

Constructing Auxiliary Dynamics for Nonequilibrium Stationary States by Variance Minimization

Ushnish Ray* and Garnet Kin-Lic Chan†

*Division of Chemistry and Chemical Engineering,
California Institute of Technology, Pasadena, CA 91125, USA*

(Dated: December 23, 2019)

We present a strategy to construct guiding distribution functions (GDFs) based on variance minimization. Auxiliary dynamics via GDFs mitigates the exponential growth of variance as a function of bias in Monte Carlo estimators of large deviation functions. The variance minimization technique exploits the exact properties of eigenstates of the tilted operator that defines the biased dynamics in the nonequilibrium system. We demonstrate our techniques in two classes of problems. In the continuum, we show that GDFs can be optimized to study interacting driven diffusive systems where the efficiency is systematically improved by incorporating higher correlations into the GDF. On the lattice, we use a correlator product state ansatz to study the 1D WASEP. We show that with modest resources we can capture the features of the susceptibility in large systems that marks the phase transition from uniform transport to a traveling wave state. Our work extends the repertoire of tools available to study nonequilibrium properties in realistic systems.

I. INTRODUCTION

Large deviation theory (LDT) is a framework to extend the formalism of equilibrium statistical mechanics to nonequilibrium systems [1]. Much of LDT is concerned with summarizing the dynamics of the system as expressed via the fluctuations of typical trajectories. Ensembles of rare trajectories display fascinating behavior reminiscent of phase transitions and criticality. Recent work illustrates such dynamical behavior in both lattice systems such as simple exclusion processes [2–9], constrained kinetic models [10, 11], models of self-assembly [12–14], dissipative hydrodynamics [15, 16], and in continuum systems in the form of driven or active Brownian particles [14, 17–21], as well as open quantum systems [22–24]. Recently LDT has also been shown to offer a route to calculating nonlinear transport coefficients [25].

Accessing properties of interest, in all but the simplest systems, requires numerical tools. In the context of LDT, this takes the form of sampling techniques such as the cloning algorithm or diffusion Monte Carlo (DMC) [26–31] and transition path sampling (TPS) or path integral Monte Carlo [32]; or representing the non-equilibrium distribution with an explicit ansatz, for example, matrix product states [9, 11, 33]. The primary challenge in sampling methods is the problem of exponential variance in the estimator for the large deviation function as a function of the bias. Several approaches have been suggested to ameliorate this variance problem [10, 13, 20, 31, 34]. These techniques can be interpreted as different forms of importance sampling. In previous work, we showed that guiding distribution functions (GDF), as introduced for quantum diffusion Monte Carlo calculations [35], define an auxiliary dynamics that importance samples the

dynamics underlying the large deviation function under bias [20], and highlighted the connection to the generalized Doob’s transform [36–38]. In the current work, we describe a practical numerical technique to generate good guiding distribution functions in both lattice and continuum simulations of large deviation functions, using the idea of variance minimization. This again draws from the quantum field, and in particular the methods of variational Monte Carlo (VMC) [39]. There has been other recent work on optimizing auxiliary dynamics, for example in Ref. [40]; our work provides a different perspective, based on different techniques.

II. THEORY

In the current work, the quantity of interest in large deviation theory is the cumulant generating function (CGF) $\psi(\lambda)$ which is analogous to the free energy of equilibrium statistical mechanics. It is computed via an ensemble average over trajectories given by,

$$\begin{aligned}\psi(\lambda) &= \lim_{t_N \rightarrow \infty} \frac{1}{t_N} \ln \langle e^{\lambda \mathcal{O}[t_N]} \rangle \\ &= \lim_{t_N \rightarrow \infty} \frac{1}{t_N} \ln \sum_{\mathcal{C}(t_N)} P[\mathcal{C}(t_N)] e^{\lambda \mathcal{O}[t_N]}\end{aligned}\quad (1)$$

where λ is a field conjugate to the observable $\mathcal{O} = \sum_{t=1}^{t_N} o(\mathcal{C}_{t+}, \mathcal{C}_{t-})$, with o an arbitrary function of configurations (\mathcal{C}_t) at adjacent times, $t+$ and $t-$, and t_N is the final trajectory time. $P[\mathcal{C}(t_N)]$ is the likelihood of a given trajectory $\mathcal{C}(t_N) = \{\mathcal{C}_0, \mathcal{C}_1, \dots, \mathcal{C}_{t_N}\}$. The trajectories are generated by the master equation $\partial_t p_t(\mathcal{C}) = \mathcal{W} p_t(\mathcal{C})$, where $p_t(\mathcal{C})$ is the probability of a configuration of the system, \mathcal{C} , at time t , and \mathcal{W} is a linear operator. It can be shown that $\psi(\lambda)$ is the largest eigenvalue of a “tilted” operator \mathbb{W}_λ , i.e., $\langle \Xi | \mathbb{W}_\lambda = \langle \Xi | \psi(\lambda)$ where $\langle \Xi |$ is the corresponding dominant left eigenvector [41]. The effect of the tilt is to reweight the transition probabilities of \mathcal{W} .

