Finite-time and finite-size scalings in the evaluation of large-deviation functions: Numerical approach in continuous time

Esteban Guevara Hidalgo,[1,2,*] Takahiro Nemoto,[2,3,†] and Vivien Lecomte[2,4,‡]

[1]*Institut Jacques Monod, CNRS UMR 7592, Université Paris Diderot, Sorbonne Paris Cité, F-750205, Paris, France*

[2]*Laboratoire de Probabilités et Modèles Aléatoires, Sorbonne Paris Cité, UMR 7599 CNRS, Université Paris Diderot, 75013 Paris, France*

[3]*Philippe Meyer Institute for Theoretical Physics, Physics Department, École Normale Supérieure and PSL Research University, 24 rue Lhomond, 75231 Paris Cedex 05, France*

[4]*LIPhy, Université Grenoble Alpes and CNRS, F-38042 Grenoble, France*

Rare trajectories of stochastic systems are important to understand because of their potential impact. However, their properties are by definition difficult to sample directly. Population dynamics provides a numerical tool allowing their study, by means of simulating a large number of copies of the system, which are subjected to selection rules that favor the rare trajectories of interest. Such algorithms are plagued by finite simulation time and finite population size, effects that can render their use delicate. In this paper, we present a numerical approach which uses the finite-time and finite-size scalings of estimators of the large deviation functions associated to the distribution of rare trajectories. The method we propose allows one to extract the infinite-time and infinite-size limit of these estimators, which—as shown on the contact process—provides a significant improvement of the large deviation function estimators compared to the standard one.

I. INTRODUCTION

Rare events and rare trajectories can be analyzed through a variety of numerical approaches, ranging from importance sampling [1] and adaptive multilevel splitting [2] to transition path sampling [3] (see, e.g., Refs. [4,5] for reviews). In this paper, we focus on population dynamics algorithms, as introduced in Refs. [6,7], which aims at studying rare trajectories by exponentially biasing their probability. This makes it possible to render typical the rare trajectories of the original dynamics in the simulated dynamics. The idea is to perform the numerical simulation of a large number of copies N_c of the original dynamics, supplemented with selection rules which favor the rare trajectories of interest.

The version of the population dynamics algorithm introduced by Giardinà, Kurchan, and Peliti [6] provides a method to evaluate the large deviation function (LDF) associated to the distribution of a trajectory-dependent observable. The LDF is obtained as the exponential growth rate that the population would present if it was not kept constant [8]. Under this approach, the corresponding LDF estimator is in fact valid only in the limits of infinite simulation time t and infinite population size N_c. The usual strategy that is followed in order to obtain those limits is to increase the simulation time and the population size until the average of the estimator over several realizations does not depend on those two parameters, up to numerical uncertainties. The limitations and associated improvements of the population dynamics algorithm have been studied in Refs. [9–12]. In this paper, following a different approach, we propose an original and simple method that takes into account the exact scalings of the finite-t and finite-N_c corrections in order to provide significantly better LDF estimators.

In Ref. [13], we performed an analytical study of a discrete-time version of the population dynamics algorithm. We derived the finite-N_c and finite-t scalings of the systematic errors of the LDF estimator, showing that these behave as $1/N_c$ and $1/t$ in the large-N_c and large-t asymptotics respectively. In principle, knowing the scaling *a priori* means that the asymptotic limit of the estimator in the $t \to \infty$ and $N_c \to \infty$ limits may be interpolated from the data at finite t and N_c. However, whether this idea is actually useful or not is a nontrivial question, as there is always a possibility that onset values of N_c and t scalings are too large to use these scalings. In the present paper, we consider a continuous-time version of the population dynamics algorithms [14,15]. We show numerically that one can indeed make use of these scaling properties in order to improve the estimation of LDF, in an application to a system with many-body interactions (a contact process). We illustrate in Fig. 1 the improvement in the determination of the LDF estimator. We emphasize that the two versions of the algorithm differ on a crucial point, which makes it impossible for an extension of the analysis developed in Ref. [13] to be done straightforwardly in order to comprehend the continuous-time case (see Appendix A). We thus stress that the observation of these scalings themselves is also nontrivial.

The paper is organized as follows. In Sec. II, we introduce the continuous-time cloning algorithm. We define the large deviations of the additive observable of interest and we detail how to estimate them. In Sec. III A, we study the behavior of the LDF estimator as a function of the duration of the observation time (for a fixed population N_c) and we see how its infinite-time limit can be extracted for the numerical data. In Sec. III B, we analyze the behavior of the estimator as we increase the number of clones (for a given final simulation time) and the infinite-size limit of the LDF estimator. Based on these results, we present in Sec. IV, a method which allows us to extract the infinite-time, infinite-size limit of the large deviation function estimator from a finite-time, finite-size scaling analysis. Our conclusions are made in Sec. V. In order

*esteban_guevarah@hotmail.com

†nemoto@lpt.ens.fr

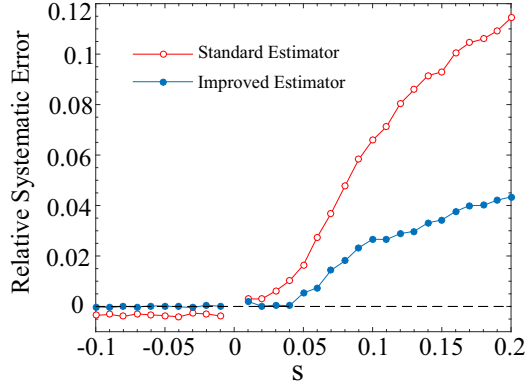‡vivien.lecomte@univ-grenoble-alpes.fr

FIG. 1. Relative systematic error $[\Psi(s) - \psi(s)]/\psi(s)$ between the numerical estimators $\Psi(s)$ and the analytical LDF $\psi(s)$. The error for the standard estimator $\overline{\Psi_s^{(N_c)}(T)}$ is shown in filled circles and for the improved one, f_∞^∞ [Eq. (23)] in empty circles. The scaling method proposed in this paper was tested on the contact process (see Sec. II E 2) (with $L = 6$, $h = 0.1$, and $\lambda = 1.75$) for a set of populations $\tilde{N}_c = \{20, \ldots, 200\}$, a simulation time $T = 100$, and $R = 1000$ realizations. As can be seen, the errors due to finite-size and finite-time effects can be reduced through the improved estimator.

to complement the main discussion done through the paper we also present the following: In Appendix A, an analysis of the difficulty of an analytical approach to the continuous-time algorithm. Then, in Appendix B, an alternative way of defining the LDF estimator is discussed. Finally, in Appendix C, we study the fluctuations of the LDF estimator.

II. CONTINUOUS-TIME CLONING ALGORITHM

A. Large deviations of additive observables

We consider a general Markov process on a discrete space of configurations $\{C\}$, with transition rates $W(C \to C')$. The probability $P(C,t)$ for the system to be in a configuration C at time t verifies a master equation of the form $\partial_t P = \mathbb{W}P$, where the master operator $\mathbb{W}$ is a matrix of elements

$$(\mathbb{W})_{C'C} = W(C \to C') - r(C)\delta_{CC'} \tag{1}$$

and where $r(C) = \sum_{C'} W(C \to C')$ is the escape rate from configuration C. A trajectory of configurations generated in this process is denoted by $(C_0, \ldots, C_K)$, starting from C_0 and presenting K jumps occurring at times $(t_k)_{1 \leqslant k \leqslant K}$. We denote by $C(t')$ the state of the system at time t': When $t_k \leqslant t' < t_{k+1}$, $C(t') = C_k$ $(k = 0,1,2,\ldots,K-1)$ with $t_0 = 0$. We are especially interested in the large deviations of additive observables of the form

$$\mathcal{O} = \sum_{k=1}^{K-1} a(C_k, C_{k+1}) + \int_0^t dt'\, b[C(t')], \tag{2}$$

for trajectories of fixed duration t. The functions a and b describe the elementary increments of the observables: a accounts for quantities associated with transitions (of state), whereas b does for static quantities. A simple example of observables of this form is that of the activity $\mathcal{O} = K$, which is the number of configuration changes on the time interval $[0,t]$ (in this case one has $a(C,C') = 1$ and $b \equiv 0$). We denote the

joint distribution function of the state C and these observables $\mathcal{O}$ at time t by $P(C,\mathcal{O},t)$.

