Finite time large deviations via matrix product states

Mari Carmen Bañuls, Luke Causer, and Juan P. Garrahan

Recent work has shown the effectiveness of tensor network methods for computing large deviation functions in constrained stochastic models in the infinite time limit. Here we show that these methods can also be used to study the statistics of dynamical observables at *arbitrary finite time*. This is a harder problem because, in contrast to the infinite time case where only the extremal eigenstate of a tilted Markov generator is relevant, for finite time the whole spectrum plays a role. We show that finite time dynamical partition sums can be computed efficiently and accurately in one dimension using matrix product states, and describe how to use such results to generate rare event trajectories on demand. We apply our methods to the Fredrickson-Andersen (FA) and East kinetically constrained models, and to the symmetric simple exclusion process (SSEP), unveiling dynamical phase diagrams in terms of counting field and trajectory time. We also discuss extensions of this method to higher dimensions.

Introduction. – Large deviation (LD) theory provides a powerful framework to investigate the statistical fluctuations of time-averaged observables in stochastic systems (for reviews, see e.g. Refs. [**?** **?** **?** **?**]). At long times (assuming finite correlation times) the probabilities of such observables obey a LD principle, and the corresponding scaled cumulant generating function (SCGF, see below) can be retrieved from the leading eigenvalue of the *tilted* (or deformed or biased) generator [**?**]. For large systems, estimating this eigenvalue is difficult, so one resorts to sampling the corresponding biased trajectory ensemble via numerical methods such as trajectory importance sampling [**?** **?** **?** **?**], population dynamics [**?** **?** **?**], optimal control [**?** **?** **?** **?** **?** **?** **?**], or machine learning approaches [**?** **?** **?** **?** **?** **?**]. For lattice models, recent work has focused on the use of tensor network (TN) techniques to approximate the leading eigenvector of the titled generator through variational means [**?** **?** **?**] or power methods [**?**].

A harder problem is that of computing the statistics of time-averaged observables for *finite time*. The reason is that away from the long time limit the corresponding dynamical partition sums (i.e., moment generating functions) do not obey a LD principle in time - only obeying an LD principle in space for large sizes - and as a consequence they are not determined only by the leading eigenvalue of the tilted generator, but by their whole spectrum. If time is very short, one can get away with direct sampling, but for intermediate times the usual sampling approaches fall short [**?**]. Here we develop a scheme to study these rare events by implementing well-developed TN techniques to simulate time evolution. This allows us to calculate dynamical partition functions for trajectories of arbitrary time extent. Furthermore, we show how to use the results here to directly simulate stochastic trajectories in finite-time tilted ensembles at small computational cost, thus generalising the method of Ref. [**?**].

We focus for concreteness on one dimensional kinetically constrained models (KCMs) - often used in the modelling of structural glasses [**?** **?** **?** **?**] - specifically the Fredrickson-Andersen (FA) [**?**] and the East [**?**] models, and on the symmetric simple exclusion process (SSEP). Both KCMs and SEPs display phase transitions in their dynamical LDs in the long-time limit [**?** **?** **?** **?** **?** **?** **?**]. With the methods developed here we are able to construct the dynamical phase diagram both as a function of counting field and of trajectory time, determining finite time scaling of active-inactive phase transitions in these models, and uncovering the emergence with time of the correlated structure of the active phase in the East model and the SSEP.

Models. – The three models we consider live in a one dimensional lattice of N sites, with binary variables $n_j = 0, 1$ for each $j = 1 \cdots N$, evolving under continuous-time Markov dynamics with local transitions. The probability for each configuration $|x\rangle = |n_1 \cdots n_N\rangle$ at time t, encoded in a vector $|P(t)\rangle = \sum_x P(t, x) |x\rangle$, evolves deterministically via a Master equation, $\partial_t |P(t)\rangle = \mathbb{W} |P(t)\rangle$, where $\mathbb{W}$ is the Markov generator. Being a stochastic operator $\mathbb{W}$ has a structure $\mathbb{W} = \mathbb{K} - \mathbb{R}$, with an off-diagonal matrix of transition rates $\mathbb{K}$, and a positive diagonal matrix of escape rates $\mathbb{R}$.