* uray@caltech.edu

† garnetc@caltech.edu

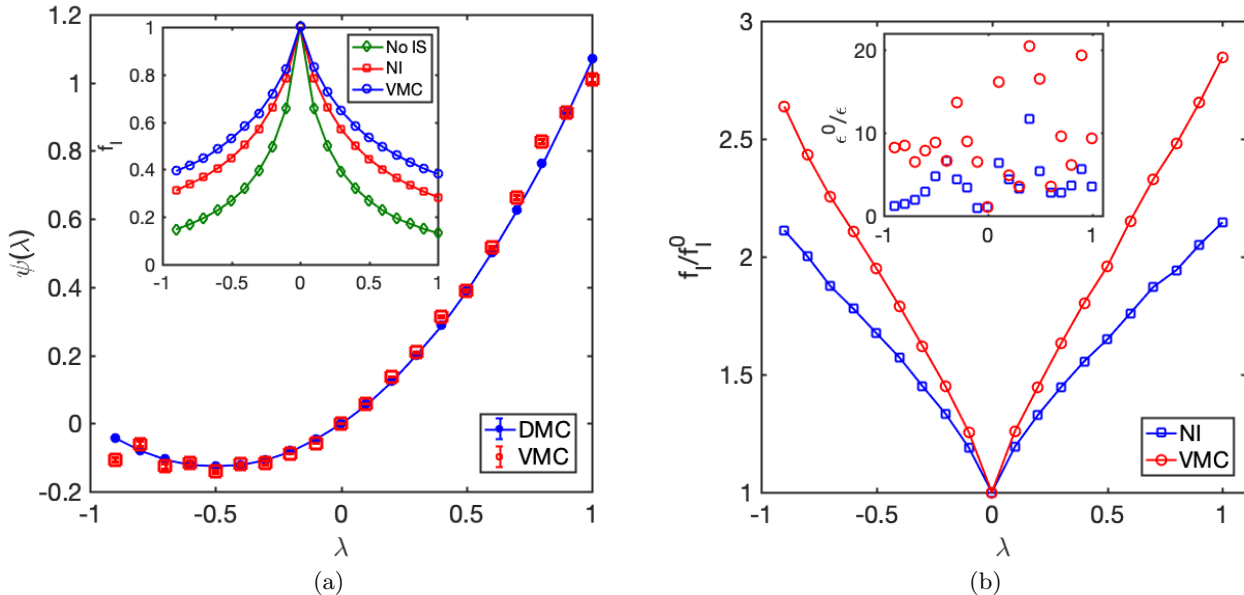


FIG. 1. (a) Large deviation function for the entropy production of $N = 10$ driven Brownian particles in a periodic potential with $v_0 = 2$, $f = 1.0$. The Gaussian interaction force is given by $r_c = 0.10$ and $\alpha = 10.0$. The main figure shows the CGF as a function of λ computed using DMC using the guiding function, and directly from the GDF via the VMC estimator (Eq. (5)). The DMC calculations are done with different GDF (no GDF, non-interacting (NI) GDF, variational form in main text) but all converge to the same estimate although the efficiencies are different. The inset shows these efficiencies measured via the fraction of independent walkers (f_I). (b) Improvement in the fraction of independent walkers (f_I) relative to calculations done without any auxiliary dynamics (f_I^0) for the same 1D Brownian problem. The inset shows the corresponding improvement in the standard deviation of the CGF ($\epsilon(\psi)$). Error bars are smaller than symbol sizes. No error estimate available for the standard deviation (see text).

In the discrete case, $\mathbb{W}_\lambda(\mathcal{C}, \mathcal{C}') = \mathcal{W}(\mathcal{C}, \mathcal{C}')e^{-\lambda\sigma(\mathcal{C}, \mathcal{C}')} (1 - \delta_{\mathcal{C}, \mathcal{C}'} - R(\mathcal{C})\delta_{\mathcal{C}, \mathcal{C}'})$, where $R(\mathcal{C}) = \sum_{\mathcal{C}' \neq \mathcal{C}} \mathcal{W}(\mathcal{C}, \mathcal{C}')$ is the exit rate.

\mathbb{W}_λ is not a Markovian operator (i.e. the sum of transition probabilities is not normalized). Consequently, when calculating $\psi(\lambda)$ via Monte Carlo techniques it is necessary to track this additional normalization constant or weight, whose variance grows exponentially with $|\lambda|$. It is desirable thus to instead consider an auxiliary dynamics generated by a modified operator $\tilde{W}_\lambda = \hat{\Xi}\mathbb{W}_\lambda\hat{\Xi}^{-1}$ (related to the generalized Doob's transformation [18, 36]) from which the CGF can be obtained as

$$\psi(\lambda) \sim \frac{1}{t_N} \ln \langle \mathbb{1} | \hat{\Xi}^{-1} e^{t_N \tilde{W}_\lambda} \hat{\Xi} | p_0 \rangle. \quad (2)$$

where the diagonal matrix $\hat{\Xi} = \sum_{\mathcal{C}} \hat{\Xi}(\mathcal{C}) |\mathcal{C}\rangle \langle \mathcal{C}|$ is constructed from the dominant left eigenvector of \mathbb{W}_λ (i.e., $\hat{\Xi}(\mathcal{C}) = \langle \Xi | \mathcal{C} \rangle$), $\langle \mathbb{1} |$ is the uniform left vector and $|p_0\rangle$ is the initial distribution of configurations. The normalization of \tilde{W}_λ is completely independent of configuration and the variance due to the bias is removed, thus the transformation by $\hat{\Xi}$ carries out a form of importance sampling. Although propagation with \tilde{W}_λ requires knowledge of the exact eigenvectors, in the GDF approach we simply approximate these eigenvectors with guiding functions of our own construction and carry out dynamics with \tilde{W}_λ using the diffusion Monte Carlo algorithm [20]. The quality of the GDF importance sampling

then depends on the degree of overlap between the approximate and the exact left eigenvector of \mathbb{W}_λ . The problem is therefore reduced to finding a high quality GDF in order to compute the CGF and its associated cumulants with good statistical efficiency.