In order to analyze large deviations of these additive observables, we follow the standard procedure as explained, for example, in Refs. [14,15]. For this, we consider the moment generating function

$$Z(s,t) = \langle e^{-s\mathcal{O}} \rangle, \tag{3}$$

where $\langle \cdot \rangle$ is the expected value with respect to trajectories of duration t. The parameter s biases the statistical weight of histories and fixes the average value of $\mathcal{O}$, so that $s \neq 0$ favors its nontypical values. Since the observable $\mathcal{O}$ is additive and the system is described by a Markov process, $Z(s,t)$ satisfies at large times the scaling

$$Z(s,t) \sim e^{t\psi(s)} \quad \text{for } t \to \infty, \tag{4}$$

where $\psi(s)$ is the growth (or decay) rate of $Z(s,t)$ with respect to time. This exponent, known as the scaled cumulant generating function (CGF), is the quantity of interest in this paper. It allows ones to recover the large-time limit of the cumulants of $\mathcal{O}$ as derivatives of $\psi(s)$ in $s = 0$, and more generically, the distribution of $\mathcal{O}/t$ from the Legendre transform of $\psi(s)$ [16], known as a (large deviation) rate function. Hereafter, we use the term "large deviation function" to refer both to the CGF and to the rate function by assuming these two are equivalent. Note that this equivalence is at least satisfied in systems that do not show any phase transition (a singularity in the rate function).

B. The mutation-selection mechanism

The moment generating function $Z(s,t)$ can be computed numerically using the cloning algorithm [6,7]. In order to do that, we introduce the Laplace transform of the probability distribution $P(C,\mathcal{O},t)$, defined as

$$\hat{P}(C,s,t) = \int d\mathcal{O}\, e^{-s\mathcal{O}} P(C,\mathcal{O},t). \tag{5}$$

This Laplace transform allows us to recover the moment generating function as $Z(s,t) = \sum_C \hat{P}(C,s,t)$. The probability $\hat{P}(C,s,t)$ satisfies an "s-modified" master equation for its time evolution (see, e.g., Ref. [17]),

$$\partial_t \hat{P} = \mathbb{W}_s \hat{P}, \tag{6}$$

where the "s-modified" master operator $\mathbb{W}_s$ is defined as

$$(\mathbb{W}_s)_{C'C} = W_s(C \to C') - r_s(C)\delta_{CC'} + \delta r_s(C)\delta_{CC'}. \tag{7}$$

Here, $\delta r_s(C) = r_s(C) - r(C) - sb(C)$,

$$W_s(C \to C') = e^{-sa(C,C')} W(C \to C'), \tag{8}$$

and

$$r_s(C) = \sum_{C'} W_s(C \to C'). \tag{9}$$

Contrary to the original operator (1), the "s-modified" operator (7) does not conserve probability (since $\delta r_s(C) \neq 0$), implying that there is no obvious way to simulate (6). However, this time-evolution equation can be interpreted not as the evolution of a single system but as a population dynamics on a large number N_c of copies of the system which evolve in a coupled way [6,7].

More precisely, reading the operator of the modified master equation (6) as in (7), we find that this evolution equation can be seen as a stochastic process of transition rates $W_s(C \rightarrow C')$ and a selection mechanism of rates

$$\delta r_s(C) = r_s(C) - r(C) - sb(C), \qquad (10)$$

where a copy of the system in configuration C is copied at rate $\delta r_s(C)$ (if $\delta r_s(C) > 0$) or killed at rate $|\delta r_s(C)|$ (if $\delta r_s(C) < 0$). As detailed below, the CGF $\psi(s)$ is recovered from the exponential growth (or decay) rate of a population evolving with these rules.

C. Continuous-time population dynamics (constant-population approach)

The mutation-selection mechanism we just described can be performed in a number of ways. One of them consists in keeping the total number of clones constant for each pre-fixed time interval (see Refs. [6,13] for example). Another one, which we use throughout this paper, consists in performing these selection mechanisms along with each evolution of the copies [5,14,15]. A detailed description of this approach is presented below. See also Appendix A for a brief explanation about important differences between these two techniques.

1. The cloning algorithm

We consider N_c clones (or copies) of the system. The dynamics is continuous in time: For each copy, the actual changes of configuration occur at times (which we call "evolution times") which are separated by intervals whose duration is distributed exponentially. At a given step of the algorithm, we denote by $\mathbf{t} = \{t^{(i)}\}_{i=1,\ldots,N_c}$ the set of the future evolution times of all copies and by $c = \{c_i\}_{i=1,\ldots,N_c}$ the configurations of the copies. Their initial configurations do not affect the resulting scaled cumulant generating function in the large-time limit. However, for the concreteness of the discussion, without loss of generality, we assume that these copies have the same configuration C at $\mathbf{t} = 0$. The cloning algorithm is constituted of the repetition of the following procedures.

1. Find the clone whose next evolution time is the smallest among all the clones: Find $j = \mathrm{argmin}_i t^{(i)}$.

2. Compute $y_j = \lfloor Y(c_j) + \epsilon \rfloor$, where the cloning factor $Y(c_j)$ is defined as $e^{\Delta t(c_j) \delta r_s(c_j)}$, $\Delta t(c_j)$ is the time spent by the clone j in the configuration c_j since its last configuration change, and ϵ is a random number uniformly distributed on [0,1].

3. If $y_j = 0$, remove this copy from the ensemble, and if $y_j > 0$, make $y_j - 1$ new copies of this clone.

4. For each of these y_j copies (if any), the state c_j is changed independently to another state c'_j, with probability $W_s(c_j \rightarrow c'_j)/r_s(c_j)$.

5. Choose a waiting time Δt from an exponential law of parameter $r_s(c'_j)$ for each of these copies. Its next change of configuration will occur at the evolution time $t^{(j)} + \Delta t$.

6. In order to keep the total number of copies constant, we choose randomly and uniformly (i) a clone k, $k \neq j$ and we copy it (if $y_j = 0$), or (ii) $y_j - 1$ clones and we erase them (if $y_j > 1$).

D. Cumulant generating function estimator

The CGF estimator $\Psi_s^{(N_c)}$ can be obtained from the algorithm we just described from the exponential growth rate that the population would present if it was not kept constant [5]. More precisely, this estimator is defined as

$$\Psi_s^{(N_c)} = \frac{1}{t} \log \prod_{i=1}^{\mathcal{K}} X_i, \qquad (11)$$

where $X_i = (N_c + y_i - 1)/N_c$ are the "growth" factors at each step j of the procedure described above, and $\mathcal{K}$ is the total number of configuration changes in the full population up to time t (which has not to be confused with K). It is important to remark (as was discussed in Ref. [8] in a nonconstant population context) that this growth rate can be also computed from a linear fit over the reconstructed log population and the initial transient regime, where the discreteness effects are present, can be discarded in order to obtain a better estimation.

