeroth-order” MF) to indicate a sampling strategy where the exponential $e^{-\lambda \mathcal{O}}$ has been absorbed directly into the proposed Monte Carlo moves. This essentially means $\xi_i = 1$ for equations S30-S32. We note that this last way of sampling should always be used if no higher order MF solutions are available. For the 2D WASEP system we discuss next, the bare sampling strategy is to at least use a zeroth-order MF. Generating trajectories from the unbiased distribution, i.e., without directly incorporating the exponential in the proposed Monte Carlo moves makes it impossible to converge calculations for the range of λ we explore and the number of walkers we deploy.

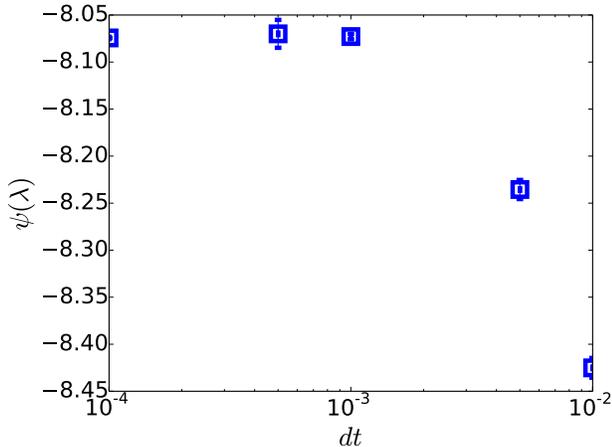


FIG. S1. Convergence of the 2D SEP DMC simulations with integration time step for $\lambda = -6$.

2D WASEP: The 2D WASEP system discussed in the main text is a generalization of the 1D model to an $L \times L$ lattice with periodic boundary conditions. Along the principal direction, x , particles are subject to biased hopping rates that are scaled by the length of the system $p_x = e^{-E/L_x}/2$ along x and $q_x = e^{E/L_x}/2$ along $-x$ with $L_x = 12$ and $E = 10$. Along the y direction particles diffuse symmetrically, with rates $p_y = q_y = 0.5$ [3]. These calculations are resource intensive so a careful study of convergence and statistical properties is needed to be able to compute the susceptibility $\chi(\lambda)$. Towards this end we determine the time step error associated with Trotterization as shown in Fig. S1 for the largest λ value that we wish to access, since this sets the upper-bound on the error. We find that $dt = 0.001$ is sufficient to converge the error.

The second major source of error in these calculations is the systematic error due to finite walker population. Since the variance grows exponentially with $|\lambda|$ [4], it is sufficient to determine the largest number of walkers needed for maximal $|\lambda|$ we use in our simulations. Fig. S2 shows the convergence of $\psi(\lambda)$ with N_w for $\lambda = -5.0$ both using and not using importance sampling. It is evident that using auxiliary dynamics, even with $N_w = 5 \times 10^5$ our results would have been sufficiently converged. Comparing against calculations without auxiliary dynamics we have ~ 4 -fold increase in efficiency mirroring the ratio of independent walkers of Fig. 3c in the main text. Despite the mean field GDF employed for this model not being particularly good at large $|\lambda|$, as the traveling wave state is not well described by a small single cluster, we still get a factor of 2-4 reduction in the required number of walkers to converge results (c.f. Fig. 3c in main text).

Following this analysis, we used $N_w = 5 \times 10^5$ walk-

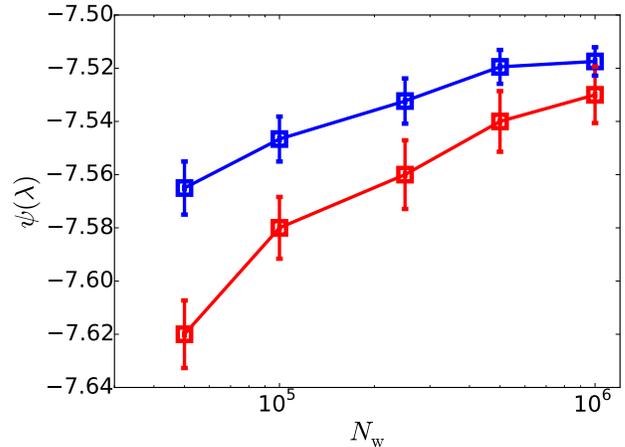


FIG. S2. Systematic error associated with walker population for 2D SEP DMC simulations for the large deviation function. Shown in red are the results for no importance sampling and in blue the results of the cluster GDF. Both were computed at $\lambda = -5.0$.

ers for $\lambda \leq -3.0$ with $t_N = 100$ where branching was done every $t_{\text{int}} = 0.01$. For $\lambda > -3.0$ simulations used $N_w = 10^6$ walkers with $t_N = 72$ and $t_{\text{int}} = 1.4$. In Fig. 3c of the main text the ratio of independent clones was determined using independent clone counts at time $t = 0.32t_N$. Fig. S3 shows a comparison of the fraction of independent walkers, $f_I(t)$, along the entire t_N trajectory for $\lambda = -5.0$. It highlights the importance of using GDF for sampling purposes to ensure that there are enough uncorrelated contributions to the estimator.

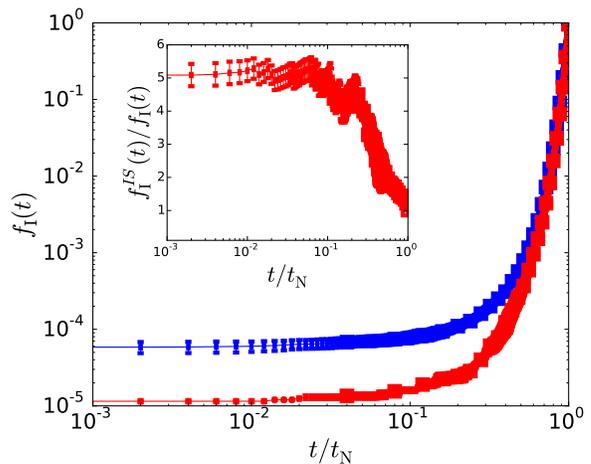


FIG. S3. Fraction of independent walkers $f_I(t)$ as a function of simulation time for $\lambda = -5.0$ using GDF (blue) and not using GDF (red). Inset shows the ratio of fraction of independent walkers with GDF and without as a function of time.

-
- [1] J. Mehl, T. Speck, and U. Seifert, *Physical Review E* **78**, 011123 (2008).
- [2] R. Chetrite and H. Touchette, in *Annales Henri Poincaré*, Vol. 16 (Springer, 2015) pp. 2005–2057.
- [3] N. Tizón-Escamilla, C. Pérez-Espigares, P. L. Garrido, and P. I. Hurtado, *Physical Review Letters* **119**, 090602 (2017).
- [4] U. Ray, G. K. Chan, and D. T. Limmer, arXiv:1708.00459 (2017).