For the KCMs the generator reads

$$\mathbb{W}^{\text{KCM}} = \sum_i f_i \big[c\sigma_i^+ + (1-c)\sigma_i^- $$
$$ - c(1 - n_i) - (1 - c)n_i \big], \quad (1)$$

where $c \in (0, 1/2]$ defines the site occupation at equilibrium, and $\sigma_i^{\pm}$ are the Pauli raising and lowering operators at site i. Spin flips are only permitted if the kinetic constraint, f_i, is satisfied. We consider two paradigmatic KCMs, the Fredrickson-Andersen (FA) [**?**] model and the East [**?**] model. They are defined by the respective constraint functions

$$ f_i^{\text{FA}} = n_{i-1} + n_{i+1}, \quad f_i^{\text{East}} = n_{i-1}. \quad (2)$$

We consider lattices with open boundary conditions (OBC) to allow for efficient tensor network contractions. For numerical convenience, we choose the fixed boundaries $n_1 = n_N = 1$ for the FA [**?**] model and $n_1 = 1$ for

the East model. The corresponding stationary states are product states,

$$|ss^{FA}\rangle = |1\rangle \otimes [(1-c)|0\rangle + c|1\rangle]^{\otimes N-2} \otimes |1\rangle, \quad (3)$$

$$|ss^{East}\rangle = |1\rangle \otimes [(1-c)|0\rangle + c|1\rangle]^{\otimes N-1}. \quad (4)$$

The third model we consider is the symmetric simple exclusion process (SSEP) whose generator reads

$$\mathbb{W}^{SSEP} = \frac{1}{2} \sum_i \left[\sigma_i^+ \sigma_{i+1}^- + \sigma_i^- \sigma_{i+1}^+ \right. \quad (5)$$
$$\left. - (n_i + n_{i+1}) + 2n_i n_{i+1} \right]$$

For the SSEP we consider OBC such that particles can come in and out at the boundaries with rate 1/4. The stationary state is $|ss^{SSEP}\rangle = 2^{-N} |-\rangle = 2^{-N} \sum_x |x\rangle$, with the "flat" state $\langle -|$ being the leading left eigenvector of each generator above.

Dynamical rare events and LDs. – We now consider the ensemble of all possible trajectories $\{\omega_\alpha\}$ with trajectory time t, where $\omega_\alpha = \{x_0 \to x_{t_1} \to \cdots \to x_t\}$ defines jumps to configurations x_{t_k} at times t_k. The probability of observing the value $K(\omega_\alpha) = K$ of some time-integrated observable K is

$$P_t(K) = \sum_\alpha \pi(\omega_\alpha)\delta(K(\omega_\alpha) - K), \quad (6)$$

where $\pi(\omega_\alpha)$ defines the probability of observing the trajectory. The corresponding moment generating function (or trajectory partition sum) is

$$Z_t(s) = \sum_\alpha \pi(\omega_\alpha)e^{-sK(\omega_\alpha)} \quad (7)$$

where the *counting field* s is conjugate to the observable.

For large times, both Eqs. (6-7) take a large deviation (LD) form in time [**? ? ? ?**], $P_t(K) \asymp e^{-t\varphi(K/t)}$ and $Z_t(s) \asymp e^{t\theta(s)}$. The LD rate function $\varphi(K/t)$ and the scaled cumulant generating function (SCGF) $\theta(s)$ play the roles of a trajectory entropy density and a free-energy density, respectively, and are related through the Legendre transform $\varphi(k) = -\min_k [\theta(s) + sk]$.