In practice, in order to obtain a good estimation of the CGF, it is normal to launch the simulation several times (where we denote by R the number of realizations of the same simulation) and to estimate the arithmetic mean of the obtained values of (11) over these R simulations. Strictly speaking (as discussed in Sec. 3.2 of Ref. [8]), as the simulation does not stop exactly at the final simulation time T but at some time $t_r^{\mathcal{F}} \leqslant T$ (which is different for every $r \in \{1,\ldots,R\}$), the average over R realizations of $\Psi_s^{(N_c)}$ is then correctly defined as

$$\overline{\Psi_s^{(N_c)}} = \frac{1}{R} \sum_{r=1}^{R} \frac{1}{t_r^{\mathcal{F}}} \log \prod_{i=1}^{\mathcal{K}_r} X_i^r. \qquad (12)$$

However, we have observed that for not too short simulation times, $|\overline{\Psi_s^{(N_c)}(T)} - \overline{\Psi_s^{(N_c)}(t_r^{\mathcal{F}})}|$ is small. By assuming $t_r^{\mathcal{F}} \approx T$, Eq. (12) can be approximated by replacing $t_r^{\mathcal{F}}$ by T (which is what we do in practice). It is important to remark that the CGF estimator can be defined differently from Eq. (12). This is done by using an alternative way of computing the average over R realizations (for an example on this topic, see Appendix B). Equation (12) allows us to estimate the CGF using the constant-population approach of the continuous-time cloning algorithm for a s-biased Markov process, given a fixed number of clones N_c, a simulation time T, and R realizations of the algorithm.

E. Example models

In order to analyze the finite-time and finite-N_c scaling of the CGF estimator, we introduce two specific models: a simple two-state annihilation-creation dynamics and a contact process on a one-dimensional periodic lattice [14,18]. In both cases, we consider the activity K as the additive observable $\mathcal{O}$ and the analytical expression of the CGF $\psi(s)$ was obtained by solving the largest eigenvalue of the operator $\mathbb{W}_s$ given by (7). Below we define these models.

1. Annihilation-creation dynamics

The dynamics occurs in one site where the only two possible configurations C are either 0 or 1. The transition rates are

$$W(0 \rightarrow 1) = c, \quad W(1 \rightarrow 0) = 1 - c, \qquad (13)$$

where $c \in [0,1]$. The analytical expression for the CGF of the activity in this case corresponds to

$$\psi(s) = -\tfrac{1}{2} + \tfrac{1}{2}[1 - 4c(1-c)(1 - e^{-2s})]^{1/2}. \qquad (14)$$

2. Contact process

Each position i of a L-site, one-dimensional lattice is occupied by a spin which is either in the state $n_i = 0$ or $n_i = 1$. The configuration C is then constituted by the states of these spins, i.e., $C = (n_i)_{i=1}^L$. The dynamics occurs on this lattice with periodic boundary conditions with transition rates $W(n_i = 1 \rightarrow n_i = 0) = 1$ and

$$W(n_i = 0 \rightarrow n_i = 1) = \lambda(n_{i-1} + n_{i+1}) + h, \qquad (15)$$

where λ and h are positive constants. This model is an example of contact processes [18], which have been studied in many contexts especially for the spread of infections [19]. It has been known that the corresponding CGF develops a singularity in $L \rightarrow \infty$, showing a dynamical phase transition [14,20].

III. FINITE-TIME AND FINITE-N_c BEHAVIOR OF CGF ESTIMATOR

In this section, we focus on the annihilation-creation process for a peculiar value of s ($s = -0.2$), which is representative of the full range of s on which we study large deviations.

A. Finite-time scaling

Here, we study the large-time behavior of the CGF estimator, at fixed number of clones N_c. Figure 2 presents the average over $R = 10^4$ realizations of the CGF estimator $\overline{\Psi_s^{(N_c)}}$ as a function of the (simulation) time for given numbers of
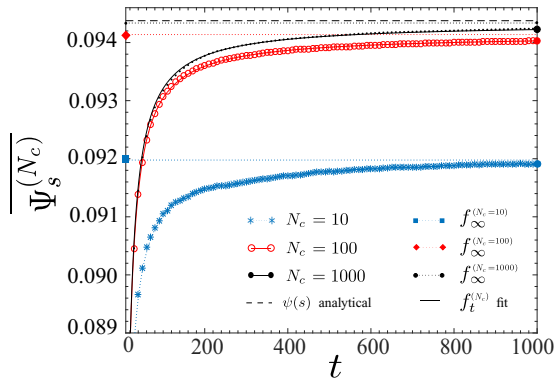

FIG. 2. Average over $R = 10^4$ realizations of the CGF estimator $\Psi_s^{(N_c)}$ [Eq. (12)] as a function of duration t of the observation window, for $N_c \in \{10, 100, 1000\}$ clones, for the annihilation-creation dynamics (13) with $c = 0.3$. The analytical expression for the large deviation function $\psi(s)$ [Eq. (14)] is shown with a black dashed line and the fitting functions $f_t^{(N_c)}$ encoding the finite-t scaling [Eq. (17)] are shown with continuous curves. The (a priori) best estimation of the large deviation function (to which we refer as standard estimator) is given by $\overline{\Psi_s^{(N_c)}}(t)$ at the largest simulation time $T = 1000$, which are shown with solid circles (at the right end of the figure). The extracted infinite-time limits $f_\infty^{(N_c)}$ are shown as dotted lines and squares ($N_c = 10$), diamonds ($N_c = 100$), and circles ($N_c = 1000$).

clones $N_c = \{10, 100, 1000\}$. It is compared with the analytical value $\psi(s)$ [Eq. (14)], which is shown with a black dashed line.

As can be seen in Fig. 2 for a small number of clones ($N_c = 10$), the CGF estimator $\overline{\Psi_s^{(N_c)}}$ is highly deviated from the analytical value $\psi(s)$. However, as N_c and the simulation time t become larger, the CGF estimator get closer to the analytical value $\psi(s)$.

One can expect that in the $t \rightarrow \infty$ and $N_c \rightarrow \infty$ limits, $\psi(s)$ will be obtained from the estimator as

$$\lim_{N_c \to \infty} \lim_{t \to \infty} \overline{\Psi_s^{(N_c)}}(t) = \psi(s), \qquad (16)$$

as it was derived in Ref. [13]. However, in a practical implementation of the algorithm, this infinite-time and infinite-size limits are not achievable and we use large but *finite* simulation time t and number of clones N_c. This fact motivates our analysis of the actual dependence of the estimator with t and N_c. The standard estimator of the large deviation function is the value of $\overline{\Psi_s^{(N_c)}}$ at the largest simulation time T and for the largest number of clones N_c, (e.g., $\overline{\Psi_s^{(N_c)}}(T)$ for $N_c = 1000$ and $T = 1000$, the black solid circle $\bullet$ in Fig. 2). This value provides the (a priori) best estimation of the large deviation function that we can obtain from the continuous-time cloning algorithm. However encouragingly, as we detail later, this estimation can be improved by taking into account the convergence speed of the CGF estimator.

The result of fitting $\overline{\Psi_s^{(N_c)}}(t)$ with the curve $f_t^{(N_c)}$ defined as

$$f_t^{(N_c)} \equiv f_\infty^{(N_c)} + b_t^{(N_c)} t^{-1} \qquad (17)$$

is shown with solid lines in Fig. 2. The fitting parameters $f_\infty^{(N_c)}$ and $b_t^{(N_c)}$ can be determined from the least squares method by minimizing the deviation from $\overline{\Psi_s^{(N_c)}}(t)$. The clear coincidence between $\overline{\Psi_s^{(N_c)}}(t)$ and the fitting lines indicates the existence of a $1/t$ convergence of $\overline{\Psi_s^{(N_c)}}(t)$ to $\lim_{t \to \infty} \overline{\Psi_s^{(N_c)}}(t)$ (that we call $1/t$ scaling). This property can be derived from the assumption that the cloning algorithm itself is described by a Markov process: In Ref. [13] with a different version of the algorithm, we constructed a meta-Markov process to describe the cloning algorithm by expressing the number of clones by a birth-death process. Once such meta process is constructed, the CGF estimator (11) is regarded as the time average of the observable X_i within such meta-Markov process.[1]

We now recall that time-averaged quantities converge to their infinite-time limit with an error proportional to $1/t$ when the distribution function of the variable converges exponentially (as in Markov processes). This leads to the $1/t$ scaling of CGF estimator (17). We note that constructing such meta-Markov process explicitly is not a trivial task, and for the algorithm discussed here, such a construction remains as an open problem.

