

**Beyond the adiabatic limit in systems with fast environments: A  $\tau$ -leaping algorithm**Ernesto Berríos-Caro<sup>1,\*</sup> and Tobias Galla<sup>1,2,†</sup><sup>1</sup>*Theoretical Physics, Department of Physics and Astronomy, School of Natural Sciences, Faculty of Science and Engineering, The University of Manchester, Manchester M13 9PL, United Kingdom*<sup>2</sup>*Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB), Campus Universitat Illes Balears, E-07122 Palma de Mallorca, Spain* (Received 21 November 2020; revised 17 May 2021; accepted 21 June 2021; published 19 July 2021)

We propose a  $\tau$ -leaping simulation algorithm for stochastic systems subject to fast environmental changes. Similar to conventional  $\tau$ -leaping the algorithm proceeds in discrete time steps, but as a principal addition it captures environmental noise beyond the adiabatic limit. The key idea is to treat the input rates for the  $\tau$ -leaping as (clipped) Gaussian random variables with first and second moments constructed from the environmental process. In this way, each step of the algorithm retains environmental stochasticity to subleading order in the timescale separation between system and environment. We test the algorithm on several toy examples with discrete and continuous environmental states and find good performance in the regime of fast environmental dynamics. At the same time, the algorithm requires significantly less computing time than full simulations of the combined system and environment. In this context we also discuss several methods for the simulation of stochastic population dynamics in time-varying environments with continuous states.

DOI: [10.1103/PhysRevE.104.014122](https://doi.org/10.1103/PhysRevE.104.014122)**I. INTRODUCTION**

The modeling of dynamical systems in biology and other disciplines necessarily requires simplifying assumptions and a level of coarse graining. If all processes we know about are included, then the model becomes so complicated that it cannot be simulated or analysed. Even if simulation or analysis is possible further study of such a model will rarely be enlightening. Excessive detail makes hard to identify the key mechanisms at work and to understand what model components are responsible for these mechanisms. At the same time, some element of realism must be maintained. The model must not be so stylized to miss the key ingredients and behavior it is meant to capture. The principal challenge, therefore, is to find the right level of detail, given the intended purpose.

The choice between stochastic and deterministic modeling approaches is one aspect of this discussion. If more detailed stochastic models mark one end of the spectrum, then many traditional models in mathematical biology or chemistry sit at the opposite end. These models are often built on a small number of ordinary or partial differential equations (e.g., Refs. [1,2]). This deterministic approach is valid if one can assume that the same initial conditions will always lead to the same outcome. For many applications involving very large systems this is a perfectly sensible approach.

However, it is now also universally recognized that stochasticity in the time-evolution of many systems is key in shaping the outcome, see, e.g., Refs. [3–5]. Consequently a number of analytical and computational methods has been developed for the study of stochastic systems. One focus is on systems with discrete interacting individuals. What these

individuals represent depends on the context, they could be members of different species in population dynamics, individual animals or humans in models of an epidemic, or molecules in chemical reaction systems [3,6–8].

One particular point of interest within this class of individual-based systems are models operating in a time-dependent environment. This environment is not part of the system proper, but its state has an effect on what happens in the system. In a model of a population of bacteria for example, the reproduction or death rates could depend on external conditions such as the availability of nutrients or the presence of toxins [9,10]. In population dynamics, the carrying capacity could vary in time [11–13], and in epidemics the infection rate is subject to seasonal changes [14]. The focus of our paper is on such individual-based models in time-varying external environments.

Analytical approaches to stochastic systems with discrete individuals usually start from the chemical master equation. In limited cases direct solution is possible, for example using generating functions. However, this is the exception, and a number of approximation schemes have consequently been developed. These include Kramers–Moyal and system-size expansions, leading to Fokker–Planck equations and descriptions in terms of stochastic differential equations [8,15]. These schemes sacrifice the granularity of a discrete-agent system, and instead describe the dynamics in terms of continuous densities. This approach can be successful in particular for large populations. Any particular event then only results in a small change in the composition of the population relative to its size. Individual-based approaches and descriptions based on deterministic differential equations have been extended to models of population dynamics in switching environments; for a selection of work, see Refs. [11,16–26].

There are however situations in which one would rather avoid giving up the discrete nature of the population. For

\*ernesto.berrios@postgrad.manchester.ac.uk

†tobias.galla@ifisc.uib-csic.es

example, granularity is crucial for extinction processes (the number of individuals of the species about to go extinct is small by definition). In other situations the population may not be large enough to warrant a description in terms of continuous densities. For example, copy numbers in genetic circuits can be of the order of tens to hundreds (see, e.g., Ref. [27]), and it is difficult to justify a continuum limit. It then becomes necessary to carry out numerical simulations of the discrete individual-based process. The method of choice is the Gillespie algorithm [28,29], generating a statistically accurate ensemble of sample paths of the continuous-time dynamics.

In most applications the rate of events scales with the size of the population so that each individual experiences an  $\mathcal{O}(1)$  number of reactions per unit time. The Gillespie method then runs into difficulties when the population is large, and with it the number of reactions per unit time. The computational cost of generating sample paths to up the desirable end point can then become very high. Similarly, a timescale separation between the dynamics in the population and the environment may make simulations challenging. If the environment is very fast compared to the population, then a significant number of environmental events needs to be executed between events in the population. This aggravates the above limitations for large populations, and simulations can become problematic even for intermediate population sizes. One possible approach to this consists of assuming that the environment is “infinitely” fast compared to the population. This is known as quasi steady state approximation [30,31], or the “adiabatic limit” [26,32]. For related work see also [23,33–36]. If this limit is taken, then the environmental dynamics can be “averaged out” and effective reaction rates can be used for the population. While computationally convenient, this approach discards any stochasticity from the environmental process. This sets another limitation, in particular when it is not valid to assume that the environment is infinitely fast compared to the population.

The objective of this work is therefore to design and test an algorithm for systems with fast environmental dynamics, but which also captures some elements of the environmental noise. We call this discrete-time algorithm  $\tau$ FE— $\tau$ -leaping for fast environments. It is built on the ideas of the conventional  $\tau$ -leaping algorithm [37], but with modifications such as to preserve elements of the stochasticity of the environmental process. To do this, we assume that the environment is fast compared to the population, but not infinitely fast. More precisely, in each step of the algorithm we take into account subleading contributions in the timescale separation.

The key new element of our algorithm is how we deal with the environment. We do not take the adiabatic limit, instead we treat the reaction rates in the population as random variables during each step. The rates are drawn from a distribution at the beginning of each step, and then remain fixed during the time step. The distribution of rates can change from one step to the next, and is constructed to reflect statistical features of the original environmental dynamics.

Each step of the  $\tau$ FE algorithm consists of two parts: First a realization of reaction rates is drawn from the appropriate distribution. Then a conventional  $\tau$ -leaping step is carried out with these rates. The core of our paper consists of the construction of the “appropriate distribution” for the reaction rates. These ideas were introduced in a previous work [38]

for a simple case of a two-species birth-death process in an environment which can take two discrete states. In the present paper we develop this further. We develop and test a more general algorithm for environments with more than two discrete states. As we will show, the algorithm can also be extended to continuous environmental dynamics.

The remainder of the paper is organized as follows. In Sec. II we describe the general setup of the type of system we simulate. We also outline the general principles of the  $\tau$ FE algorithm. In Sec. III we then make the necessary preparations for the introduction of the algorithm. In particular, we compute the statistics of reaction rates which are fed into the conventional  $\tau$ -leaping step. We then describe the algorithm in detail. In Sec. IV, we test the  $\tau$ FE algorithm in different models with discrete environmental states. In Sec. V, we then describe how the  $\tau$ FE algorithm can be used when the environment takes continuous states. Specifically, we consider an Ornstein-Uhlenbeck process. In this context we also describe how known algorithms can be adapted to simulate continuous environments. Finally, we provide a discussion of our results and overall conclusions in Sec. VII.

## II. MODEL SETUP AND GENERAL PRINCIPLES OF THE ALGORITHM

### A. Model definitions and notation

We look at systems composed of discrete individuals. We will refer to this synonymously as the “system proper” or “the population.” Each of the individuals is of one of  $S$  species (or types), labeled  $i = 1, \dots, S$ . We write  $n_i$  for the number of individuals of species  $i$  in the population, and  $\mathbf{n} = (n_1, \dots, n_S)$ . The system evolves in an external environment, whose state we write as  $\sigma$ . These states are time dependent, and can either take discrete values or be continuous.

The dynamics in the population proceeds through reactions  $r = 1, \dots, R$ . Each of these reactions converts a number of individuals from one type into another. Time in the model is continuous, and we assume that the dynamics is Markovian. We then write  $R_{r,\sigma}(\mathbf{n})$  for the rate of reaction  $r$  if the environment is in state  $\sigma$  and the population in state  $\mathbf{n}$ . The stoichiometric coefficient  $\nu_{r,i}$  indicates how the number of individuals of type  $i$  changes when a reaction of type  $r$  occurs. Each  $\nu_{r,i}$  is an integer, which can be negative, zero, or positive. We write  $\mathbf{v}_r = (\nu_{r,1}, \dots, \nu_{r,S})$ . The rates  $R_{r,\sigma}(\mathbf{n})$  and the stoichiometric coefficients fully specify the dynamics of the population when the environment is in state  $\sigma$ .