The partition sum Eq. (7) can be written as

$$Z_t(s) = \langle -|e^{t\mathbb{W}_s}|ss\rangle \quad (8)$$

in terms of the *tilted generator* $\mathbb{W}_s$ [**? ? ? ?**]. In what follows we focus on the *dynamical activity* [**? ? ? ?**]), that is, the total number of spin flips, as a trajectory observable. In this case, $\mathbb{W}_s = e^{-s}\mathbb{K} - \mathbb{R}$. While for large times all that is needed to determine Eq. (8) is the dominant eigenstate of $\mathbb{W}_s$, for finite times the whole spectrum of $\mathbb{W}_s$ is required.

Finite time statistics from MPS. – The models we consider obey detailed balance. This allows us to write

FIG. 1. **Demonstration of the methods.** East model at $c = 0.5$, $N = 100$ and $s = 0.1$. (a) Dynamical activity $\langle k \rangle$ from tMPS (black line), TPS with no auxiliary dynamics (red circles), TPS with the LD eigenvector auxiliary dynamics via vMPS (blue squares), and TPS with a tMPS reference dynamics (green pentagons). (b) Time-dependent occupations (top) and instantaneous activity (bottom) from MPS time-evolution (black line) from direct sampling with a tMPS auxiliary dynamics (green pentagons / bars).

$\mathbb{W}_s$ in a Hermitian form through a similarity transformation independent of s [**?**], $\mathbb{H}_s = \mathbb{P}^{-1/2} \mathbb{W}_s \mathbb{P}^{1/2}$, where $\mathbb{P}^{1/2}$ is a diagonal matrix of probability amplitudes at equilibrium (for the SSEP, $\mathbb{P}$ is the identity). As a consequence the leading eigenvalue of $\mathbb{H}$ obeys a Rayleigh-Ritz variational principle, allowing the application of variational methods such as the density matrix renormalisation group (DMRG) [**?**]. We then write Eq. (8) as

$$Z_t(s) = \langle \psi_0|e^{t\mathbb{H}_s}|\psi_0\rangle, \quad (9)$$

where $|\psi_0\rangle = \mathbb{P}^{-1/2} |ss\rangle = [\langle -|\mathbb{P}^{1/2}]^\dagger$. It is useful to define the time evolved vector $|\psi_\tau\rangle = e^{\tau\mathbb{H}_s} |\psi_0\rangle$ ($\tau \leq t$). The partition function can then be written as $Z_t(s) = \langle \psi_{t-\tau}|\psi_\tau\rangle$, and in particular can be determined by only evolving the vector by $\tau = t/2$.

The average dynamical activity (per unit time and site) follows from the partition sum,

$$k(s) = -\frac{1}{Nt}\frac{d}{ds}\log(Z_t(s)). \quad (10)$$

We can also calculate time-dependent configurational observables for any $0 \leq \tau \leq t$,

$$\langle O(\tau)\rangle_s = Z_t(s)^{-1} \langle \psi_0|e^{(t-\tau)\mathbb{H}_s}Oe^{\tau\mathbb{H}_s}|\psi_0\rangle$$
$$= Z_t(s)^{-1} \langle \psi_{t-\tau}|O|\psi_\tau\rangle. \quad (11)$$

In order to compute the time-evolved state $|\psi_t\rangle$ we use methods from quantum many-body physics, in particular, matrix product states (MPS) (for reviews, see Refs. [**?** **?**]). Here we use both variational optimization of MPS (vMPS, e.g. [**?** **?**]), and time-evolved MPS (tMPS, e.g. [**?**]). Notice that for long times, $|\psi_\tau\rangle$ becomes close to the leading eigenvector of $\mathbb{W}_s$. We exploit this fact to simulate evoluti0n for long times with higher precision, see the Supplemental Material [**?**] for details.

In Ref. [**?**], we used the MPS approximation (from DMRG) to the ground state of $\mathbb{H}_s$ to construct a near-optimal dynamics, which supplemented with trajectory importance sampling (specifically transition path sampling or TPS [**?**]), allowed to efficiently simulate trajectories in large time tilted ensembles. Here we apply the same scheme, but instead use the time-evolved state $|\psi_{t/2}\rangle$. We construct a time-independent dynamics which approximates the optimal (or Doob) dynamics at the centre of finite-time trajectories under tilting [**?**].