By assuming the validity of the scaling form (17), it is possible to extract the infinite-time limit of the CGF estimator from finite-time simulations. We denote this infinite-time limit

---

[1] In other words, $t\Psi_s^{(N_c)}$ is an additive observable of the meta process describing the cloning algorithm, as read from (11).
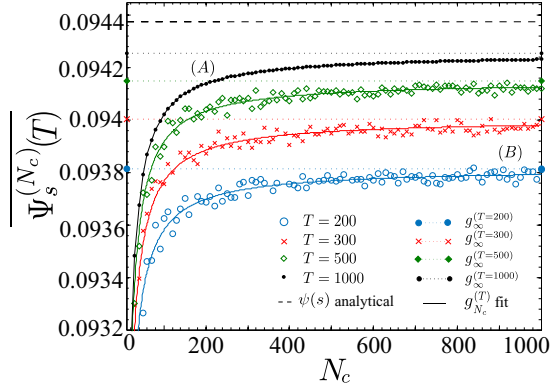
FIG. 3. CGF estimator $\overline{\Psi_s^{(N_c)}(T)}$ [Eq. (12)] for given final (simulation) times $T = \{200,300,500,1000\}$ as a function of the number of clones N_c (on the range $10 \leqslant N_c \leqslant 1000$). The analytical value $\psi(s)$ (14) is shown with a dashed line and the fits $g_{N_c}^{(T)}$ [Eq. (19)] with continuous curves. A large simulation time for a small number of clones, shown in (A), produces a better estimation compared to the one given by the largest number of clones with a relatively short simulation time, which is shown in (B). The best CGF estimation we can naively obtain would be given by $\overline{\Psi_s^{(N_c)}(T)}$ at largest simulation time T and largest number of clones N_c. However, the extracted infinite-size limits $g_\infty^{(T)}$ provide a better estimation in comparison. These limits are shown with dotted lines and circles ($T = 200$), crosses ($T = 300$), diamonds ($T = 500$), and dots ($T = 1000$). Additionally, $c = 0.3$ and $s = -0.2$.

as $f_\infty^{(N_c)}$ and it is expected to be a the better estimator of CGF than $\overline{\Psi_s^{(N_c)}(T)}$ at finite T, provided that

$$f_\infty^{(N_c)} = \lim_{t \to \infty} \overline{\Psi_s^{(N_c)}(t)}. \tag{18}$$

In Fig. 2, we show $f_\infty^{(N_c)}$ with dotted lines and circles ($N_c = 10$), diamonds ($N_c = 100$), and squares ($N_c = 1000$). As can be seen, this parameter indeed provides a better numerical estimate of $\psi(s)$ than $\overline{\Psi_s^{(N_c)}(T)}$.

B. Finite-N_c scaling

Here, we study the behavior of the CGF estimator $\overline{\Psi_s^{(N_c)}(T)}$ as we increase the number of clones N_c, for a given final (simulation) time T. Similar to what we did in Sec. III A, we consider a curve in the form

$$g_{N_c}^{(T)} = g_\infty^{(T)} + \tilde{b}_{N_c}^{(T)} N_c^{-1}, \tag{19}$$

where $g_\infty^{(T)}$ and $\tilde{b}_{N_c}^{(T)}$ are fitting parameters which are determined by the least squares fitting to $\overline{\Psi_s^{(N_c)}(T)}$. The obtained $g_{N_c}^{(T)}$ as a function of N_c are shown in Fig. 3 as solid lines. We considered four values of final simulation time $T = \{200,300,500,1000\}$ and population sizes in the range $10 \leqslant N_c \leqslant 1000$. As can be seen, these curves describe well the dependence in N_c of $\overline{\Psi_s^{(N_c)}(T)}$, indicating that $\overline{\Psi_s^{(N_c)}(T)}$ converges to its infinite-N_c limit with an error proportional to $1/N_c$ (that we call $1/N_c$ scaling). This scaling could be proved under general assumptions in Ref. [13], (i) however, without covering the continuous-time algorithm discussed here and

(ii) for the CGF estimator $\overline{\Psi_s^{(N_c)}(T)}$ considered the $T \to \infty$ limit, instead of finite T. The generalization of the argument presented in Ref. [13] in order to cover the general cases (i) and (ii) is an important open direction of research.

By assuming the validity of such $1/N_c$ scaling, we can evaluate the $N_c \to \infty$ limit of $\overline{\Psi_s^{(N_c)}(T)}$ as the fitting parameter $g_\infty^{(T)}$ obtained from finite N_c simulations as

$$g_\infty^{(T)} = \lim_{N_c \to \infty} \overline{\Psi_s^{(N_c)}(T)}. \tag{20}$$

These parameters $g_\infty^{(T)}$ (to which we refer as infinite-size limit) are shown in Fig. 3 as dotted lines. As shown in the figure, $g_\infty^{(T)}$ provides better estimations of $\psi(s)$ than the one given by the standard estimator $\overline{\Psi_s^{(N_c)}(T)}$.

Complementary to the discussion in this section, in Appendix C we analyze the fluctuations of the CGF estimator.

IV. FINITE-TIME AND FINITE-N_c SCALING METHOD TO ESTIMATE LARGE DEVIATION FUNCTIONS

In the previous section, we have shown how it is possible to extract $f_\infty^{(N_c)}$ and $g_\infty^{(T)}$ from finite T and finite N_c simulations, respectively. In this section, we combine both of these $1/t$- and $1/N_c$-scaling methods in order to extract the infinite-time and infinite-size limit of the CGF estimator. This limit gives a better evaluation of the large deviation function within the cloning algorithm than the standard estimator.

We first note that either of $f_\infty^{(N_c)}$ or $g_\infty^{(T)}$ is expected to converge to $\psi(s)$ as $N_c \to \infty$ or as $T \to \infty$. We checked numerically this property by defining the distance D between $\psi(s)$ and its numerical estimator $\overline{\Psi_s^{(N_c)}}$,

$$D\left[\overline{\Psi_s^{(N_c)}},\psi(s)\right] = \left|\overline{\Psi_s^{(N_c)}} - \psi(s)\right|. \tag{21}$$
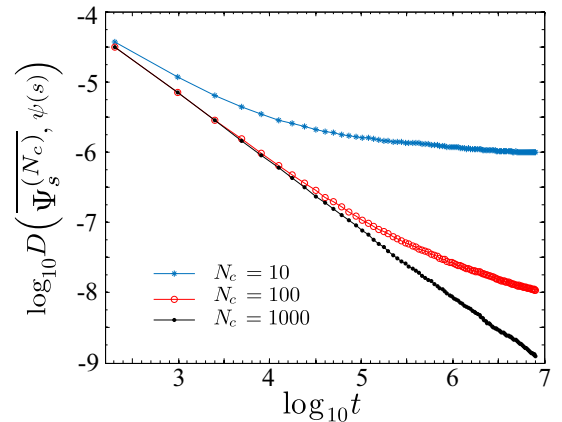

FIG. 4. Distance D [Eq. (21)] between the analytical CGF $\psi(s)$ and its numerical estimator $\overline{\Psi_s^{(N_c)}}$, as a function of time t in log-log scale. The distances are computed from the values in Fig. 2. This distance behaves as a power law of exponent -1 on a time window, where the size of the time window increases as N_c increases. This illustrates the scaling (22). The parameters of the model are $c = 0.3$, $s = -0.2$.
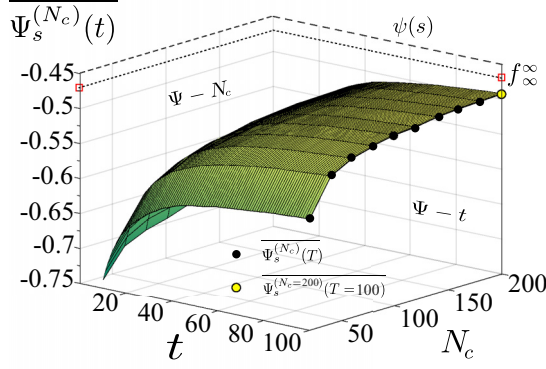
FIG. 5. Estimator of the large deviation function $\overline{\Psi_s^{(N_c)}(t)}$ as a function of time and the number of clones. The estimator $\overline{\Psi_s^{(N_c)}(T)}$ at final simulation time $T = 100$ as a function of the number of clones (up to $N_c = 200$) is shown as black circles. The best CGF estimation under this configuration given by the standard estimator, i.e., $\overline{\Psi_s^{(N_c=200)}(T = 100)}$ is shown as a yellow circle. The analytical value of the CGF $\psi(s)$ is obtained from the largest eigenvalue of the operator (7) and shown as a black dashed line. The extracted limit f_∞^∞ is shown with red squares. Additionally, $L = 6$, $s = 0.15$, $h = 0.1$, $\lambda = 1.75$, and $R = 10^3$.