As mentioned earlier, we can determine GDFs using ideas that originate from quantum diffusion Monte Carlo calculations, where the analogous problem is to determine a guiding function that best approximates the ground-state of a quantum Hamiltonian [39]. This is termed a variational Monte Carlo (VMC) calculation. While energy minimization is commonly used for this purpose [42, 43], \mathbb{W}_λ is not Hermitian and thus its spectrum is not necessarily bounded. However, we can use an associated property of eigenstates, *viz.* that the variance of the quantity,

$$\Lambda(\mathcal{C}) \equiv \hat{\Xi} \mathbb{W}_\lambda |\mathcal{C}\rangle \langle \mathcal{C}| \hat{\Xi}^{-1} \quad (3)$$

must vanish for eigenstates. This quantity, which we call the *local CGF*, is the analogue of the *local energy* for which variance minimization has previously been explored in quantum Monte Carlo [39]. The nonequilibrium variance minimization problem thus corresponds to minimizing

$$\sigma^2(\{p\}, \lambda) = \sum_{\mathcal{C}} w(\mathcal{C}) [\Lambda(\{p\}, \mathcal{C}) - \langle \Lambda(\{p\}) \rangle]^2, \quad (4)$$

where $\langle \Lambda(\{p\}) \rangle \equiv \sum_{\mathcal{C}} w(\mathcal{C}) \Lambda(\{p\}, \mathcal{C})$ is the average of the local CGF, $\{p\}$ are the parameters used to characterize

the GDF: $\tilde{\Xi}(\{p\}, \mathcal{C})$, and $w(\mathcal{C})$ is a normalized sampling distribution. The calculation involves minimizing Eq. (4) with respect to $\{p\}$ over a fixed set of configurations $\{\mathcal{C}\}$. Note that (4) can be sampled from any $w(\mathcal{C})$ but for $w(\mathcal{C}) = \frac{\tilde{\Xi}(\{p\}, \mathcal{C})}{\|\tilde{\Xi}(\{p\})\|}$, we obtain

$$\psi(\lambda) \approx \langle \Lambda \rangle = \frac{\langle \tilde{\Xi}(\{p\}) | \mathbb{W}_\lambda | \mathbb{1} \rangle}{\langle \tilde{\Xi}(\{p\}) | \mathbb{1} \rangle} \quad (5)$$

which is an estimator for $\psi(\lambda)$ in the sense that the approximate sign is replaced by an equality for a GDF that is the exact eigenstate. Despite the absence of a bounded variational principle, this estimator of $\psi(\lambda)$ in practical terms is often useful, and we refer to this as the VMC estimator for $\psi(\lambda)$, in complete correspondence to its quantum counterpart.

Whereas a strictly zero variance is guaranteed for eigenstates, we emphasize that a smaller (non-zero) variance of the local CGF does not strictly imply a better GDF. A more rigorous metric is the actual reduction of the standard deviation in the estimate of the CGF ($\varepsilon(\psi)$) from DMC (or indeed TPS, which can easily be adapted to use a GDF). Decreasing $\varepsilon(\psi)$ thus defines a better GDF, and this is reflected in statistical independence of samples (the trajectories which are being generated). For the DMC algorithm the indicator of statistical independence is the fraction of independent walkers f_I [20, 31]. Empirically f_I is seen to be less susceptible to statistical noise than $\varepsilon(\psi)$ but (for a given λ) still bears a monotonic relationship with it, and we will therefore primarily use f_I as the measure of quality of a GDF. In the case of perfect importance sampling, using the exact auxiliary dynamics, f_I is equal to 1 at all times. In the other limit, if all walkers are correlated, $f_I = 0$. Because $f_I(t)$ measures the correlation among walkers as a function of time, it must be smallest at $t = 0$ [14, 20], which is what we will report. It is important to note that although improvements in f_I yield improvements in $\varepsilon(\psi)$ the relationship between the two is not linear.

III. RESULTS

To demonstrate our procedure, we carry out simulations on a continuum system and a lattice model. For the continuum system, we consider the prototypical driven Brownian walker, where the observable of interest is the entropy production. This system consists of $N = 10$ particles (at location $\mathbf{R} = \{r_i\}$) moving in a periodic potential ($V(r) = v_0 \cos(2\pi r)$) on a ring of size $L = 1.0$ under the influence of an external driving force (f) and a fluctuating field represented by Gaussian white noise. These particles also interact via a pairwise repulsive force ($\hat{\mathbf{f}}(r_i, r_j) = -\hat{\mathbf{f}}(r_j, r_i)$ and $|\hat{\mathbf{f}}| = \alpha \exp[-(|r_i - r_j|/r_c)^2]$). The transformed tilted propagator that includes auxil-

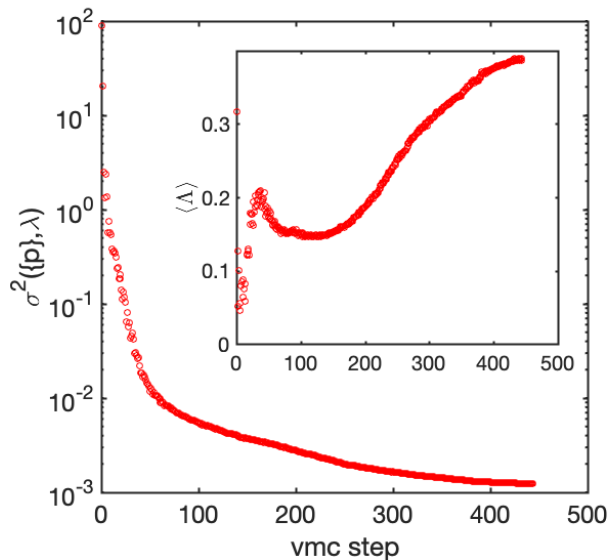


FIG. 2. VMC minimization trace of the variance for the driven Brownian walker system for $\lambda = 0.5$. Inset shows how the local CGF changes as the minimization proceeds (see text).