Figure 1(a) compares various sampling methods in the East model at $s > 0$. The dynamics is active at short times (due to initial conditions) and inactive at large times [**?** **?** **?**]. We show the activity from the partition sum calculated via MPS time-evolution (black line) as a function of trajectory length. We also show sampling with TPS with the original dynamics (red circles); this method only accounts for the dynamical activity $\langle k \rangle$ at short times, and fails at long times. The scheme introduced in Ref. [**?**] (blue squares) which uses TPS with the MPS approximation to the long-time optimal (Doob) dynamics account for $\langle k \rangle$ at long times [**?**], but fails at short times. If we replace the ground state of $\mathbb{H}_s$ by the time evolved state in the definition of the auxiliary dynamics (green pentagons) we get accurate results for the activity for all trajectory lengths. Despite the fact that the exact Doob dynamics for finite time is in general time-dependent [**?**], this latter approach with a time-independent dynamics for each trajectory length t is efficient enough for TPS to converge to the actual finite-time tilted ensemble, thus correcting any discrepancies. In [**?**] we provide a detailed comparison.

Finite time scaling of active-inactive transition. – The three models we study here display an LD phase transition [**?** **?** **?** **?**] in the long time and large size limit between a dynamical phase where activity is extensive in space and one where activity is subextensive. The finite size scaling analysis of this transition in the long time limit has been studied theoretically [**?** **?** **?** **?** **?**] and numerically [**?** **?**]: for finite size the active-inactive transition is smoothed into a sharp crossover located at $s_c(N, t=\infty) > 0$, which decreases as an inverse power of the system size. In general, however, the location of the transition point depends both on time and size, $s_c(N, t)$, but a detailed numerical analysis of the finite time scaling has not been possible to date due to the difficulty of simulating efficiently rare trajectories at intermediate

times [**?**]. With the approach presented above we can now investigate this issue in detail.

Figure 2(a) shows the dynamical activity $k(s)$ as a function of s and inverse time t^{-1} (East model, top row; FA model, middle; SSEP, bottom). There is a transition from a high activity (light) to low activity (dark) as s is increased which becomes sharper and moves to smaller s with increasing time. The point $s_c(N, t)$ (shown by the red dashed line) is that of the peak in the dynamical susceptibility $\chi(s, t) = dk(s)/ds$. These dynamical phase diagrams are reminiscent of those of (first-order) quantum phase transitions [**?**], with s as an applied field and the inverse time as temperature.

The scaling of the critical point is shown as a function of (inverse) time for multiple system sizes $N \in [20, 200]$ in column (b) of Fig. 2. For small times the critical point scales approximately as $s_c \sim t^{-1}$ for the three models. When time becomes large enough finite-size effects start to dominate. For simplicity, we use the approximate form

$$s_c(N, t) \approx s_c(N) + s_c(t), \tag{12}$$

where $s_c(N) \sim N^{-\alpha}$ can be extracted from vMPS [**?** **?**]. For the FA and East models the exponent $\alpha > 1$ [**?**], while for the SSEP we find the expected $\alpha \approx 2$ [**?**]. In column (c) of Fig. 2 we show how the $s_c(N, t)$ curves can be collapsed, allowing us to estimate $s_c(t) \sim t^{-\beta}$. We find $\beta \approx 1$ for all models.

Also important to the rare event statistics is the probability distribution of the dynamical activity, $P_t(K)$. While for finite times $Z_t(s)$ and $P_t(K)$ do not obey a LD principle in time, for large sizes they still obey one in system size, $Z_t(s) \asymp e^{N\Theta(s,t)}$ and $P_t(K) \asymp e^{-N\Phi_t(K/N)}$. We can therefore obtain the time-dependent rate function $\Phi_t(K)$ through the Legendre transform

$$\Phi_t(k) = -\max_s \left[\Theta(s, t) + sk \right], \tag{13}$$

where $\Theta(s, t) = N^{-1} \log Z_t(s)$. From the numerical estimate of $Z_t(s)$ we can therefore estimate $\Phi_t(K/N)$ for all times. Column (d) of Fig. 2 shows the corresponding rate functions for system size $N = 100$. For small times, the distribution of the activity is close to Poissonian (dashed line), in agreement with the absence of a transition. As time increases the rate function broadens into the characteristic shape of a first order phase transition [**?** **?**].