This quantity is shown in Fig. 4 as a function of t in log-log scale. As we can see, as N_c increases, $\log D$ behaves as straight line with slope -1 on a time window which grows with N_c. In other words, when $N_c \to \infty$,

$$\left| \overline{\Psi_s^{(N_c)}} - \psi(s) \right| \sim t^{-1}. \tag{22}$$

Inspired by this observation, we assume the following scaling for the fitting parameter $f_\infty^{N_c}$. If we consider a set of simulations performed at population sizes $\vec{N}_c = \{N_c^{(1)}, \ldots, N_c^{(j)}\}$, the obtained infinite-time limit of the CGF estimator $f_\infty^{N_c}$ behaves

as a function of N_c as

$$f_\infty^{(N_c)} \simeq f_\infty^\infty + b_\infty^{(N_c)} N_c^{-1}, \tag{23}$$

which means that $f_\infty^{(N_c)}$ itself exhibits $1/N_c$ corrections for large but finite N_c. By using this scaling, we detail below in Sec. IV A the method to extract the infinite-time infinite-N_c limit of the CGF estimator $\overline{\Psi_s^{(N_c)}(T)}$ from finite-time and finite-N_c data. We note that this method can be used for a relatively short simulation time and a relatively small number of clones (see Fig. 6). In Sec. IV B, we present numerical examples of the application of this method to the contact process.

A. The scaling method

The procedure is summarized as follows:

1. Determine the average over R realizations $\overline{\Psi_s^{(N_c)}(t)}$ [Eq. (12)] up to a final simulation time T for each $N_c \in \vec{N}_c$.

2. Determine the fitting parameter $f_\infty^{(N_c)}$ defined in the form $f_t^{(N_c)} = f_\infty^{(N_c)} + b_t^{(N_c)} t^{-1}$ from each of the obtained $\overline{\Psi_s^{(N_c)}(t)}$.

3. Determine f_∞^∞ from a fit in size $f_\infty^{(N_c)} = f_\infty^\infty + b_\infty^{(N_c)} N_c^{-1}$ [Eq. (23)] on $f_\infty^{(N_c)}$.

The result obtained for f_∞^∞ renders a better estimation of $\psi(s)$ than the standard estimator $\overline{\Psi_s^{(N_c)}(t)}$ evaluated for $N_c = \max \vec{N}_c$ and for $t = T$.

B. Application to the contact process

We apply the scaling method to the one-dimensional contact process (see Sec. II E for the definition). We set $L = 6$, $h = 0.1$, $\lambda = 1.75$, $T = 100$, and $s = 0.15$. As we detail below, we compare the improved estimator f_∞^∞ obtained from the application of the scaling method (for $\vec{N}_c = \{20, 40, \ldots, 180, 200\}$) with the standard estimator $\overline{\Psi_s^{(N_c)}(T)}$ (for $N_c = \max \vec{N}_c = 200$).
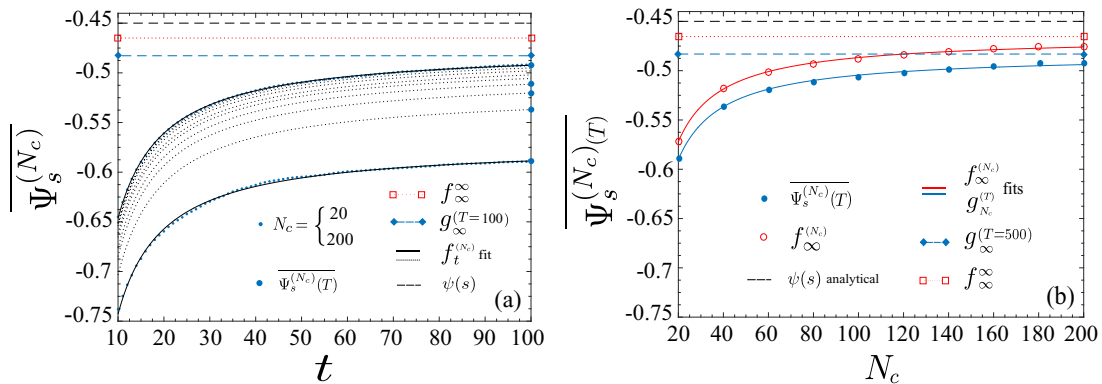

FIG. 6. (a) Projection of the surface represented in Fig. 5 over the plane $\Psi - t$. $\overline{\Psi_s^{(N_c)}(t)}$ is represented for $N_c = 20$ and $N_c = 200$ with blue dots. The estimations $\overline{\Psi_s^{(N_c)}(T)}$ of the large deviation (at the final simulation time $T = 100$) are shown in large blue dots for all the values of N_c considered. The fit in time [Eq. (17)] over $\overline{\Psi_s^{(N_c)}(t)}$ is shown as black solid lines (for $N_c = 20$ and $N_c = 200$) and dotted lines (for other values of N_c). (b) Projection at the final simulation time $T = 100$ on the plane $\Psi - N_c$, $\overline{\Psi_s^{(N_c)}(T)}$ is shown in large blue dots. The infinite-time limit $f_\infty^{(N_c)}$ as a function of N_c [see Eq. (17)] is represented in red circles. The results of fitting $\overline{\Psi_s^{(N_c)}(T)}$ [Eq. (19)] and $f_\infty^{(N_c)}$ [Eq. (23)] are shown with blue and red solid curves respectively. The infinite-N_c limit $g_\infty^{(T)}$ is shown with blue dashed line and diamonds meanwhile the infinite-size and infinite-time limit f_∞^∞ is shown with a red dotted line in both panels (a) and (b). The extracted limit f_∞^∞ renders a better estimate of the large deviation function than $\overline{\Psi_s^{(N_c=200)}(T = 100)}$ (and also than $g_\infty^{(T)}$), demonstrating the efficacy of the method proposed.

Figure 5 represents the behavior of the estimator $\overline{\Psi_s^{(N_c)}}(t)$ as a function of the simulation time t and of the number of clones N_c. The values of the estimator at the final simulation time T are represented with black circles for each $N_c \in \vec{N}_c$ and with a yellow circle for $N_c = \max \vec{N}_c$. The analytical expression for the large deviation function $\psi(s)$ is shown in a black dashed line.