The state of the environment undergoes a Markovian stochastic process, governed by a master equation if states are discrete or by a stochastic differential equation in the case of continuous environmental states. These dynamics can depend on the state of the population  $\mathbf{n}$ . If the environmental states are discrete, then we write  $q_{\sigma \rightarrow \sigma'}(\tau)$  for the probability of finding the environment in state  $\sigma'$  at a particular point in time, given that  $\tau$  units of time earlier it was in state  $\sigma$ . If the environment is continuous, then  $q_{\sigma \rightarrow \sigma'}(\tau)$  is a probability density for  $\sigma'$  (at given  $\sigma$ ). We call  $q_{\sigma \rightarrow \sigma'}(\tau)$  the transition kernel of the environmental process. We write  $\rho^*$  for the stationary distribution of the environmental dynamics. For discrete environmental states the entries  $\rho_\sigma^*$  denote the prob-

ability of finding the system in state  $\sigma$  in the stationary state. For continuous environments  $\rho_\sigma^*$  is the stationary probability density for  $\sigma$ .

### B. General principles of the $\tau$ -leaping algorithm for systems in fast environments

A conventional reaction system (without external environment) is governed by a wchemical master equation of the form

$$\frac{d}{dt}P(\mathbf{n}, t) = \sum_r [R_r(\mathbf{n} - \mathbf{v}_r)P(\mathbf{n} - \mathbf{v}_r, t) - R_r(\mathbf{n})P(\mathbf{n}, t)] \quad (1)$$

The notation is as in Sec. II A; the only difference is that there is no subscript  $\sigma$ , as there is no environment. Sample paths entail events (reactions) which can occur at any point in continuous time, separated by exponentially distributed random waiting times. In each such event the state of the system  $\mathbf{n}$  changes, and accordingly the reaction rates  $R_r(\mathbf{n})$  can also change. Sample paths can be generated for example using the celebrated Gillespie algorithm [28,29].

The  $\tau$ -leaping algorithm for such conventional reaction systems is built around the idea of keeping reaction rates constant over finite time steps of length  $\tau$  [37]. That is to say, if the state of the population is  $\mathbf{n}$  at time  $t$ , then the assumption is made that this state  $\mathbf{n}$  and the rates  $R_r(\mathbf{n})$  do not change until the end of the time step. The algorithm does not account for potential changes of the rates as individual reactions occur, and instead directly “leaps” to time  $t + \tau$ . This is justified provided the so-called “leap condition” is fulfilled [37]: broadly speaking the time step  $\tau$  must be sufficiently small so that the state  $\mathbf{n}$  in the continuous-time system does not change significantly in a time interval of length  $\tau$ .

Making the approximation of constant  $\mathbf{n}$  in the time interval from  $t$  to  $t + \tau$ , the number of reactions of type  $r$  that fire in this interval follows a Poissonian distribution with parameter  $\tau R_r(\mathbf{n})$ . Accordingly, realizations of Poissonian random variables  $m_1, \dots, m_R$  are drawn, and the corresponding numbers of each reactions are executed simultaneously. This generates a new state  $\mathbf{n}'$  at time  $t + \tau$ , with entries  $n'_i = n_i + \sum_{r=1}^R m_r \nu_{i,r}$ . The process then repeats with updated rates  $R_r(\mathbf{n}')$ .

The idea of the  $\tau$ -leaping algorithm we introduce for systems in external environments is similar. As in the conventional algorithm, we discretize time and keep the composition of the population  $\mathbf{n}$  fixed during each iteration. It is only updated at the end of each step. From now on we use  $\Delta t$  for the duration of a step instead of  $\tau$ .

The difference to the conventional case is the external environment. If the environmental state space is discrete, then switches of the environment can in principle be simulated in continuous time along with the other reactions (using Gillespie algorithm [28,29]). They can also be dealt with by means of the conventional  $\tau$ -leaping algorithm, again along with the other reactions. These are natural simulation approaches when the environment operates on a similar timescale as the reactions in the population. Not much can then be gained by distinguishing between environmental processes and the dynamics in the system proper.

If the environment is infinitely fast compared to the reactions in the population, then the environment reaches stationarity on very short timescales. One can average over environmental states, see for example [26,30,31,33,38,39]. If the environment is discrete, for example, then we can use average rates

$$R_r^*(\mathbf{n}) \equiv \sum_\sigma \rho_\sigma^* R_{r,\sigma}(\mathbf{n}). \quad (2)$$

In the case of continuous environments the sum is to be replaced with an integral. These rates are functions of  $\mathbf{n}$  only, the environmental process has been averaged out. Noise from the environmental process plays no role in the dynamics if these average rates are used. This corresponds to making a quasi-stationary state approximation for the fast-moving environment [30,31].

The aim of this paper is to go beyond this adiabatic limit, and to construct a  $\tau$ -leaping algorithm which captures some elements of extrinsic noise. We focus on the limit of a fast, but not infinitely fast environmental dynamics.

Broadly speaking the  $\tau$ FE algorithm is constructed around the idea of treating the reaction rates  $R_r(\mathbf{n})$  as stochastic variables in each discrete time step. These random variables represent the rates one obtains when averaging the environmental process over the time step  $\Delta t$ . Assuming that the rate of change of the environment is finite these average rates will remain stochastic. In the limit of infinitely fast environments the deterministic limit in Eq. (2) is recovered, and there is no stochasticity from the environment.

To construct the random reaction rates for each step, we make an approximation: we use a Gaussian distribution for the rates, with means as in Eq. (2) and with variances and correlations derived from the original combined process of the population and environment. We describe this in detail in the next section.

## III. CONSTRUCTION OF THE $\tau$ FE ALGORITHM FOR SYSTEMS WITH DISCRETE ENVIRONMENTS

### A. Preliminary analysis of the environmental process

Here we assume the environmental states are discrete,  $\sigma \in \{1, \dots, M\}$ . The dynamics of the environment is governed by the rates  $\lambda A_{\sigma \rightarrow \sigma'}(\mathbf{n})$  for transitions from  $\sigma$  to  $\sigma'$ . The factor  $\lambda$  is introduced to control the timescale separation between reactions in the population and the switching of the environment. We use the notation  $\mathbf{A}(\mathbf{n})$  for the  $M \times M$  matrix with elements  $A_{\sigma \rightarrow \sigma'}(\mathbf{n})$ . We also set  $A_{\sigma \rightarrow \sigma}(\mathbf{n}) = -\sum_{\sigma' \neq \sigma} A_{\sigma \rightarrow \sigma'}(\mathbf{n})$ . The combined dynamics of population and environment are then described by the master equation

$$\begin{aligned} \frac{d}{dt}P(\mathbf{n}, \sigma, t) &= \sum_r [R_{r,\sigma}(\mathbf{n} - \mathbf{v}_r)P(\mathbf{n} - \mathbf{v}_r, \sigma, t) - R_{r,\sigma}(\mathbf{n})P(\mathbf{n}, \sigma, t)] \\ &\quad + \lambda \sum_{\sigma'} [A_{\sigma' \rightarrow \sigma}(\mathbf{n})P(\mathbf{n}, \sigma', t) - A_{\sigma \rightarrow \sigma'}(\mathbf{n})P(\mathbf{n}, \sigma, t)]. \quad (3) \end{aligned}$$

The rates  $\lambda A_{\sigma \rightarrow \sigma'}(\mathbf{n})$  can depend on the state of the population,  $\mathbf{n}$ . This means that  $\mathbf{n}$  and  $\sigma$  do not necessarily evolve in time independently. However, as mentioned above the state  $\mathbf{n}$

of the system is kept constant during each  $\tau$ -leaping step. This in turn means that the transition rates for the environment also remain constant during each step.

We now focus on one such time step, starting at time  $t$  and ending at  $t + \Delta t$ . We assume that  $\mathbf{n}$  remains constant during this time interval. For the remainder of Sec. III A we suppress the potential dependence of  $\mathbf{A}$  on  $\mathbf{n}$ , although it is always implied. We write  $\rho_\sigma(t')$  for the probability that the environment is in state  $\sigma$  at time  $t' \in [t, t + \Delta t]$ . We then have the master equation

$$\frac{d\rho}{dt'} = \lambda \mathbf{A} \rho \quad (4)$$

for the environmental dynamics. The stationary distribution  $\rho^*$  for the environment is the solution of  $\mathbf{A}\rho^* = 0$ . If  $\mathbf{A}$  depends on  $\mathbf{n}$ , then  $\rho^*$  will also be a function of  $\mathbf{n}$ . Assuming that the environmental process is irreducible this stationary distribution is unique for any one  $\mathbf{n}$ .