ary dynamics due to the GDF $\tilde{\Xi}(\mathbf{R})$ is given by [20],

$$\begin{aligned} \tilde{W}_\lambda = & \sum_i \partial_i (\partial_i - [\mathbf{F}_i(\mathbf{R}) \cdot \hat{\mathbf{x}} + 2f\lambda + 2 \ln \tilde{\Xi}(\mathbf{R})]) \\ & + \tilde{\Xi}^{-1}(\mathbf{R}) \mathbb{W}_\lambda^\dagger \tilde{\Xi}(\mathbf{R}), \end{aligned} \quad (6)$$

where $\mathbf{F}_i(\mathbf{R}) = f\hat{\mathbf{x}} - \partial_i V(r_i)\hat{\mathbf{x}} + \sum_{j \neq i} \hat{\mathbf{f}}(r_i, r_j)$. The adjoint operator $\mathbb{W}_\lambda^\dagger = \sum_i \partial_i^2 + (\mathbf{F}_i(\mathbf{R}) \cdot \hat{\mathbf{x}} + 2f\lambda)\partial_i + f\lambda(f\lambda + \mathbf{F}_i(\mathbf{R}) \cdot \hat{\mathbf{x}})$, represents the norm breaking term that is handled via branching. Trajectories for \tilde{W}_λ are generated from Langevin dynamics $\partial_t r_i = \mathbf{F}_i(\mathbf{R}) \cdot \hat{\mathbf{x}} + 2f\lambda + 2\partial_i \ln \tilde{\Xi}(\mathbf{R}) + \eta_i$, where the random force, η_i , satisfies $\langle \eta_i(t) \rangle = 0$ and $\langle \eta_i(t)\eta_j(t') \rangle = 2\delta(t-t')$. The entropy production $s(t)$ is reflected in the biasing term $\exp[\lambda t_N s(t)]$ with $s(t_N)t_N = \sum_i \int_0^{t_N} f \dot{r}_i(\tau) d\tau$ which is absorbed into the dynamics of the tilted Fokker-Planck operator [14, 20, 44].

For this system we parametrize the GDF as $\tilde{\Xi}(\mathbf{R}) = \prod_i \phi(r_i) \prod_{j < i} [J(r_i, r_j)]$, where the single particle function, $\phi(r_i)$, is obtained from the non-interacting eigenstate of \tilde{W}_λ ($\hat{\mathbf{f}}(r_i, r_j) = 0$) generated with $M = 101$ plane wave modes [20] and $J(r_i, r_j) = \sum_{k_1, k_2} \tilde{J}(k_1, k_2) e^{ik_1 r_1 + ik_2 r_2}$. The parameters $\tilde{J}(k_1, k_2)$ are estimated by minimizing (4).

Shown in Fig. 1a is the large deviation function and f_I computed using no guiding function, the non-interacting guiding function ($J = 1$), and the variational form above (a similar form has been explored independently in [40]). We parametrized $J(k_1, k_2)$ with 241 parameters (21 plane waves per particle) and we minimized the variance such that $\sigma^2(\{p\}, \lambda) < 4.0 \times 10^{-3}$ for all λ s we considered. The minimization procedure is started at small $|\lambda|$ using an