In Fig. 6(a), we show the projection of the surface of Fig. 5 on the plane $\Psi - t$. The behavior in t of the estimator $\overline{\Psi_s^{(N_c)}}(t)$ is shown for $N_c = 20$ and $N_c = 200$, in blue dots in Fig. 6(a). The standard CGF estimators, $\overline{\Psi_s^{(N_c)}}(T)$, are shown in large blue dots in Fig. 6(a) (on the axis for $T = 100$). The fitting curves $f_t^{(N_c)}$ [Eq. (17)] are shown in black continuous lines (for $N_c = 20$ and $N_c = 200$) and black dotted lines (for other intermediate values of N_c). Next, we show in Fig. 6(b) the projection of the surface of Fig. 5 on the plane $\Psi - N_c$ where the time has been set to the largest $t = T$. The standard CGF estimators, $\overline{\Psi_s^{(N_c)}}(T)$ are plotted as blue filled circles, and the fitting curve $g_{N_c}^{(T)}$ [Eq. (19)] on $\overline{\Psi_s^{(N_c)}}(T)$ is shown as a blue solid line. From these curves, we determine $g_\infty^{(T)}$ (see Sec. III B), which is shown as a blue dashed line and diamonds. Finally, the parameter $f_\infty^{(N_c)}$ extracted from the fitting on $\overline{\Psi_s^{(N_c)}}(t)$ (for each value of N_c) is shown as red circles in Fig. 6(b). These values also scale as $1/N_c$ [Eq. (23)] and their fit is shown as a red solid curve. The scaling parameter f_∞^∞ obtained from this last step provides a better estimation of the large deviation function than the standard estimator $\overline{\Psi_s^{(N_c=200)}}(T = 100)$ that is widely used in the application of cloning algorithms. This improvement is valid on a wide range of values of the parameter s as can be visualized in Fig. 1, where we represented the relative systematic error $[\Psi(s) - \psi(s)]/\psi(s)$ between the standard and improved estimators $\Psi(s)$ and the analytical LDF $\psi(s)$.

V. CONCLUSION

Direct sampling of the distribution of rare trajectories is a rather difficult numerical issue (see, for instance, Ref. [21]) because of the scarcity of the nontypical trajectories. We have shown how to increase the efficiency of a commonly used numerical method (the so-called cloning algorithm) in order to improve the evaluation of large deviation functions which quantify the distribution of such rare trajectories, in the large time limit. We used the finite-size and finite-time scaling behavior of CGF estimators in order to propose an improved version of the continuous-time cloning algorithm, which provides more reliable results, less affected by finite-time and finite-size effects. We verified the results observed for the discrete-time version of the cloning algorithm [13] and we showed their validity also for the continuous case. Importantly, we showed how these results can be applied to more complex systems.

We note that the scalings which rule the convergence to the infinite-size and infinite-time limits (with corrections in $1/N_c$ and in $1/t$) have to be taken into account properly: Indeed, as power laws, they present no characteristic size and time above which the corrections would be negligible. The situation is very similar to the study of the critical depinning force in driven

random manifolds: The critical force presents corrections in one over the system size [22], which has to be considered properly in order to extract its actual value. Generically, such scalings also provide a convergence criterion to the asymptotic regimes of the algorithm: One has to confirm that the CGF estimator does present corrections (first) in $1/t$ and (second) in $1/N_c$ with respect to an asymptotic value in order to ensure that such value does represent a correct evaluation of the LDF.

It would be interesting to extend our study of these scalings to systems presenting dynamical phase transitions (in the form of a nonanalyticity of the CGF), where it is known that the finite-time and finite-size scalings of the CGF estimator can be very hard to overcome [14]. In particular, in this context, it would be useful to understand how the dynamical phase transition of the original system translates into anomalous features of the distribution of the CGF estimator in the cloning algorithm. These phase transitions are normally accompanied with an infinite system-size limit (although there was a report of dynamical phase transitions without taking a such limit [23]). To overcome these difficulties (caused by a large system size and/or by the presence of a phase transition), it may be useful to use the adaptive version of the cloning algorithm [24], which has been recently developed to study such phase transitions, with the scaling method presented in this paper.

APPENDIX A: ISSUES ON AN ANALYTICAL APPROACH

In a previous analytical study [13], we considered a *discrete-time* version of the population dynamics algorithm, where a cloning procedure is performed every small time interval Δt. We have proved the convergence of the algorithm in the large-N_c, large-t limits, and we also derived that the systematic error of the LDF estimator (i.e., the deviation of the estimator from the desired LDF) decayed proportionally to $1/N_c$ and $1/t$. From a practical point of view, however, the formulation used there had one problem. In order to prove the result, we took the large frequency limit of cloning procedure or, in other words, we took the $\Delta t \to 0$ limit. A rough estimate of the error due to noninfinitesimal Δt proves to be $O(\Delta t)$. For a faster algorithm, it is better to take this value to be larger, and indeed empirically, we expect that this error to be very small (or rather disappearing in the large-t, N_c limits). However, the detailed analytical estimation of this error is still an open problem.

In the main part of this current paper, from a different point of view, we consider the *continuous-time* version of the population dynamics algorithm [14,15]. Here, the cloning is performed at each change of state of a copy. The time intervals

Δt which separate those changes of state are noninfinitesimal, which means that the formulation we used in Ref. [13] cannot be applied to understand its convergence. Furthermore, because these time intervals are of nonconstant duration and stochastically distributed, the continuous-time algorithm is more difficult to handle analytically than the discrete-time version. Instead of pursuing the analytical study within the continuous-time algorithm, we perform a numerical study, and we show that the $1/N_c$ and $1/t$ scalings are also observed for the continuous-time algorithm. Although the proof of these scalings are beyond the scope of the current paper, these numerical observations support a conjecture that such scaling in large t and in large N_c limits are generally valid in cloning algorithms to calculate large deviation functions.

APPENDIX B: A DIFFERENT CGF ESTIMATOR

Normally, CGF estimator is defined as an arithmetic mean over many realizations, as seen in (12). Here we show that another definition of the CGF estimator can be used, which indeed provides better results than the ones from the standard estimator (in some parameter ranges). We define a new estimator as

$$\Phi_s^{(N_c)} = \frac{1}{T} \log \overline{\prod_{i=1}^{K_r} X_i^r}, \tag{B1}$$

where we note that the average with respect to realizations are taken *inside* the logarithm. As we discussed in Sec. IV C of [13], this estimator provides a correct value of CGF $\psi(s)$ in the infinite-time, infinite-N_c limits. This is thanks to the fact that the distribution of $\Psi_s^{(N_c)}$ concentrates around $\psi(s)$ in those limits (the so-called "self-averaging" property). At any finite population, one can rewrite $\Phi_s^{(N_c)}$ using the large-time LDF principle (C4) as follows:

$$\Phi_s^{(N_c)} = \frac{1}{T} \log \overline{e^{T \Psi_s^{(N_c)}}} \tag{B2}$$

$$= \frac{1}{T} \log \int d\Psi \, e^{-T[I_{N_c}(\Psi) + \Psi]}, \tag{B3}$$

which proves that in the large-T limit,

$$\Phi_s^{(N_c)} = \min_{\Psi} \left[I_{N_c}(\Psi) + \Psi \right], \tag{B4}$$

to be compared to

$$\overline{\Psi_s^{(N_c)}} = \underset{\Psi}{\operatorname{argmin}} \, I_{N_c}(\Psi). \tag{B5}$$

On one hand, the definition (B1) amounts to estimate ψ from the exponential growth rate of the average of the final-T population of many small (noninteracting) "islands," where the cloning algorithm would be operated. On the other hand, the estimator (12) amounts to estimate ψ from growth rate of a large island gathering the full set of the R populations. The later is thus expected to be a better estimator of $\psi(s)$ than the former because it corresponds to a large population, where finite-size effects are less important. As a consequence, the estimator $\Phi_s^{(N_c)}$ appears *a priori* to be a worse estimator than $\overline{\Psi_s^{(N_c)}}$ of $\psi(s)$. However, as shown in Sec. IV C of Ref. [13], at small $|s|$ and finite N_c, a supplementary bias introduced by taking (B1)
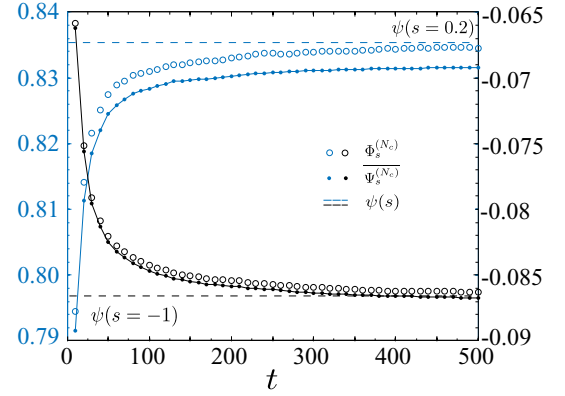