The stochastic matrix  $\mathbf{A}$  has one zero eigenvalue, which we write as  $\mu_1 = 0$ . The remaining eigenvalues are denoted by  $\mu_2, \dots, \mu_M$ . We then have  $\mu_2, \dots, \mu_M < 0$ . The corresponding (right) eigenvalues are written as  $\mathbf{v}_1 = \rho^*$  (the eigenvector corresponding to eigenvalue 0), and  $\mathbf{v}_2, \dots, \mathbf{v}_M$ , respectively, for the remaining eigenvectors. These are all understood to be column vectors of length  $M$ .

We note that the general solution of Eq. (4) can be written in the form

$$\rho(t') = \rho^* + \sum_{\ell=2}^M c_\ell e^{\lambda \mu_\ell (t' - t)} \mathbf{v}_\ell, \quad (5)$$

with coefficients  $c_\ell$  determined by the initial condition at the beginning of the time step  $t' = t$ . More precisely these coefficients can be obtained from the linear system

$$\sum_{\ell=2}^M c_\ell \mathbf{v}_\ell = \rho(t) - \rho^*. \quad (6)$$

We remark that there are  $M - 1$  coefficients,  $c_\ell$  ( $\ell = 2, \dots, M$ ). The system in Eq. (6) technically consists of  $M$  equations, but these are not independent due to normalization of the probabilities on the right-hand side. We assume that the Markov process for  $\sigma$  has a unique stationary distribution  $\rho^*$  with full support, and that all other conditions are met which are required so that set of coefficients  $c_\ell$  in Eq. (6) is well defined and unique. In other words, we assume that the environmental process  $\sigma(t)$  is a Markov process that is irreducible (i.e., all states communicate) and positive recurrent (i.e., there exists a finite expected time to return to any state) so the above conditions are met [40].

Calculating the probability  $q_{\sigma \rightarrow \sigma'}(\Delta t)$  to find the environment in state  $\sigma'$  at the end of the time step if it was in  $\sigma$  at the beginning of the step is now mainly a matter of computing the coefficients  $c_\ell$ . We write  $c_{\ell, \sigma}$  for the value the coefficient  $c_\ell$  takes when  $\rho_{\sigma'}(t) = \delta_{\sigma', \sigma}$  (i.e., when the system starts in state  $\sigma$  at the beginning of the step).

We then have

$$q_{\sigma \rightarrow \sigma'}(\Delta t) = \rho_{\sigma'}^* + \sum_{\ell=2}^M c_{\ell, \sigma} e^{\lambda \mu_\ell \Delta t} v_{\ell, \sigma'}, \quad (7)$$

where  $v_{\ell, \sigma'}$  is the  $\sigma'$ -entry of the eigenvector  $\mathbf{v}_\ell$  of  $\mathbf{A}$ .

If the matrix  $\mathbf{A}$  depends on the population state  $\mathbf{n}$ , then the parameters  $\mu_\ell$ ,  $v_{\ell, \sigma'}$ , and  $c_{\ell, \sigma}$  can also depend on  $\mathbf{n}$ . For simplicity of notation we have not included this potential dependence in the above equations.

### B. Time-averaged reaction rates as random variables

The  $\tau$ -leaping algorithm proceeds in discrete time intervals of length  $\Delta t$ . We continue to focus on one such interval  $[t, t + \Delta t]$ . The state of the population at the beginning of the step is  $\mathbf{n}$  and we assume that this state does not change until the end of the interval. We do however take into account the fact that the state of environment  $\sigma$  can undergo changes in the interval from  $t$  to  $t + \Delta t$ . As a consequence,  $R_{r, \sigma}(\mathbf{n})$  (at fixed  $\mathbf{n}$ ) is also a function of time.

We then introduce the time-averaged quantities

$$\overline{R}_r(\mathbf{n}) \equiv \frac{1}{\Delta t} \int_t^{t+\Delta t} dt R_{r, \sigma(t)}(\mathbf{n}), \quad (8)$$

noting that the time average is *over the duration of the time step only* as opposed to a long-term asymptotic time average. Given that the time step  $\Delta t$  is finite ( $\Delta t < \infty$ ) and assuming that the environment fluctuates with finite rates ( $\lambda < \infty$ ), the quantity  $\overline{R}_r(\mathbf{n})$  is a stochastic variable as it depends on the realization of the environmental process. In one given time interval, the rates  $\overline{R}_r(\mathbf{n})$  for different  $r$  will be correlated as they all derive from the same path of the environment. As we will show below, the fluctuations of the random variables  $\overline{R}_r(\mathbf{n})$  in any one time step are inversely proportional to  $\lambda \Delta t$  to leading order. In the limit  $\lambda \Delta t \rightarrow \infty$  the  $\overline{R}_r(\mathbf{n})$  become deterministic.

We assume that the distribution for  $\sigma$  at the beginning of the time step is the stationary distribution  $\rho^*$ . This is the case, for example, if then environmental state is drawn from the stationary distribution at the beginning of the simulation. The distribution for  $\sigma(t')$  is then also the stationary distribution at each time  $t' \in [t, t + \Delta t]$ . Writing  $\langle \dots \rangle$  for the average over realizations of the environmental process, we then have

$$\langle \overline{R}_r(\mathbf{n}) \rangle = R_r^*(\mathbf{n}), \quad (9)$$

with  $R_r^*(\mathbf{n})$  as in Eq. (2).

However,  $\sigma(t')$  ( $t' \in [t, t + \Delta t]$ ) will generally be correlated with  $\sigma(t)$ . Neglecting these correlations means to operate in the adiabatic limit. We would like to retain some of these correlations. To compute second moments  $\langle \overline{R}_r(\mathbf{n}) \overline{R}_s(\mathbf{n}) \rangle$  we first use Eq. (8). This leads to averages of the type  $\langle R_{r, \sigma(t_1)}(\mathbf{n}) R_{s, \sigma(t_2)}(\mathbf{n}) \rangle$ , where  $t_1$  and  $t_2$  are two times in the interval from  $t$  to  $t + \Delta t$ . The second moments can then be expressed in terms of  $q_{\sigma \rightarrow \sigma'}(\cdot)$  as follows:

$$\begin{aligned} \langle \overline{R}_r(\mathbf{n}) \overline{R}_s(\mathbf{n}) \rangle &= \frac{1}{\Delta t^2} \sum_{\sigma \sigma'} \int_t^{t+\Delta t} dt_1 \int_{t_1}^{t_1+\Delta t} dt_2 \{ \rho_\sigma^* q_{\sigma \rightarrow \sigma'}(t_2 - t_1) \\ &\quad \times [R_{r, \sigma}(\mathbf{n}) R_{s, \sigma'}(\mathbf{n}) + R_{r, \sigma'}(\mathbf{n}) R_{s, \sigma}(\mathbf{n})] \}. \end{aligned} \quad (10)$$



Further details are given in Appendix A. Using Eq. (7) we find

$$\begin{aligned} & (\bar{R}_r(\mathbf{n})\bar{R}_s(\mathbf{n})) - R_r^*(\mathbf{n})R_s^*(\mathbf{n}) \\ &= \frac{1}{\Delta t^2} \sum_{\sigma\sigma'} \sum_{\ell=2}^M \left\{ \rho_{\sigma}^* c_{\ell,\sigma} v_{\ell,\sigma'} \right. \\ & \quad \times [R_{r,\sigma}(\mathbf{n})R_{s,\sigma'}(\mathbf{n}) + R_{r,\sigma'}(\mathbf{n})R_{s,\sigma}(\mathbf{n})] \\ & \quad \times \left. \int_t^{t+\Delta t} dt_1 \int_{t_1}^{t_1+\Delta t} dt_2 e^{\lambda\mu_{\ell}(t_2-t_1)} \right\}. \end{aligned} \quad (11)$$

For fixed  $\ell \in \{2, \dots, M\}$  the integral in the last expression evaluates to

$$\int_t^{t+\Delta t} dt_1 \int_{t_1}^{t_1+\Delta t} dt_2 e^{\lambda\mu_{\ell}(t_2-t_1)} = -\frac{\Delta t}{\lambda\mu_{\ell}} + \frac{1}{(\lambda\mu_{\ell})^2} [e^{\lambda\mu_{\ell}\Delta t} - 1]. \quad (12)$$

For  $\lambda\Delta t \gg 1$  the first term dominates after inserting into Eq. (11), as also observed in Ref. [41]. We are then left with

$$\begin{aligned} & (\bar{R}_r(\mathbf{n})\bar{R}_s(\mathbf{n})) - R_r^*(\mathbf{n})R_s^*(\mathbf{n}) \\ &= -\frac{1}{\lambda\Delta t} \sum_{\sigma\sigma'} \sum_{\ell=2}^M \left\{ \frac{1}{\mu_{\ell}} R_{r,\sigma}(\mathbf{n})R_{s,\sigma'}(\mathbf{n}) \right. \\ & \quad \times \left. [\rho_{\sigma}^* c_{\ell,\sigma} v_{\ell,\sigma'} + \rho_{\sigma'}^* c_{\ell,\sigma'} v_{\ell,\sigma}] \right\}. \end{aligned} \quad (13)$$

The main challenge in implementing the  $\tau$ FE algorithm is to find the average rates from Eq. (9) for all  $r$ , and the second moments from Eq. (13) for any pair  $r, s$  of reactions affected by the environment.