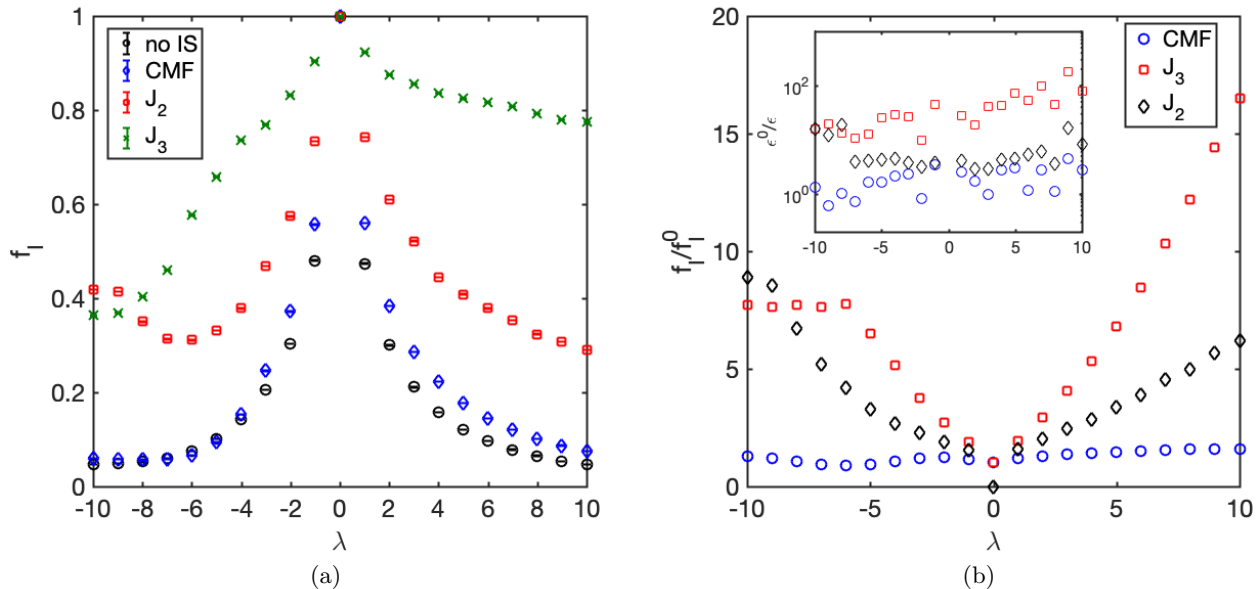


FIG. 3. (a) Fraction of independent walkers (f_I) for different types of GDF (no GDF, non-interacting GDF, optimized variational form (see main text)) for a $L = 16$ WASEP model. The most flexible GDF is provided by the J_3 ansatz. We find that the efficiency gains from using the GDF systematically improve from the CMF ansatz to the J_2 and J_3 ansatz. (b) Improvement in the fraction of independent walkers (f_I) relative to calculations done without any auxiliary dynamics (f_I^0). The inset shows the corresponding improvement in the standard deviation of the CGF ($\epsilon(\psi)$). Error bars are smaller than symbol sizes. No error estimate available for the standard deviation (see text).

initial guess for parameters that produces a uniform state since we know that for $\lambda = 0$, the exact $\langle \Xi \rangle$ is uniform. The optimized parameters are used as an initial guess for the next (nearby) λ . A trace of the minimization is shown in Fig. 2 at $\lambda = -0.5$ obtained using a simple simplex algorithm. For this system the minimization required 10000 configurations (sampled from the guess GDF as the distribution $w(\mathcal{C})$ for $-0.3 \leq \lambda \leq 0.3$ ($\lambda \neq 0$) in order to avoid getting stuck in local minima (which produced a poor GDF). For $|\lambda| > 0.3$, 2000 configurations were sufficient.

From the reduction in f_I (inset of Fig. 1a) it is evident that continuum calculations can be made much more efficient with an appropriate GDFs. As noted earlier, the improvement in f_I estimated at $t = 0$ can imply greater efficiency gains when the full trajectory space is considered. In Fig. 1b (inset) we show the improvement in the corresponding standard deviation in the subsequent DMC calculation, which can be reduced by a large factor, although statistical noise in this measure means that it is difficult to give a precise estimate of the factor. Additionally, although we used the same observation time (t_N) to compute the CGF and cumulants for all types of sampling, we find that the results converge much faster with t_N for simulations done with auxiliary dynamics.

We now consider an interacting non-equilibrium problem on a lattice, namely the current fluctuations of a periodic weakly asymmetric simple exclusion process (WASEP) [45]. The WASEP models transport of N particles on a lattice with L sites. Here N is chosen as $0.3L$. The configuration of the particles is defined by

a set of occupation numbers with hard-core constraints, $n_i = \{0, 1\}$, e.g. $\mathcal{C} = \{0, 1, \dots, 1, 1\}$. The tilted propagator for a current bias λ ,

$$\mathbb{W}_\lambda = \sum_{i=1}^L p e^{\lambda/L} \hat{b}_{i+1}^\dagger \hat{b}_i - p \hat{n}_i (\hat{\mathbb{1}} - \hat{n}_{i+1}) + q e^{-\lambda/L} \hat{b}_{i-1}^\dagger \hat{b}_i - q \hat{n}_i (\hat{\mathbb{1}} - \hat{n}_{i-1}) \quad (7)$$

yields particles hopping to the right with rate $p = \frac{1}{2} e^{E/L}$ and to the left with rate $q = \frac{1}{2} e^{-E/L}$, where the $1/L$ factor in hopping gives the weakly asymmetric limit, whose large scale behavior (universality class) is described by the Edwards-Wilkinson equation [4, 46]. Here \hat{b}_i^\dagger (\hat{b}_i) creates (destroys) particles on site i and n_i ($\mathbb{1} - n_i$) counts the number of particles (holes). For subsequent calculations we set $E = 10$, and use periodic boundary conditions, $i = L + 1 \rightarrow 1$ and $i = 0 \rightarrow L$.

Unlike the continuum system, where the soft-core interaction means the non-interacting solution is a sensible starting point to construct the GDF, the hard-core interaction requires a different treatment. Here we consider a GDF that is a product purely of n -particle correlation factors, e.g. for $n = 2$, $\Xi(\mathcal{C}) = \prod_{p < q} J_2(r_p, r_q)$ where r_p, r_q denote the positions of the particles in the configuration \mathcal{C} , for $n = 3$ we use $J_3(r_p, r_q, r_s)$ etc., and the variational parameters are the values (i.e. $J_2(r_p, r_q)$). This form is sometimes referred to as a correlator product state in quantum systems [47, 48]. To enforce PBC we use $J_2(r_{pq}), J_3(r_{pq}, r_{qr}, r_{qs})$ where the inter-particle distances are defined with a minimum image convention, i.e. $r_{pq} = \min(|r_p - r_q|, L - |r_p - r_q|)$, and J_2, J_3 are symmetric under cyclic permutations of their arguments. (Cyclic permutation symmetry, rather than full symmetry, was