FIG. 7. Comparison between two different estimators of the large deviation function, $\overline{\Psi_s^{(N_c)}}$ [Eq. (12)] shown in dots and $\Phi_s^{(N_c)}$ [Eq. (B1)] in circles, for the annihilation-creation dynamics (13). The analytical value $\psi(s)$ [Eq. (14)] is shown with a dashed line. Here we have also compared two different values of parameter $s = 0.2$ (blue upper curves) and $s = -1$ (black). Additionally, $N_c = 100$, $c = 0.4$, $T = 500$, and $R = 500$. The estimator $\Phi_s^{(N_c)}$ provides a better numerical evaluation of the CGF at small s.

in fact *compensates* for the finite-N_c systematic error presented by (12), for a simple two-state model. Namely, the error is $O(s N_c^{-1})$ for (12) while it is $O(s^2 N_c^{-1})$ for (B1). This fact is illustrated on Fig. 7, where we show that at small $s = 0.2$, $\Phi_s^{(N_c)}$ provides a better estimation of $\psi(s)$ than $\overline{\Psi_s^{(N_c)}}$, while at larger $|s|$ ($s = -1$) the two estimators yield a comparable error.

APPENDIX C: FLUCTUATIONS OF CGF ESTIMATOR

1. Central limit theorem

From relation (12), one can infer that the dispersion of the distribution of $\Psi_s^{(N_c)}$ depends on the simulation time t. This determines whether or not a large number of realizations R is required in order to minimize the statistical error. In fact, as seen in Fig. 8, the dispersion of $\Psi_s^{(N_c)}$ is concentrated around its mean value, which approaches the analytical value $\psi(s)$ as the simulation time and the number of clones increase.

We numerically confirm that these distributions are well approximated by a Gaussian distribution

$$P\left(\Psi_s^{(N_c)}\right) \sim A \, e^{-\frac{1}{C^2}\left(\Psi_s^{(N_c)} - B\right)^2}, \tag{C1}$$

where the parameter B is equal to $\overline{\Psi_s^{(N_c)}}(T)$ and the parameters A and $1/C^2$ are respectively of the order of $N_c^{1/2}$ and N_c. A mathematical argument to explain this obtained Gaussian distribution is given as follows: At any given time (not necessarily at T), let us perform the following rescaling:

$$\hat{\Psi}_s^{(N_c)} = \frac{\Psi_s^{(N_c)} - \overline{\Psi_s^{(N_c)}}}{\sigma_{\Psi_s^{(N_c)}}}, \tag{C2}$$

where $\sigma_{\Psi_s^{(N_c)}}^2$ is the variance of the R realizations of $\Psi_s^{(N_c)}$.

Then, it produces a collapse of the distributions $P(\hat{\Psi}_s^{(N_c)})$, for any t and any N_c (Fig. 9). We remark then that the CGF estimator (12) is an additive observable of the history of the
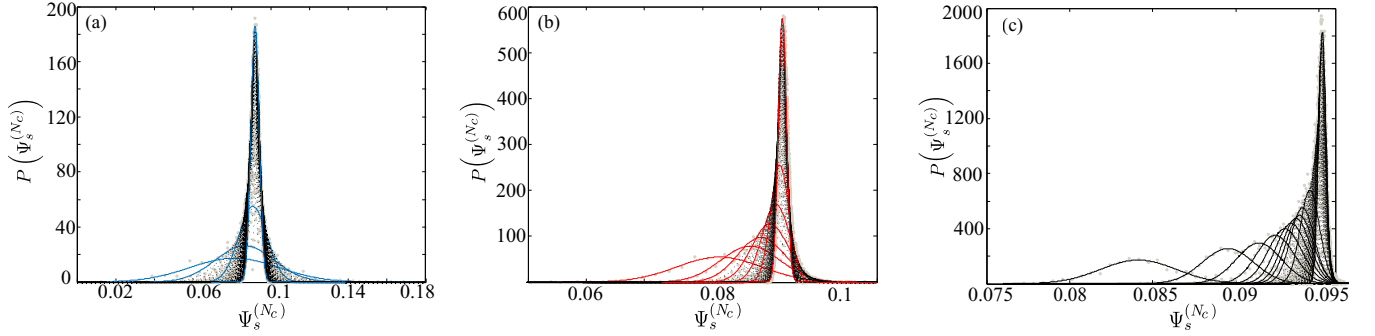
FIG. 8. Distribution $P(\Psi_s^{(N_c)})$ of the CGF estimator $\Psi_s^{(N_c)}$ for (a) $N_c = 10$, (b) $N_c = 100$, and (c) $N_c = 1000$ and for simulation times $t \in [10, 1000]$. Each realization ($R = 10^4$ for each simulation time) is shown with gray dots; meanwhile its respective Gaussian fit [Eq. (C1)] is shown with a dotted or a continuous curve. The dispersion of $\Psi_s^{(N_c)}$ is wider for shorter simulation times and small N_c. The mean value of the distribution converges to the theoretical value as the simulation time and the number of clones increase.

population, which follows a Markov dynamics. Hence, the rescaled estimator $\hat{\Psi}_s^{(N_c)}$ follows a standard normal distribution in the large time limit, according to the central limit theorem (CLT):

$$P\big(\hat{\Psi}_s^{(N_c)}\big) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\hat{\Psi}_s^{(N_c)})^2}. \qquad (C3)$$

We note that this check of the CLT allows us to ensure if the steady state of the population dynamics has been reached (note that in general the typical convergence time to the steady state is larger than the inverse of the spectral gap of the biased evolution operator [8]).

By considering the scaling (C2), we focus only on the small fluctuations of $\Psi_s^{(N_c)}$ around $\overline{\Psi_s^{(N_c)}}$. But in general, the distribution function is not Gaussian, and in that case we need to consider a large deviation principle as below.

2. Logarithmic distribution of CGF estimator

Since $\Psi_s^{(N_c)}$ is itself an additive observable of the dynamics of the ensemble of clones [13], the distribution of the CGF
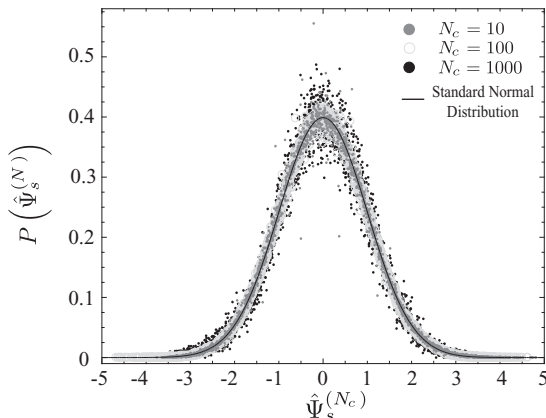

FIG. 9. The distribution function of the rescaled variable $\hat{\Psi}_s^{(N_c)}$ [Eq. (C2)]. Compatible with the central limit theorem, a collapse of the distribution function into a standard normal distribution for different number of clones is observed.

estimator $\Psi_s^{(N_c)}$ satisfies itself a large deviation principle

$$P\big(\Psi_s^{(N_c)}\big) \sim e^{-t \, I_{N_c}(\Psi_s^{(N_c)})}, \qquad (C4)$$

where $I_{N_c}(\Psi_s^{(N_c)})$ is the rate function. This rate function could be evaluated in principle from the empirical distribution $P(\Psi_s^{(N_c)})$ as

$$I_{N_c}\big(\Psi_s^{(N_c)}\big) \approx -\frac{1}{t} \log P\big(\Psi_s^{(N_c)}\big) \qquad (C5)$$

for a large t. Here we try to estimate the rate function from this equation. The numerical estimation of the right-hand side of the last expression at final simulation time T is shown in Fig. 10(a), where we have defined

$$\hat{I}_{N_c}\big(\Psi_s^{(N_c)}\big) \equiv -\frac{1}{t} \log P\big(\Psi_s^{(N_c)}\big) + \frac{1}{t} \log P\big(\overline{\Psi_s^{(N_c)}}\big) \qquad (C6)$$

so that $\hat{I}_{N_c}(\overline{\Psi_s^{(N_c)}}) = 0$. In the same figure, we also show $\overline{\Psi_s^{(N_c)}}(T)$ as vertical dotted lines which correspond to the minima of the logarithmic distribution $\hat{I}_{N_c}(\Psi_s^{(N_c)})$. As can be seen, these minima are displaced towards the analytical value $\psi(s)$ (shown with a dashed line) as $N_c \to \infty$. The logarithmic distribution $\hat{I}_{N_c}$ also becomes more concentrated as N_c increases.