We reiterate that the above considerations are for the regime  $\lambda\Delta t \gg 1$ , i.e., many environmental switches must occur in each time step  $\Delta t$ . It is this assumption that allows us to think of  $\bar{R}_r(\mathbf{n})$  as a Gaussian random variable. This in turn is a key component of the  $\tau$ FE algorithm. As a consequence the algorithm can only be expected to be accurate when  $\lambda$  is sufficiently large (fast environmental dynamics).

### C. Description of the algorithm

Without loss of generality, we assume that only the rates for the reactions  $r = 1, \dots, L$  ( $L \leq R$ ) depend on the environmental state  $\sigma$ .

The  $\tau$ FE algorithm with time step  $\Delta t$  proceeds as follows:

- (1) Initiate the population in state  $\mathbf{n}(0)$ . Set time to  $t = 0$ .
- (2) Compute  $R_r^*(\mathbf{n})$  for  $r = 1, \dots, L$  using Eq. (2), and the covariances  $\Xi_{rs}(\mathbf{n}) \equiv (\bar{R}_r(\mathbf{n})\bar{R}_s(\mathbf{n})) - R_r^*(\mathbf{n})R_s^*(\mathbf{n})$  using Eq. (13) for every pair  $r, s \in \{1, \dots, L\}$ .
- (3) (i) First consider the reactions with rates dependent on the environment: Draw correlated Gaussian random numbers  $\ell_1, \dots, \ell_L$  such that  $\langle \ell_r \rangle = R_r^*(\mathbf{n})$ , and  $\langle \ell_r \ell_s \rangle - \langle \ell_r \rangle \langle \ell_s \rangle = \Xi_{rs}(\mathbf{n})$ . If  $\ell_r < 0$  for any  $r \in \{1, \dots, L\}$  set, then  $\ell_r = 0$ .
- (4) For the remaining reactions  $r \in \{L+1, \dots, R\}$  set  $\ell_r = R_r(\mathbf{n})$ . These are the reactions with rates independent of the environment.
- (5) Draw independent Poissonian random numbers  $m_r$ ,  $r = 1, \dots, R$ , each with parameter  $\ell_r \Delta t$ .

(6) Update the state of the population,  $\mathbf{n}(t + \Delta t) = \mathbf{n}(t) + \sum_r m_r \mathbf{v}_r$ .

(7) Increment time by  $\Delta t$  and go to 2.

We note that the mean of the  $\ell_r$  in step 3(i) is of order  $(\lambda\Delta t)^0$ , and their variance of order  $(\lambda\Delta t)^{-1}$ . Truncation of the  $\ell_r$  will therefore only be required very rarely when  $\lambda\Delta t \gg 1$ .

Evaluating the expressions in Eqs. (2) and (13) in step 2 requires eigenvalues  $\mu_{\ell}$  of the transition matrix  $\mathbf{A}(\mathbf{n})$  for the environment, the eigenvectors,  $\mathbf{v}_{\ell}$  (including the stationary distribution  $\mathbf{v}_1 = \boldsymbol{\rho}^*$ ), and the coefficients  $c_{\ell,\sigma}$  for all  $\sigma$ . If the environmental process is independent of the state of the population (the  $A_{\sigma \rightarrow \sigma'}$  are not functions of  $\mathbf{n}$ ), then these quantities do not depend on  $\mathbf{n}$ , and only need to be calculated once at the beginning.

We stress that the  $\tau$ FE algorithm does not generate sample paths of the environmental process  $\sigma(t)$ . Instead the algorithm captures the statistics of  $\sigma$  in a Gaussian approximation for the transition rates  $\bar{R}_r(\mathbf{n})$ , as expressed above in step 3 (i). Using this approximation, the algorithm then propagates the state  $\mathbf{n}$  forward in time, without generating a sample path for  $\sigma$ .

In setting up the algorithm we perform an average over trajectories of the environmental process in each time step (retaining some fluctuations), assuming that the environment starts is drawn independently from its stationary distribution at the beginning of each step. In effect, the algorithm captures correlations of the environmental path within one time step, but it discards correlations of  $\sigma(t)$  from the beginning of one time step to the beginning of the next.

We also draw attention to the fact that the objects modelled as Gaussian random variables are the transition rates  $\bar{R}_r(\mathbf{n})$  (as opposed to individual parameters appearing in these rates). For example if a death rate depended on a carrying capacity  $K_{\sigma}$  (as in a Lotka-Volterra model), and was of the form  $R_{r,\sigma}(n) = n^2/K_{\sigma}$ , then our algorithm would not treat  $\bar{K}$  as a Gaussian random quantity, but instead the full rate  $\bar{R}_r(n)$ . Similarly, the adiabatic limit is obtained by averaging  $R_{r,\sigma}(n)$  over the stationary distribution of  $\sigma$ , and not  $K_{\sigma}$ . Models of this type are discussed for example in Refs. [11–13].

In Sec. IV we first test the  $\tau$ FE algorithm on different models with discrete environments. The algorithm can also be adapted to the case of environmental dynamics with continuous states. This is discussed in Sec. V.

## IV. APPLICATION OF THE $\tau$ FE ALGORITHM TO MODELS WITH DISCRETE ENVIRONMENTAL STATES

We now consider three examples of systems with discrete environmental states.

The first example (Sec. IV A) is a genetic circuit. The role of the environment is here played by a process of binding and unbinding to promoters of the genes described by the model. Gene regulatory systems can exhibit timescale separation as discussed for example in Refs. [32,42,43]. Mathematically the model describes a population with two types of individuals and an environment with two states (bound and unbound). The environmental dynamics depends on the state of the population.

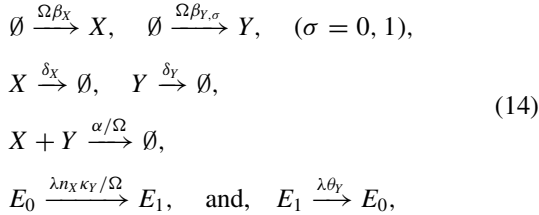
The second example (Sec. IV B) is a toy model with two species in the population and three environmental states. The

environmental process in this example is independent of the state of the population.

Section **IV C** finally focuses on a bimodal genetic switch with two species in the population, and an environmental process with three states, and with rates which depend on the state of the population.

### A. Genetic circuit: Two system-independent environments, two species

This system models the dynamics of two genes, which produce two different regulatory proteins:  $X$  (a transcription factor) and  $Y$  (an inhibitor that titrates  $X$  into an inactive complex). Specifically, we use the activator-titration circuit described in Ref. [32]. The reactions are as follows:



where the  $E_\sigma$  denote states of the environment ( $\sigma = 0, 1$ ). These two environmental states represent situations in which a transcription factor  $X$  is bound to the promoter of gene  $Y$  (state  $E_1$ ), or no transcription factor is bound ( $E_0$ ), respectively. The first two reactions in Eq. (14) describe the production of the two proteins ( $X$  and  $Y$ ). The production rates are  $\beta_X$  and  $\beta_{Y,\sigma}$ . The former is independent of the environmental state, the latter explicitly depends on  $\sigma$  (i.e., on the presence or absence of a bound transcription factor). The reactions in the second line of Eq. (14) describe degradation of  $X$  and  $Y$ , and the reaction in the third line captures titration. The binding and unbinding processes of the transcription factor are described by the reactions in the last line. The parameter  $\Omega$  in the reaction rates determines the typical number of particles in the system, for further details see [32]. We write  $n_X$  for the number of  $X$ -particles in the system, and similarly  $n_Y$  is the number of  $Y$ -particles. One finds  $n_X, n_Y = \mathcal{O}(\Omega)$  in the stationary state.

Mathematically, the model consists of two species in the population (with numbers of particles  $n_X, n_Y$ ), and two environmental states,  $\sigma = 0, 1$ . We therefore have  $S = 2, M = 2$ . Eqs. (9) and (13) can be evaluated explicitly for this case, see also Ref. [38]. The only process affected by the state of the environment is the production of  $Y$ , with rate  $\beta_{Y,\sigma}$ . This rate becomes a (clipped) Gaussian random variable in the  $\tau$ FE algorithm, with first moment

$$\langle \bar{\beta}_Y \rangle = \beta_Y^* = \frac{\theta_Y \beta_{Y,0} + n_X \kappa_Y \beta_{Y,1} / \Omega}{\theta_Y + n_X \kappa_Y / \Omega}, \tag{15}$$

and with variance

$$\begin{aligned}
 \sigma_{\beta_Y}^2 &\equiv \langle \bar{\beta}_Y^2 \rangle - \beta_Y^{*2} \\
 &= \frac{2n_X \kappa_Y \theta_Y / \Omega}{\lambda \Delta t (n_X \kappa_Y / \Omega + \theta_Y)} (\beta_{Y,0} - \beta_{Y,1})^2.
 \end{aligned} \tag{16}$$

Further details of the derivation can be found in Appendix B.