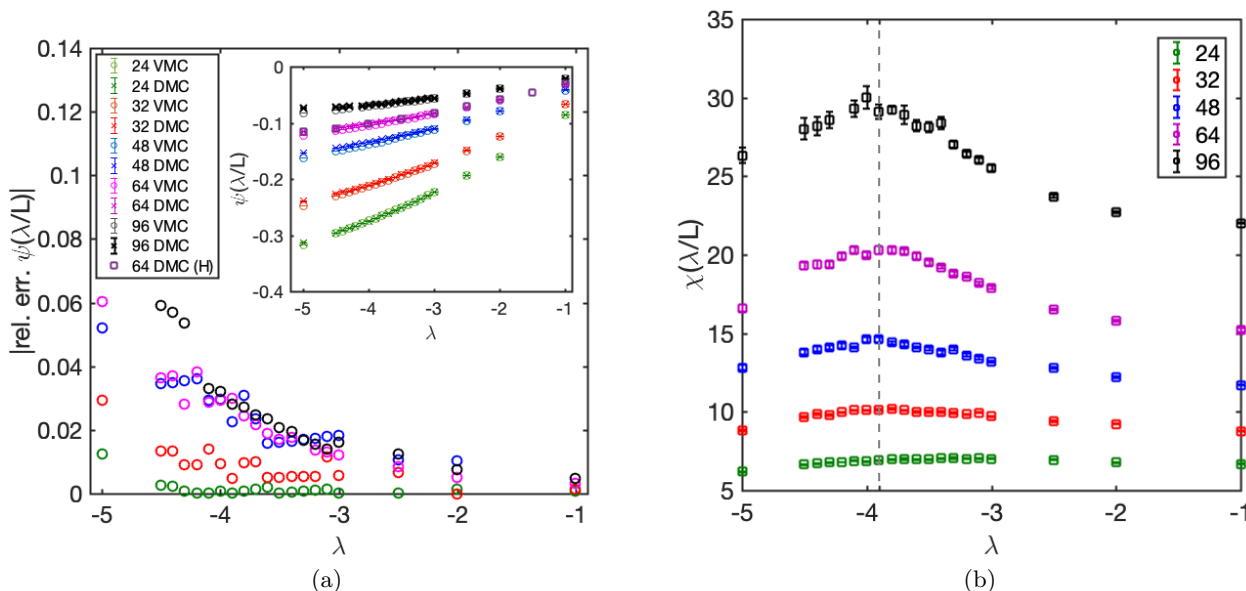


FIG. 4. (a) Relative error of CGF obtained from VMC as compared to DMC. Note that as the system size is increased and more particles are involved the GDF is unable to capture the full extent of the correlations. Nonetheless the DMC calculations still benefit from the GDF in terms of statistical efficiency. Inset shows comparison of DMC and VMC results for the CGF ($\psi(\lambda/L)$). Also included is data from [49] for $L = 64$ indicated as DMC (H). (b) Susceptibility ($\chi(\lambda/L)$) of the 1D WASEP for different system sizes. The dashed line at $\lambda = 3.9$ indicates the possible location of the continuous phase transition in the limit of $L \rightarrow \infty$.

used to reflect the handedness of hopping around in the model). To minimize the number of variational parameters in the large calculations below, for $r_{pq} > R_{\text{cut}}$ we set $J_3 = 1$ where R_{cut} is a cutoff distance. For comparison we have also considered a GDF of the cluster mean-field (CMF) form described in [20].

Fig. 3a shows a comparison of f_I as a function of λ for a $L = 16$ model with different GDFs. In the WASEP, the short-range CMF is unable to capture the long-range correlations present in the system and therefore we do not get much improvement in efficiency using this GDF. However, with the correlator product state we can obtain large improvements (e.g., we see from Fig. 3b using J_3 , $\varepsilon(\psi)$ is improved by an order to two orders of magnitude).

In order to illustrate the flexibility of this ansatz we have further performed calculations for different system sizes $L = 24-96$. Due to the reduction in standard deviation, despite the large system size we needed only a modest number of walkers in the DMC procedure $N_w = 10000$ to 120000, which was sufficient to estimate the susceptibility $\chi(\lambda/L) = \frac{d^2 \psi(\lambda/L)}{d(\lambda/L)^2}$ (computed as a correlation function). For these calculations, in order to reduce the number of parameters, we used a cutoff distance of $R_{\text{cut}} = 16$ for J_3 . The variance minimization was carried out using between 4000 to 8000 fixed configurations sampled from the GDF to optimize 85 to 155 parameters depending on the system size. The error of the VMC estimator for $\psi(\lambda)$, as a measure of the GDF quality, is shown in Fig. 4a where we see that the relative error grows with $|\lambda|$ as more particles become correlated. Note that for these calculations we did not spend a lot

of effort to optimize the more extreme λ values as we were interested in the range over which the susceptibility peaks, where the VMC error is $< 5\%$. In this model, the trajectories undergo a continuous phase transition from a uniform state to a traveling wave at some λ_c provided $E > E_c = \pi/\sqrt{\rho(1-\rho)}$ where $\rho = N/L$ [3, 49]. This can be detected from the growing susceptibility of the system shown in Fig. 4b. Macroscopic Fluctuation Theory (MFT) suggests that $\lambda_c = -E + \sqrt{E^2 - E_c^2} = -2.7198$ (for $\rho = 0.3$) [49], although this is not an exact result and the precise value of λ_c for $L \rightarrow \infty$ is not explicitly known [4]. Our DMC calculations suggest a critical point near $\lambda_c \sim 3.9$ for the largest system sizes ($L = 64, 96$). We note our results for $\psi(\lambda/L)$ are in excellent agreement with the largest system size ($L = 64$) considered in [49], also plotted in the inset of Fig. 4a.