Structure of the active phase. – Long time trajectories with an atypically large activity are known to display an interesting structure in two of the models we consider here [**?** **?** **?**]. Our finite time method allows to study how such structure depends on the trajectory length.

In Fig. 3(a), we show the average occupation at the mid point of the dynamics. The left panel shows the overall average occupation $\langle n(t/2) \rangle_s$ (the mean occupation in the configuration at time $\tau = t/2$ average over the ensemble of trajectories of total time t) as a function of s for various

t, at $c = 0.05$. The panels on the right show the average spatial profile $\langle n_i(t/2) \rangle_s$ at $s = -0.1$ for $c = 0.05$ (top) and $c = 0.5$ (bottom). In both cases, the average density is spatially featureless at short times, but arranges to maximise activity at long times. For $c = 0.05$ it does so by forming an anti-correlated structure, while these anti-correlations are absent for $c = 0.5$ (cf. the long time case [?]). Figure 3(b) shows the same for the FA model, where there is no appreciable structure forming for small c. Notice also from the left panels the longer times needed to reach the LD behaviour in the East compared to the FA model.

In Fig. 3(c) we quantify the local structure of the SSEP in terms of the nearest-neighbour correlations

$$C_i(t) = \langle n_i n_{i+1}(t) \rangle_s - \langle n_i(t) \rangle_s \langle n_{i+1}(t) \rangle_s , \qquad (14)$$

and the lattice average $C(t) = (N-1)^{-1} \sum_{i=1}^{N-1} C_i(t)$. The right panels show a growth of anti-correlated order with increasing trajectory length towards the "hyperuniform" arrangement at long times, cf. Ref. [?].