Simulation results for this model are shown in Fig. 1. In Figs. 1(a) and 1(b) we illustrate typical sample paths obtained

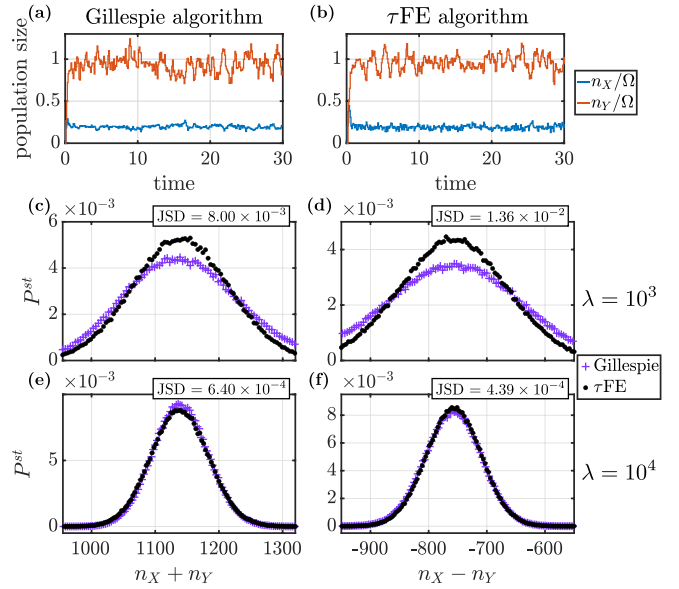


FIG. 1. Simulation output for the model of the genetic-circuit in Eq. (14). Panel (a) shows a sample path obtained from Gillespie simulations of the full model. Panel (b) is a sample path from the  $\tau$ FE algorithm [ $\lambda = 10^3$  in panels (a) and (b)]. Panels (c) and (d) show the stationary distributions of  $n_X + n_Y$  and  $n_X - n_Y$ , respectively, for  $\lambda = 10^3$ , while panels (e) and (f) are for  $\lambda = 10^4$ . In panels (c)–(f) we report the Jensen-Shannon divergence (JSD) between the distributions obtained using the two different simulation methods. Remaining parameters:  $\Omega = 10^3$ ,  $\beta_X = 2$ ,  $\beta_{Y,0} = 0$ ,  $\beta_{Y,1} = 10$ ,  $\delta_X = \delta_Y = 1$ ,  $\kappa_Y = 1$ ,  $\theta_Y = 0.5$ , and  $\alpha = 10$ . For the  $\tau$ FE we have used a time step  $\Delta t = 0.1$ .

from Gillespie simulations of the full model (population and environment), and from the  $\tau$ FE algorithm, respectively. We also show the stationary distributions for the quantities  $n_X + n_Y$  and  $n_X - n_Y$  as obtained from both simulation algorithms. The distributions in Figs. 1(c) and 1(d) are for  $\lambda = 10^3$  (i.e., moderately fast environmental dynamics), there are then remaining discrepancies between the  $\tau$ FE algorithm and simulations of the full model. In Figs. 1(e) and 1(f) the timescale separation is larger ( $\lambda = 10^4$ ). The agreement improves as indicated by the Jensen-Shannon divergence (JSD) [44,45] given in the figure.

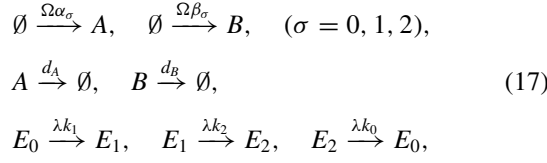
We note at this point that the average CPU time to run a sample path up to time  $t = 10^3$  with parameters as in Figs. 1(e) and 1(f) is 2.94 s for the Gillespie algorithm, and 0.03 s for the  $\tau$ FE algorithm (with a time step  $\Delta t = 0.1$ ). These average simulation times are obtained from ten runs. They indicate that the  $\tau$ FE algorithm can significantly increase efficiency while producing results of the quality shown in Fig. 1. We stress that our primary interest is the relative comparison of computing times, and not on absolute simulation times [46]. It is also important to keep in mind that, by construction, the  $\tau$ FE algorithm can only be expected to be accurate for sufficiently large values of  $\lambda$ , i.e., sufficiently fast environmental processes.

### B. Birth-death process: Three environments, two species

Next, we consider a two-species birth-death process subject to an external environment which can be in one of three

different states. This is a toy model chosen for illustration and does not represent any specific natural system. However, it captures elements of models of population dynamics.

The species in the population are labeled  $A$  and  $B$ , and the environmental states  $\sigma = 0, 1, 2$ . Particles of type  $A$  are produced with rate  $\Omega\alpha_\sigma$ , and particles of type  $B$  with rate  $\Omega\beta_\sigma$ . The subscript  $\sigma$  indicates explicit dependence on the state of the environment. Particles are removed with constant per capita rates  $d_A$  and  $d_B$ , respectively. The parameter  $\Omega$  again sets the typical size of the population. We write  $n_A$  and  $n_B$  for the number of individuals of either species. The environmental states cycle stochastically through the sequence  $\sigma = 0, 1, 2, 0, \dots$ , with rate constants  $\lambda k_1, \lambda k_2$  and  $\lambda k_0$  for the three transitions. Mathematically, the reactions in this model are



where as before  $E_\sigma$  denotes the environment. The rates  $k_0, k_1$ , and  $k_2$  are constant parameters, independent of the population state.

Details of the calculation of the average rates and their second moments can be found in Appendix C. The average production rates for the two types of particles are

$$\begin{aligned} \alpha^* &= \frac{k_0\alpha_0 + k_1\alpha_1 + k_2\alpha_2}{k_0 + k_1 + k_2}, \\ \beta^* &= \frac{k_0\beta_0 + k_1\beta_1 + k_2\beta_2}{k_0 + k_1 + k_2}, \end{aligned} \quad (18)$$

while the covariance  $\sigma_{\alpha\beta} = \sigma_{\beta\alpha} \equiv \langle \bar{\alpha}\bar{\beta} \rangle - \alpha^*\beta^*$  takes the form

$$\begin{aligned} \sigma_{\alpha\beta} &= \frac{\theta^2}{\lambda\Delta t} \left\{ (\alpha_0 - \alpha_1)(\beta_0 - \beta_1)(3k_0^2 - k_{0,1}k_{0,2}) \right. \\ &\quad + (\alpha_1 - \alpha_2)(\beta_1 - \beta_2)(3k_1^2 - k_{1,0}k_{1,2}) \\ &\quad \left. + (\alpha_0 - \alpha_2)(\beta_0 - \beta_2)(3k_2^2 - k_{2,0}k_{2,1}) \right\}, \end{aligned} \quad (19)$$

with

$$\theta^2 \equiv \frac{k_0k_1k_2}{(k_0k_1 + k_1k_2 + k_2k_0)^3}, \quad (20)$$

and  $k_{\sigma,\sigma'} = k_\sigma - k_{\sigma'}$ , for  $\sigma, \sigma' \in \{0, 1, 2\}$ . The variance  $\sigma_{\alpha\alpha} \equiv \langle \bar{\alpha}\bar{\alpha} \rangle - (\alpha^*)^2$ , is obtained by replacing all instances of  $\beta_\sigma$  on the right-hand side of Eq. (19) with  $\alpha_\sigma$ . The analog  $\sigma_{\beta\beta}$  is obtained similarly by replacing  $\alpha_\sigma$  with  $\beta_\sigma$ .

In Fig. 2 we report results from numerical simulations for this model, both from conventional Gillespie algorithm of the full systems of environment and population, and using the  $\tau$ FE algorithm. Figures 2(a)–2(f) show the stationary distributions of  $n_A + n_B$  and  $n_A - n_B$ . As seen from the data for example in Fig. 2(a) the  $\tau$ FE algorithm displays deviations from Gillespie simulations of the full model when the environmental process is not sufficiently fast. We quantify these deviations again through the Jensen–Shannon divergence between the two distributions. The deviations reduce as the timescale separation  $\lambda$  is increased, i.e., when the environ-