IV. CONCLUSIONS

In this article we showed how to compute guiding distribution functions using the technique of variance minimization originating in variational Monte Carlo calculations of quantum systems. This provides a systematic route to statistically efficient Monte Carlo computation of large deviation functions using the diffusion Monte Carlo or cloning algorithm, which we demonstrated in the continuum for the problem of the Brownian walker, as well as on the lattice for the WASEP model. The very general nature of variance minimization means that the possibilities for different guiding distribution func-

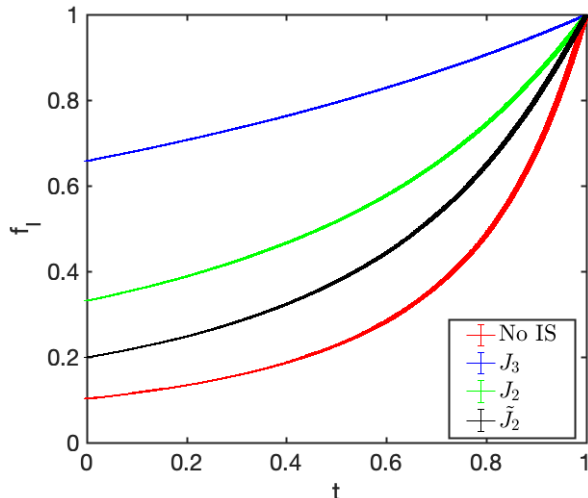


FIG. 5. Fraction of independent walkers (f_I) as a function of normalized observation time for different types of auxiliary dynamics for $\lambda = -5.0$. (See text)

tions are limited only by one's imagination, and they can be adapted to very complex systems. Finally, we note that obtaining a good form for the guiding distribution function, much like obtaining a compact wavefunction in a quantum problem, is likely to provide important analytic insights into the behaviour of the non-equilibrium system of interest.

ACKNOWLEDGMENTS

The authors would like to thank Rob Jack, Vivien Lecomte, Juan P. Garrahan and David Limmer for fruitful and engaging discussions. 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 performed with CANSS, available at <https://github.com/ushnishray/CANSS>.

Appendix A: Fokker-Planck operator for the continuum

In the main text we described how the computation of the cumulant generating large deviation function (CGF) can be greatly enhanced with the appropriate form of auxiliary dynamics. The GDF structure we have introduced in the 1D case is of the form $\tilde{\Xi}(\mathbf{R}) = \prod_i \phi(r_i) \prod_{j < i} [J(r_i, r_j)]$. To carry out simulations with this GDF, we require the adjoint of the transformed Fokker-Planck operator. For the Brownian walker

GDF	$\psi(\lambda)$	$\varepsilon(\lambda)$
No IS	-0.5012	1.3e-04
\tilde{J}_2	-0.50113	9.2e-05
J_2	-0.50106	2.9e-05
J_3	-0.501049	4.9e-06

TABLE I. CGF estimate and standard deviation for different auxiliary dynamics at $\lambda = -0.5$ (see text).

in the main text, we find

$$\frac{\mathbb{W}_\lambda^\dagger \tilde{\Xi}(\mathbf{R})}{\tilde{\Xi}(\mathbf{R})} = \sum_{i=1}^N \left\{ \frac{P_1^\dagger \phi(r_i)}{\phi(r_i)} + \sum_{j \neq i} \frac{\partial^2 J(r_i, r_j)}{\partial r_i^2} + \left[\gamma_i(r_i) + 2\lambda + \frac{2}{\phi(r_i)} \frac{d\phi(r_i)}{dr_i} \right] \left[\sum_{j \neq i} \frac{\partial J(r_i, r_j)}{\partial r_i} \right] + \left[\sum_{j \neq i} \hat{\mathbf{f}}(r_i, r_j) \cdot \hat{\mathbf{x}} \right] \left[\frac{1}{\phi(r_i)} \frac{d\phi(r_i)}{dr_i} + \sum_{j \neq i} \frac{\partial J(r_i, r_j)}{\partial r_i} \right] \right\}. \quad (\text{A1})$$

where the non-interacting single-particle operator is given by,

$$\frac{P_1^\dagger \phi(r_i)}{\phi(r_i)} = \frac{1}{\phi(r_i)} \frac{d\phi(r_i)}{dr_i} + \frac{\gamma_i(r_i) + 2\lambda}{\phi(r_i)} \frac{d\phi(r_i)}{dr_i} + \lambda(\lambda + \gamma_i(r_i)) \quad (\text{A2})$$

where $\gamma_i(r_i) = f - \partial_i V(r_i)$ is the effective single-particle force acting on a particle. This expression provides the norm-breaking term for a subsequent DMC calculation and can be generalized straightforwardly to higher dimensions.

Appendix B: Statistical Independence and Efficiency Improvement with Auxiliary Dynamics

In the main text we have mentioned that an improvement of the GDF is indicated by the improvement in the fraction of independent walkers (f_I) estimated from DMC (this would be reflected in a corresponding decrease in the autocorrelation time for TPS) and we reported f_I evaluated for $t = 0$. Here we show the full $f_I(t)$ in Fig. 5 for calculations using different types of dynamics at $\lambda = -5.0$ for the 1D WASEP system. Note that $f_I(1) = 1$ but as one goes back in simulation time, the walkers are descended from a smaller and smaller set of ancestors. The GDFs J_2 and \tilde{J}_2 use the same J_2 ansatz in the main text but illustrate the effect of optimizing for more steps in the variance minimization of J_2 , leading to an improved $f_I(t)$. This is also reflected in Table I where the VMC estimator of the CGF appears to converge with increasingly flexibility of the GDF, while the standard deviation is systematically improved.