Next, in order to study this decreasing of the width, we show a rescaled logarithmic distribution function $(1/N_c)\hat{I}_{N_c}(\Psi_s^{(N_c)})$ in Fig. 10(b). The minimum converges to the analytical value $\psi(s)$ (black dashed line) as $N_c \to \infty$. In the infinite-time, infinite-size limit of $\Psi_s^{(N_c)}$, it would be thus compatible with a logarithmic distribution function given by

$$I\big(\Psi_s^{(N_c)}\big) = -\lim_{N_c \to \infty} \frac{1}{N_c} \lim_{t \to \infty} \frac{1}{t} \log P(\Psi_s^{(N_c)}(t)), \qquad (C7)$$

which is shown (rescaled) with black dots in Fig. 10(b). By performing the shift $\check{\Psi}_s^{(N_c)} = (\Psi_s^{(N_c)} - \overline{\Psi_s^{(N_c)}})$ we can see in the inset of Fig. 10(b) the superposition of quadratic deviations of the numerical estimator $\Psi_s^{(N_c)}$ around the minimum of $\hat{I}_{N_c}$ (especially for $N_c = 100, 1000$). This indicates the decreasing of the fluctuation of CGF estimator proportional with both of T and N_c (see Ref. [13] for more detailed explanation).
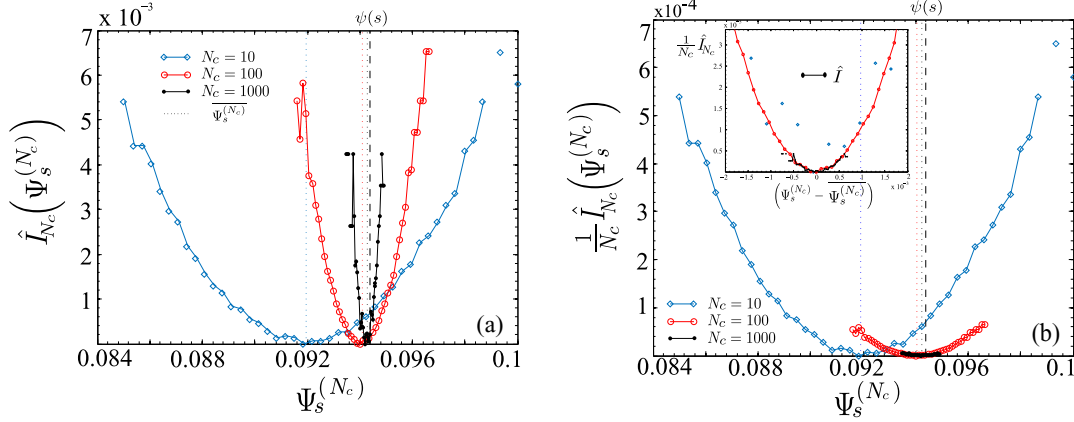
FIG. 10. (a) Logarithmic distribution $\hat{I}_{N_c}(\Psi_s^{(N_c)})$ [Eq. (C6)]. Numerical evaluations were made for three fixed population sizes $N_c \in \{10, 100, 1000\}$ with a fixed simulation time $T = 1000$. The logarithmic distribution presents a smaller width as N_c increases. The average over R realizations of the CGF estimator $\overline{\Psi_s^{(N_c)}}(T)$ corresponds to the minimum of $\hat{I}_{N_c}(\Psi_s^{(N_c)})$ (dotted lines) and converges to the analytical value $\psi(s)$ (dashed lines) as $N_c \to \infty$. (b) Rescaled logarithmic distribution $\frac{1}{N_c}\hat{I}_{N_c}(\Psi_s^{(N_c)})$ as a function of $\Psi_s^{(N_c)}$ and as a function of $\check{\Psi}_s^{(N_c)} = (\Psi_s^{(N_c)} - \overline{\Psi_s^{(N_c)}})$ (inset) for a final simulation time $T = 1000$.

The obtained logarithmic distribution is well approximated by a quadratic form, although these large deviations are in general not quadratic [13]. This means that the direct observation discussed here cannot capture the large deviations of the CGF estimator (see also Ref. [21] for more detailed study of the direct estimation of rate functions). However, we note that, for practical usage of the algorithm, we only consider small fluctuations described by the central limit theorem, although these large fluctuations might play an important role in more complicated systems, such as the ones presenting dynamical phase transitions.

[1] H. Kahn and T. E. Harris, National Bureau of Standards Applied Mathematics Series **12**, 27 (1951).

[2] F. Cérou and A. Guyader, Stoch. Anal. Appl. **25**, 417 (2007).

[3] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Annu. Rev. Phys. Chem. **53**, 291 (2002).

[4] J. Bucklew, *Introduction to Rare Event Simulation* (Springer, Berlin, 2013).

[5] C. Giardinà, J. Kurchan, V. Lecomte, and J. Tailleur, J. Stat. Phys. **145**, 787 (2011).

[6] C. Giardinà, J. Kurchan, and L. Peliti, Phys. Rev. Lett. **96**, 120603 (2006).

[7] J. Tailleur and J. Kurchan, Nat. Phys. **3**, 203 (2007).

[8] E. Guevara Hidalgo and V. Lecomte, J. Phys. A: Math. Theor. **49**, 205002 (2016).

[9] P. I. Hurtado and P. L. Garrido, J. Stat. Mech. (2009) P02032.

[10] M. Tchernookov and A. R. Dinner, J. Stat. Mech. (2010) P02006.

[11] A. Kundu, S. Sabhapandit, and A. Dhar, Phys. Rev. E **83**, 031119 (2011).

[12] T. Nemoto, F. Bouchet, R. L. Jack, and V. Lecomte, Phys. Rev. E **93**, 062123 (2016).

[13] T. Nemoto, E. Guevara Hidalgo, and V. Lecomte, Phys. Rev. E **95**, 012102 (2017).

[14] V. Lecomte and J. Tailleur, J. Stat. Mech. (2007) P03004.

[15] J. Tailleur and V. Lecomte, Modeling and Simulation of New Materials, *Proceedings of Modeling and Simulation of New Materials: Tenth Granada Lectures, Granada (Spain)*, AIP Conf. Proc. 1091 (AIP, Melville, NY, 2009), pp. 212–219.

[16] H. Touchette, Phys. Rep. **478**, 1 (2009).

[17] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, J. Phys. A **42**, 075007 (2009).

[18] T. E. Harris, Ann. Prob. **2**, 969 (1974).

[19] C. Bezuidenhout and G. Grimmett, Ann. Probab. **18**, 1462 (1990).

[20] V. Lecomte, C. Appert-Rolland, and F. van Wijland, J. Stat. Phys. **127**, 51 (2007).

[21] C. M. Rohwer, F. Angeletti, and H. Touchette, Phys. Rev. E **92**, 052104 (2015).

[22] A. B. Kolton, S. Bustingorry, E. E. Ferrero, and A. Rosso, J. Stat. Mech. (2013) P12004.

[23] P. T. Nyawo and H. Touchette, Europhys. Lett. **116**, 50009 (2016).

[24] T. Nemoto, R. L. Jack, and V. Lecomte, Phys. Rev. Lett. **118**, 115702 (2017).