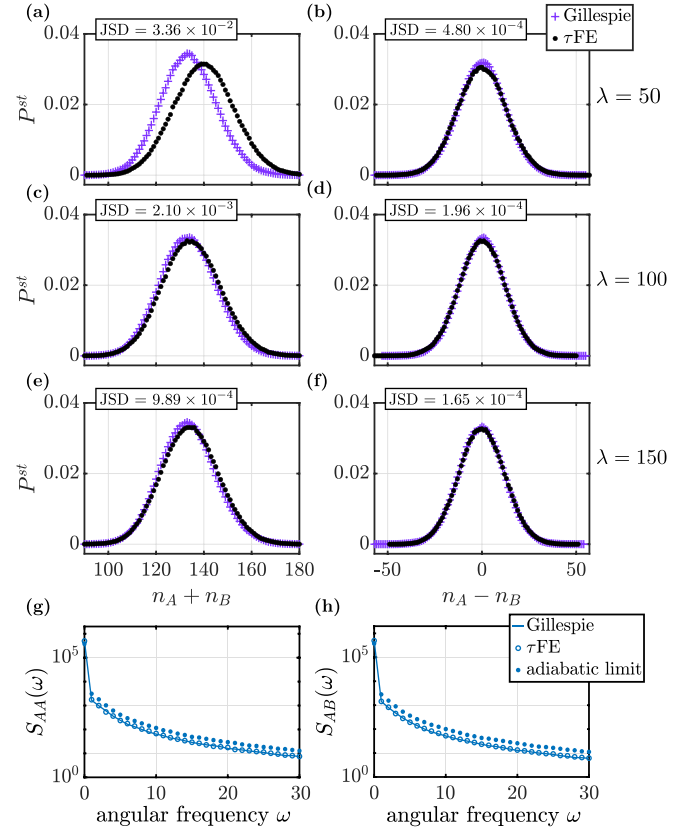


FIG. 2. Simulation output of the two-species birth-death process in an environment with three states [Eq. (17)]. Panels (a), (c), and (e) show the stationary distribution of  $n_A + n_B$  obtained using the Gillespie algorithm for the full model, and the  $\tau$ FE algorithm. Data is shown for different values of  $\lambda$ . Panels (b), (d), and (f) show the stationary distribution of  $n_A - n_B$ . Parameters used:  $k_0 = k_1 = k_2 = 1$ ,  $\Omega = 20$ ,  $d_A = d_B = 0.1$ ,  $\alpha_0 = \beta_0 = \beta_1 = \alpha_2 = 0$ , and  $\alpha_1 = \beta_2 = 1$ . In panels (a)–(f) we report the Jensen–Shannon divergence (JSD) between the two distributions obtained from Gillespie simulations of the full model and from the  $\tau$ FE algorithm. Panels (g) and (h): Spectral densities  $S_{AA}(\omega)$  and  $S_{AB}(\omega)$  [Eq. (21)] obtained from simulations using the Gillespie algorithm (full line), the  $\tau$ FE algorithm (open circles), and from conventional  $\tau$ -leaping simulations of the model in the adiabatic limit (asterisks). Parameters in panels (g) and (h) are as in panels (e) and (f), i.e.,  $\lambda = 150$ . The time step for the  $\tau$ FE algorithm and for conventional  $\tau$ -leaping in the adiabatic limit is  $\Delta t = 0.1$ .

mental process becomes faster relative to the dynamics within the population.

To examine if the  $\tau$ FE algorithm accurately reproduces dynamical features (i.e., properties of the system beyond the stationary distribution), we show spectral densities of the time series for  $n_A$  and  $n_B$  in Figs. 2(g) and 2(h). The spectral densities are defined as

$$\begin{aligned} S_{AA}(\omega) &= \langle |\hat{n}_A(\omega)|^2 \rangle, \\ S_{AB}(\omega) &= \langle \hat{n}_A^\dagger(\omega) \hat{n}_B(\omega) \rangle, \end{aligned} \quad (21)$$

where  $\hat{n}_A(\omega)$  and  $\hat{n}_B(\omega)$  are the Fourier transforms of  $n_A(t)$  and  $n_B(t)$ , respectively. The dagger denotes complex conjugation. The data from the  $\tau$ FE algorithm (open symbols) in

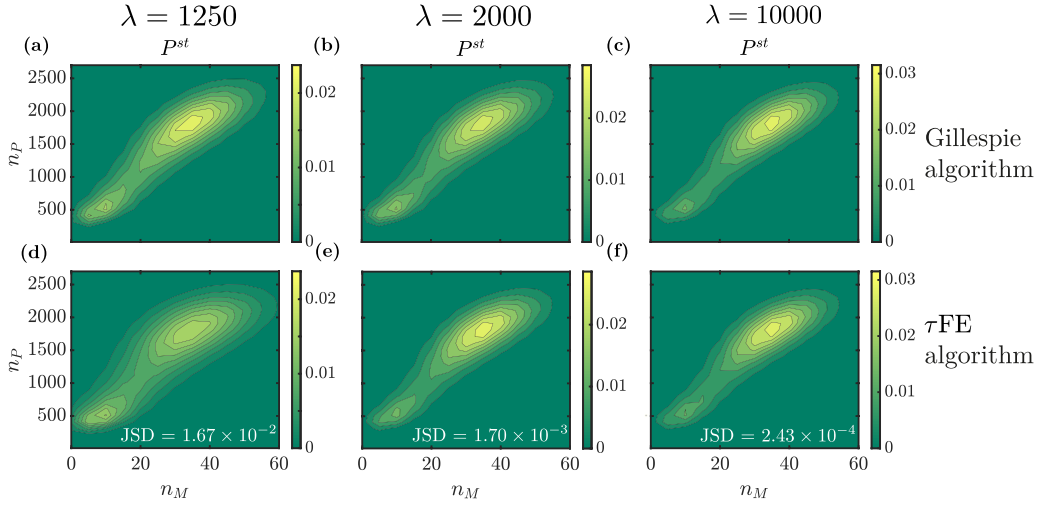
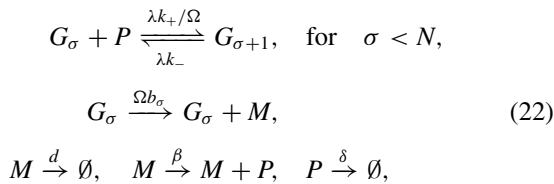


FIG. 3. Stationary distribution for the numbers of mRNA and protein molecules ( $n_M$  and  $n_P$ , respectively) in the model of a genetic switch [Eq. (17)]. Data is shown for different values of  $\lambda$  and for simulations of the full model by means of the Gillespie algorithm, and the  $\tau$ FE algorithm. Parameters:  $N = 2$ ,  $\Omega = 50$ ,  $b_0 = b_1 = 1$ ,  $b_2 = 20$ ,  $d = 9.2$ ,  $\beta = 50$ ,  $\delta = 1$ ,  $k_+ = 0.025$ ,  $k_- = 1$ . The stationary distribution is obtained from a long run up to time  $t = 10^5$ . For the  $\tau$ FE algorithm we use  $\Delta t = 100/\lambda$ . For each of the three values of  $\lambda$  we report the Jensen-Shannon divergence (JSD) between the distributions obtained from the two simulation methods.

Figs. 2(g) and 2(h) compares well with spectra obtained from direct Gillespie simulations of the full model (solid lines). This shows the  $\tau$ FE method indeed captures the dynamics of  $n_A$  and  $n_B$ . We also provide a comparison against the spectral densities obtained from conventional  $\tau$ -leaping simulations in the adiabatic limit, i.e., simulations with constant rates  $\alpha^*$  and  $\beta^*$  for the production events [Eq. (18)]. These are shown as full markers in Figs. 2(g) and 2(h). One then finds more substantial systematic deviations. This is because environmental fluctuations are discarded in the adiabatic limit. The  $\tau$ FE algorithm on the other hand captures the stochasticity of the environment to subleading order in  $\lambda^{-1}$  in each iteration step.

### C. Bimodal genetic switch: Three system-state dependent environments, two species

We now consider a model studied in Refs. [26,47], describing a single gene  $G$  with a promoter site which can bind to a total of up to  $N$  molecules of protein. The number of protein molecules bound,  $\sigma$ , plays the role of the environment in this setting. The rate for transitions from  $\sigma$  to  $\sigma + 1$  depends on the number of protein molecules. The reactions in this model can be summarized as follows:



where  $M$  and  $P$  refer to molecules of mRNA and protein, respectively. The production rate  $b_\sigma$  for mRNA depends on the number of protein molecules bound to the promoter. We refer to Refs. [26,47] for further details. In the following we write  $n_M$  and  $n_P$  for the numbers of particles of either type. One interesting feature of this model is that the distribution of

the protein and mRNA populations can become bimodal, as illustrated in Fig. 3. This leads to bistability, with trajectories transitioning between the two modes of the joint distribution of  $n_P$  and  $n_M$ . Hence, the model describes a genetic switch.

In this model only the production rate of mRNA molecules is affected by the state of the environment. The average mRNA-production rate is found as

$$b^* = \frac{k_-^2 b_0 + k_- \tilde{k}_+ b_1 + \tilde{k}_+^2 b_2}{k_-^2 + k_- \tilde{k}_+ + \tilde{k}_+^2}, \tag{23}$$

with  $\tilde{k}_+ = k_+ n_P / \Omega$ . The second moment of the production rate takes the form

$$\begin{aligned}
 \sigma_{bb}^2 &\equiv \langle \bar{b}^2 \rangle - b^{*2} \\
 &= \frac{\theta^2}{\lambda \Delta t} \{ (b_0 - b_1)^2 k_- (k_-^2 + \tilde{k}_+ k_- - \tilde{k}_+^2) \\
 &\quad + (b_0 - b_2)^2 2k_- \tilde{k}_+ (k_- + \tilde{k}_+) \\
 &\quad + (b_1 - b_2)^2 \tilde{k}_+ (\tilde{k}_+^2 + \tilde{k}_+ k_- - k_-^2) \},
 \end{aligned} \tag{24}$$

with

$$\theta^2 = \frac{2k_- \tilde{k}_+}{(k_-^2 + k_- \tilde{k}_+ + \tilde{k}_+^2)^3}. \tag{25}$$

Details of the calculation leading to Eqs. (23) and (24) can be found in Appendix D.