Conclusions.– We have implemented a time evolution scheme using MPS to study the rare events of one dimensional KCMs in finite-time trajectories. In this way we have extended recent efforts on the long-time LD statistics via TNs to the arguably harder problem of the LDs away from the long time limit. We showed how to directly compute dynamical partition sums, and we derived an efficient sampling scheme for finite-time rare trajectories. A next step would be to extend these ideas to dimensions larger than one. A possibility could be to implement sampling through two dimensional TNs, such as PEPS (e.g. [? ?]), which have already proven useful in studying the LDs in the long-time limit of two dimensional exclusion processes [?]. While bond dimensions will be limited in this case, using a time evolution scheme like we presented here one could approximate the reference dynamics for the centre of trajectories (i.e. evolve by $e^{t\mathbb{W}_s/2}$) alongside a scheme such as TPS to obtain reliable results. We hope to report on such studies in the near future.
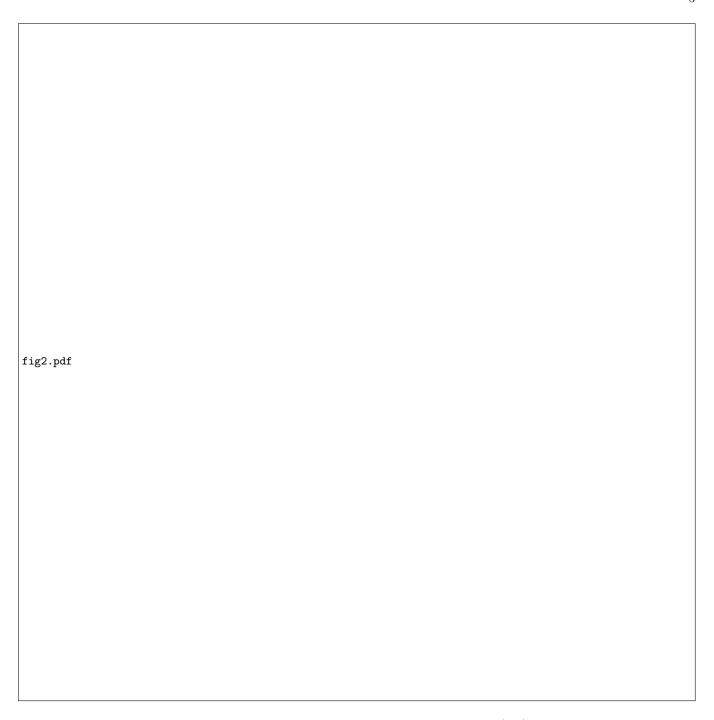
FIG. 2. **Rare event statistics.** The rare event statistics for the East model with $c = 0.5$ (top) and the FA model with $c = 0.2$ (middle) and SSEP (bottom). (a) The dynamical activity $k(s,t)$ as a function of s and inverse time $1/t$ for $N = 100$. The red dotted line indicates our estimate of the critical point. (b) The critical point for various system sizes $N \in [10, 200]$. The dotted lines indicate the infinite time value (see Refs. [? ?]), and the dashed line shows $s_c(N, t) \sim t^{-1}$. (c) The same data is shown in (b) but with $s_c(N, t)$ scaled by $s_c(N)$ and time scaled by $N^{-\alpha}$, where α is the critical exponent extracted from $s_c(N)$. The dotted line shows where the y-axis is one, and the dashed line shows $t^{-\beta}$. The sum of both lines is given by the dashed-dotted line. (d) The estimate of the rate function $\Phi_t(k)$ defined in Eq. (13). The dashed line shows a Poisson distribution with the equilibrium average as its mean.
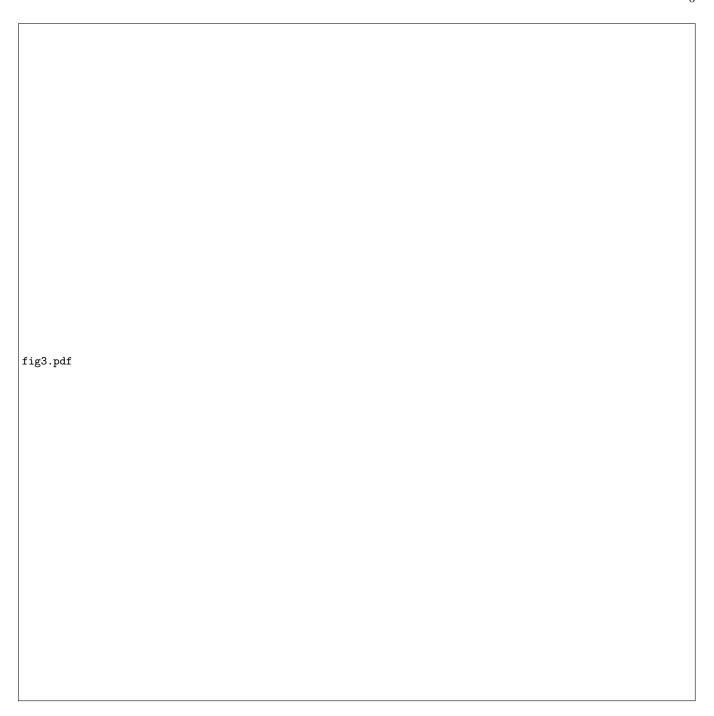
FIG. 3. **Structures in the active phase.** We show the average occupations at the center of the trajectory $\langle n_i(t/2) \rangle_s$ for $s = -0.1$ for the (a) East model and (b) FA models. The left panels of each show the lattice average for a range of s and t with $c = 0.05$, whilst the right panels show the occupations at each site for $c = 0.05$ (top) and $c = 0.5$ (bottom), with $s = -0.1$. We show the same for the SSEP in (c) but with the nearest neighbour correlations $C_i(t/2)$. The right panels are for $s = -0.1$ (top) and $s = -1.0$ (bottom). **[LC: Hoping to replace (a) with $N = 100$.]**