-
- [1] H. Touchette, *Physics Reports* **478**, 1 (2009).
- [2] B. Derrida, J. L. Lebowitz, and E. R. Speer, *Journal of Statistical Physics* **110**, 775 (2003).
- [3] T. Bodineau and B. Derrida, *Physical Review E* **72**, 066110 (2005).
- [4] S. Prohac and K. Mallick, *Journal of Physics A: Mathematical and Theoretical* **42**, 175001 (2009).
- [5] J. de Gier and F. H. L. Essler, *Phys. Rev. Lett.* **107**, 010602 (2011).
- [6] P. I. Hurtado and P. L. Garrido, *Physical review letters* **107**, 180601 (2011).
- [7] M. Gorissen, A. Lazarescu, K. Mallick, and C. Vanderzande, *Phys. Rev. Lett.* **109**, 170601 (2012).
- [8] A. Lazarescu, *Journal of Physics A: Mathematical and Theoretical* **48**, 503001 (2015).
- [9] P. Helms, U. Ray, and G. K.-L. Chan, *Phys. Rev. E* **100**, 022101 (2019).
- [10] T. Nemoto, R. L. Jack, and V. Lecomte, *Physical Review Letters* **118**, 115702 (2017).
- [11] M. C. Bañuls and J. P. Garrahan, arXiv e-prints, arXiv:1903.01570 (2019), arXiv:1903.01570 [cond-mat.stat-mech].
- [12] S. Whitelam, L. O. Hedges, and J. D. Schmit, *Physical review letters* **112**, 155504 (2014).
- [13] K. Klymko, P. L. Geissler, J. P. Garrahan, and S. Whitelam, *Phys. Rev. E* **97**, 032123 (2018).
- [14] U. Ray, G. K.-L. Chan, and D. T. Limmer, *The Journal of Chemical Physics* **148**, 124120 (2018).
- [15] A. Prados, A. Lasanta, and P. I. Hurtado, *Phys. Rev. Lett.* **107**, 140601 (2011).
- [16] A. Prados, A. Lasanta, and P. I. Hurtado, *Physical Review E* **86**, 355 (2012).
- [17] J. Mehl, T. Speck, and U. Seifert, *Physical Review E* **78**, 011123 (2008).
- [18] R. Chetrite and H. Touchette, *Journal of Statistical Mechanics: Theory and Experiment* **2015**, P12001 (2015).
- [19] P. Tsohngni Nyawo and H. Touchette, *Phys. Rev. E* **94**, 032101 (2016).
- [20] U. Ray, G. K.-L. Chan, and D. T. Limmer, *Physical Review Letters* **120**, 210602 (2018).
- [21] T. GrandPre and D. T. Limmer, *Phys. Rev. E* **98**, 060601 (2018).
- [22] F. Carollo, J. P. Garrahan, I. Lesanovsky, and C. Pérez-Espigares, *Phys. Rev. A* **98**, 010103 (2018).
- [23] F. Carollo, J. P. Garrahan, and I. Lesanovsky, *Phys. Rev. B* **98**, 094301 (2018).
- [24] A. J. Schile and D. T. Limmer, *The Journal of Chemical Physics* **149**, 214109 (2018), <https://doi.org/10.1063/1.5058281>.
- [25] C. Y. Gao and D. T. Limmer, arXiv e-prints, arXiv:1812.01470 (2018), arXiv:1812.01470 [cond-mat.stat-mech].
- [26] P. Grassberger, *Computer Physics Communications* **147**, 64 (2002).
- [27] P. Del Moral and J. Garnier, *Ann. Appl. Probab.* **15**, 2496 (2005).
- [28] C. Giardinà, J. Kurchan, and L. Peliti, *Phys. Rev. Lett.* **96**, 120603 (2006).
- [29] C. Giardinà, J. Kurchan, V. Lecomte, and J. Tailleur, *Journal of statistical physics* **145**, 787 (2011).
- [30] F. Crou, A. Guyader, T. Lelivre, and D. Pommier, *The Journal of chemical physics*, **134**, 054108 (2011).
- [31] T. Nemoto, F. Bouchet, R. L. Jack, and V. Lecomte, *Phys. Rev. E* **93**, 062123 (2016).
- [32] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, *Annual review of physical chemistry* **53**, 291 (2002).
- [33] M. Gorissen, J. Hooyberghs, and C. Vanderzande, *Physical Review E* **79**, 020101 (2009).
- [34] D. Jacobson and S. Whitelam, arXiv e-prints, arXiv:1903.06098 (2019), arXiv:1903.06098 [cond-mat.stat-mech].
- [35] D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).
- [36] J. L. Doob, *Classical Potential Theory and Its Probabilistic Counterpart* (Springer-Verlag, 1984).
- [37] R. Chetrite and H. Touchette, in *Annales Henri Poincaré*, Vol. 16 (Springer, 2015) pp. 2005–2057.
- [38] R. L. Jack and P. Sollich, *Progress of Theoretical Physics Supplement* **184**, 304 (2010).
- [39] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, *Phys. Rev. Lett.* **60**, 1719 (1988).
- [40] A. Das and D. T. Limmer, arXiv e-prints, arXiv:1909.03589 (2019), arXiv:1909.03589 [cond-mat.stat-mech].
- [41] J. L. Lebowitz and H. Spohn, *Journal of Statistical Physics* **95**, 333 (1999).
- [42] C. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, *Physical review letters* **98**, 110201 (2007).
- [43] S. Sorella, *Physical Review B* **71**, 241103 (2005).
- [44] U. Seifert, *Reports on Progress in Physics* **75**, 126001 (2012).
- [45] B. Schmittmann and R. K. Zia, *Phase transitions and critical phenomena* **17**, 3 (1995).
- [46] S. F. Edwards and D. R. Wilkinson, *Proceedings of the Royal Society of London Series A* **381**, 17 (1982).
- [47] M. P. Nightingale and H. W. J. Blöte, *Phys. Rev. B* **33**, 659 (1986).
- [48] H. J. Changlani, J. M. Kinder, C. J. Umrigar, and G. K.-L. Chan, *Physical Review B* **80**, 245116 (2009).
- [49] C. P. Espigares, P. L. Garrido, and P. I. Hurtado, *Physical Review E* **87**, 032115 (2013).