Figure 3 shows the stationary joint distribution of the number of mRNA and protein molecules for different values of the timescale separation parameter  $\lambda$ . The figure shows data from Gillespie simulations of the full model [Figs. 3(a)–3(c)], and data from the  $\tau$ FE algorithm [Figs. 3(d)–3(f)]. The  $\tau$ FE algorithm captures the distribution profile with two local maxima. For low values of  $\lambda$  (i.e., a relatively slow environmental process) the distribution obtained from  $\tau$ FE tends to be wider than those from the Gillespie algorithm. The agreement



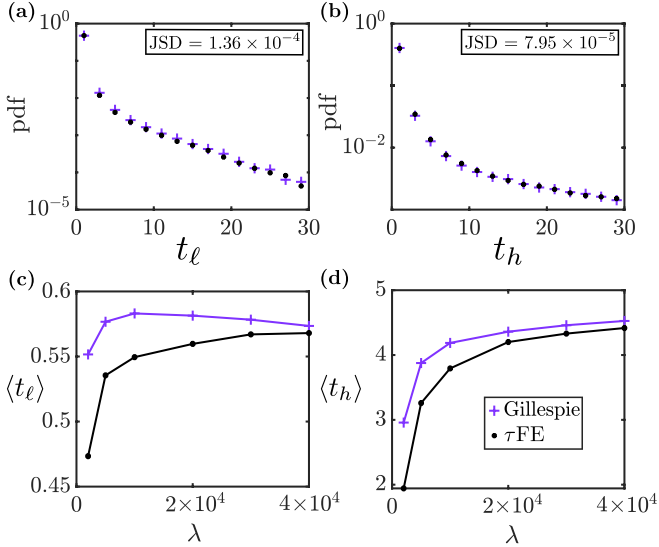


FIG. 4. Sojourn times  $t_\ell$  and  $t_h$  near the two modes of the bistable genetic switch (see also Fig. 3). Panels (a) and (b) show the distribution of the time spent in the vicinity of each mode (see text for details); data obtained from  $\tau$ FE algorithm is shown along with results from exact Gillespie simulations of the full model ( $\lambda = 2000$ ). In each panel we report the Jensen–Shannon distance between the two distributions. Panels (c) and (d) show the mean sojourn times as a function of the timescale parameter  $\lambda$ . The parameters are as in Fig. 3, the lower mode is  $(n_M, n_P) = (10, 500)$ , and the upper mode  $(n_M, n_P) = (30, 1800)$ . For the  $\tau$ FE algorithm we use a time step of  $\Delta t = 100/\lambda$ .

improves for faster environments, as indicated again by the Jensen–Shannon distances in Fig. 3.

In Fig. 4, we show the distribution and means of the sojourn times  $t_\ell$  and  $t_h$  near the lower and higher modes of the stationary distribution. More precisely this is the time between entering and leaving a designated region around each of the modes. The lower maximum of the stationary distribution is sharper than the upper maximum (Fig. 3). Accordingly, we have chosen a smaller region at the lower mode than at the upper mode. For the lower mode, we use the region  $0 \leq n_M \leq 20$ ,  $0 \leq n_P \leq 1100$  which encloses the mode at  $(n_M, n_P) = (10, 500)$ . For the higher mode we use the region  $20 \leq n_M \leq 80$ ,  $1100 \leq n_P \leq 2700$  enclosing the mode at  $(n_M, n_P) = (30, 1800)$ .

The data shown in the figure is constructed from one long sample path (run until  $t = 10^6$ ), recording the points in time at which the system enters or leaves either region. Gillespie simulations operate in continuous time and the  $\tau$ FE algorithm in discrete time. To remove any artefacts resulting from this difference, the same time resolution (0.05) is used in both algorithms for the measurement of arrival and departure times. Because the lower mode is sharper than the upper maximum and because the sizes of the two detection regions are different the sojourn time  $t_\ell$  at the lower mode is found to be smaller than that at the higher mode,  $t_h$ .

The distributions of sojourn times in Figs. 4(a) and 4(b) indicate that the  $\tau$ FE algorithm captures this dynamic quantity, provided the environmental process is sufficiently fast. This is confirmed in Figs. 4(c) and 4(d), where we show the

TABLE I. Mean computation time (in seconds) required to simulate one sample path up to  $t = 10^3$  of the bimodal genetic-switch system defined in Eq. (22). Measurements are from ten independent sample runs, using Gillespie simulations of the full model, and the  $\tau$ FE algorithm, respectively. Parameters are as in Fig. 3. For the  $\tau$ FE algorithm we set  $\Delta t = 100/\lambda$ .

$\lambda$	Gillespie	$\tau$ FE
1250	1.35	0.04
2500	1.89	0.08
5000	2.62	0.17
10 000	4.30	0.31
20 000	7.77	0.57

mean sojourn times as a function of the relative speed  $\lambda$  of the environment compared to the population dynamics. As seen in both panels, the  $\tau$ FE algorithm generates accurate measurements of the mean sojourn times  $\langle t_\ell \rangle$  and  $\langle t_h \rangle$  in the limit  $\lambda \gg 1$ .

At the same time, stochastic effects due to the random environmental process are captured for large but finite  $\lambda$ . This can be seen in Fig. 4(d): the mean sojourn time  $\langle t_h \rangle$  drops significantly as the environmental process becomes slower, and hence additional noise is injected into the population (there is no environmental noise in the adiabatic limit). While there are quantitative differences compared to exact simulations, the  $\tau$ FE algorithm captures this reduction of  $\langle t_h \rangle$ . Figure 4(c) reveals that there are also limitations to the precision of the  $\tau$ FE algorithm. The mean sojourn time  $\langle t_\ell \rangle$  near the lower mode is affected much less by a reduction of the timescale separation parameter  $\lambda$  than the mean sojourn time at the upper mode. This indicates that the escape from this region is driven mostly by intrinsic noise rather than by environmental stochasticity. While the data from the two algorithms remains within approximately 10% for sufficiently fast environmental dynamics ( $\lambda \gtrsim 10^4$ ) the  $\tau$ FE algorithm is unable to capture the small rise of  $\langle t_\ell \rangle$  observed in Gillespie simulations for intermediate values of  $\lambda$ .

In Table I we compare the the computing time required for both the Gillespie algorithm and the  $\tau$ FE method for different values of  $\lambda$ . The data in the table is the CPU time required to generate one sample path up to time  $t = 10^3$ , averaged over ten runs. The model parameters are as in Figs. 3 and 4.

The full model comprises the reactions in the population and the environmental switching. The rates for the former reactions are independent of  $\lambda$ , the rates for the latter scale linearly in  $\lambda$ . Accordingly, one expects the computing time for Gillespie simulations of the full model to be linear in  $\lambda$ , with a nonzero intercept. The data in the table is consistent with this. We note that Gillespie algorithm does not require any time discretization.

The running time for the  $\tau$ FE algorithm depends on the choice of the time step. The time step in turn affects the accuracy of the outcome. If  $\Delta t$  is large, then  $\tau$ FE simulations are fast, but the approximation to the continuous-time full model becomes less good. This is because the “leap condition” must be met [37], for the models studied here effectively dictating that  $\Delta t$  must be (much) smaller than unity. On the other hand,

the time step must not be too small as the construction of the  $\tau$ FE algorithm requires sufficient averaging of the environmental process in each step [Eqs. (11)–(13)].

Due to these constraints, the choice of time step for the  $\tau$ FE algorithm is not a trivial matter. Broadly speaking, we calibrated the time step by requiring that the outcome of  $\tau$ FE simulations reproduces results from full Gillespie simulations in a regime in which  $\lambda$  is sufficiently large that we can expect the fast-switching approximation to be valid, but where  $\lambda$  is not so large that Gillespie simulations become unrealistic. Once this agreement has been checked, the choice of time step can then be extrapolated to regimes of even faster environments in which full Gillespie simulations are no longer possible within reasonably computing time.

Following this idea, the time step for the  $\tau$ FE algorithm in Figs. 3 and 4, and in Table I is chosen inversely proportional to  $\lambda$  ( $\Delta t = 100/\lambda$ ). This choice ensures that each time step captures a sufficient number of switches of the environmental state, and for sufficiently large  $\lambda$  the leap condition is also met. For  $\lambda \lesssim 1000$ , however, the leap condition is no longer met for this choice. This is the reason for the choice  $\Delta t = 0.1$  in Fig. 2. We make the same choice in Fig. 1, but we have checked that this latter figure does not materially change if  $\Delta t = 100/\lambda$ .

Accordingly, we expect the computing time for the  $\tau$ FE algorithm to scale linearly in  $\lambda$ , with no intercept. Again, the running times we measured in our simulations are consistent with this expectation. Overall, Table I shows that the  $\tau$ FE

algorithm is able to generate data of the accuracy as in Figs. 3 and 4 while reducing the computing effort approximately tenfold compared to full Gillespie simulations.

## V. NUMERICAL SIMULATION OF CONTINUOUS-ENVIRONMENTAL SYSTEMS

### A. Setup

We turn now to systems which are subject to an environment with continuous states. Specifically, we follow Ref. [20] and assume that the environmental state  $\sigma$  follows an Ornstein–Uhlenbeck process (see also Refs. [22,48]),

$$\frac{d\sigma}{dt} = \lambda(m - \sigma) + \sqrt{2\lambda v^2} \eta(t), \quad (26)$$

where  $\eta(t)$  is Gaussian white noise of unit amplitude, in particular,  $\langle \eta(t)\eta(t') \rangle = \delta(t - t')$ . The parameter  $m$  is the average value of  $\sigma$  in the long run, while  $v$  controls the magnitude of noise. As before, the parameter  $\lambda > 0$  indicates how quickly the environment changes relative to the dynamics in the population;  $\lambda$  is the equivalent of  $1/\tau_c$  in the notation of Ref. [20].

The probability distribution of finding the environment in state  $\sigma$  at time  $t$ , given that was in state  $\sigma'$  at time  $t'$ , can be obtained from the Fokker–Planck equation for the Ornstein–Uhlenbeck process, and is given by (see, e.g., Refs. [49,50])

$$q_{\sigma' \rightarrow \sigma}(t - t') = \sqrt{\frac{1}{2\pi v^2(1 - e^{-2\lambda(t-t')})}} \exp\left\{-\frac{[\sigma - \sigma' e^{-\lambda(t-t')} - m(1 - e^{-\lambda(t-t')})]^2}{2v^2(1 - e^{-2\lambda(t-t')})}\right\}. \quad (27)$$

For  $t \rightarrow \infty$  (and  $t'$  fixed) this quantity tends to the stationary distribution

$$\rho_{\sigma}^* = \sqrt{\frac{1}{2\pi v^2}} \exp\left[-\frac{(\sigma - m)^2}{2v^2}\right]. \quad (28)$$

We note that it is not a requirement for the  $\tau$ FE algorithm that the environment follows an Ornstein–Uhlenbeck process. However, both functions  $q_{\sigma' \rightarrow \sigma}(t - t')$  and  $\rho_{\sigma}^*$  are required, as discussed in more detail below.

We proceed to describe how the  $\tau$ FE algorithm can be implemented for models with continuous environments (Sec. VB).

In the case of discrete environments, continuous-time sample paths of the full model can be generated using the conventional Gillespie algorithm. This is an exact procedure: the ensemble of these sample paths faithfully describes the statistics of the full model. In Sec. IV we have used this as a benchmark to test the  $\tau$ FE algorithm. We are not aware of any analogous exact simulation method for models of discrete populations in a stochastic environment with continuous states. To test the  $\tau$ FE algorithm we therefore compare outcomes against those from approximation methods to generate paths of the combined set of the population and the environment. Several such methods exist, we describe these in

Sec. VC. The tests of the  $\tau$ FE algorithm against the baseline of these methods are described in Sec. VI.

### B. Implementation of the $\tau$ FE algorithm for continuous environments

We proceed similar to discrete case in Sec. III, replacing the sums over  $\sigma$  in Eqs. (2) and (10) with integrals. We then have

$$R_r^*(\mathbf{n}) = \int_{-\infty}^{\infty} d\sigma \rho_{\sigma}^* R_{r,\sigma}(\mathbf{n}), \quad (29)$$

and the relation for the second moments turns into

$$\begin{aligned} \langle \bar{R}_r(\mathbf{n}) \bar{R}_s(\mathbf{n}) \rangle &= \frac{1}{\Delta t^2} \int_{-\infty}^{\infty} d\sigma \int_{-\infty}^{\infty} d\sigma' \int_{t_1}^{t_1+\Delta t} dt_1 \\ &\times \int_{t_1}^{t_1+\Delta t} dt_2 i\{\rho_{\sigma}^* q_{\sigma \rightarrow \sigma'}(t_2 - t_1) \\ &\times [R_{r,\sigma}(\mathbf{n}) R_{s,\sigma'}(\mathbf{n}) + R_{r,\sigma'}(\mathbf{n}) R_{s,\sigma}(\mathbf{n})]\}. \end{aligned} \quad (30)$$

Depending on the form of the stationary distribution  $\rho_{\sigma}^*$ , the kernel  $q_{\sigma \rightarrow \sigma'}(t_2 - t_1)$  and the rates  $R_{r,\sigma}(\mathbf{n})$  the integrals in Eqs. (29) and (30) can be carried out, and closed-form analytical expressions can be obtained. In Sec. VI we explore

a number of different examples, further scenarios are also discussed Appendix F. Once the average rates and the second moments are calculated, the  $\tau$ FE algorithm is implemented as described in Sec. III C.

### C. Conventional simulation approaches for discrete populations in continuous environments

In this section we summarize “conventional” approaches to simulating discrete Markovian systems subject to environmental dynamics with continuous states. By “conventional” we mean methods which produce explicit (approximate) sample paths of the environmental process. This is in contrast to the  $\tau$ FE algorithm, which generates paths only of the system proper.

#### 1. Gillespie algorithm with discretized environmental states (GADE)

This approach is based on a discretization of the space of environmental states, time remains continuous. Once such a discretization for the environmental states is carried out, the combined states of the population and environment are also discrete. Simulations can be carried out using the conventional Gillespie method. We will refer to this method as GADE (Gillespie approach with discretized environment).

The key step in this approach is to find an appropriate dynamics in the space of discretized environmental states. We describe this in the context of the Ornstein–Uhlenbeck process in Eq. (26). We discretize the environmental state into integer multiples of  $\Delta\sigma$ , i.e., the environment takes states  $\dots, -2\Delta\sigma, -\Delta\sigma, 0, \Delta\sigma, 2\Delta\sigma, \dots$ . Transitions from one state  $k\Delta\sigma$  can only occur to states  $(k \pm 1)\Delta\sigma$ . The transition rates are constructed such that this discrete process recovers the continuous Ornstein–Uhlenbeck dynamics in the limit  $\Delta\sigma \rightarrow 0$ . The details of the construction are described in Appendix E, we here only report the main outcome. Specifically, the rates to transition from state  $k\Delta\sigma$  to  $(k \pm 1)\Delta\sigma$  can be chosen as

$$T_k^\pm = \frac{\lambda}{2\Delta\sigma} \left[ \pm(m - k\Delta\sigma) + \frac{2v^2}{\Delta\sigma} \right]. \quad (31)$$

This process can then be simulated using the standard Gillespie algorithm, along with the events in the population. We note that the rates  $T_k^\pm$  need to be nonnegative, i.e., we require  $|m - k\Delta\sigma| < 2v^2/\Delta\sigma$ , for all  $k$ . In practice, this can be achieved by truncating the set of possible states  $k\Delta\sigma$ . More precisely, we disallow transitions out of the region  $\{k : |m - k\Delta\sigma| \leq K\}$ , with a given cutoff  $K$ . Provided that  $K$  is sufficiently large truncations will only be required rarely. Once a cutoff  $K$  is chosen we must require  $\Delta\sigma \leq 2v^2/K$  to guarantee nonnegativity of the  $T_k^\pm$ . The variance of the Ornstein–Uhlenbeck process for  $\sigma$  is given by  $v^2$  in the long run [Eq. (28)], so  $K \propto v$  is a sensible choice. This results in maximum value for  $\Delta\sigma$  which is also proportional to  $v$ .

#### 2. Discrete-time simulation with explicit environmental dynamics (DEED)

Approximate sample paths of the combined system of population and environment can also be generated in a

discrete-time simulation. We refer to this as DEED (discrete-time simulation with explicit environmental dynamics). The time step  $\Delta t$  needs to be sufficiently small to capture the details of the environmental process with characteristic timescale  $\tau_c = \lambda^{-1}$ . We therefore require  $\Delta t \lesssim \lambda^{-1}$ . One possible implementation is as follows:

(1) Suppose we have arrived at time  $t$ , and the state of the population is  $\mathbf{n}(t)$  and that of the environment  $\sigma(t)$ . Obtain  $\sigma(t + \Delta t)$  from Eq. (26) using the Euler–Maruyama method [51].

(2) Use  $\sigma(t)$  and  $\mathbf{n}(t)$  to calculate the rates  $p_r(t) = \Delta t \times R_{r,\sigma(t)}[\mathbf{n}(t)]$  for  $r = 1, \dots, R$ .

(3) Provided  $\Delta t$  is small enough, the  $p_r(t)$  are all less than one. To lowest order in  $\Delta t$  they are the probabilities that a reaction of type  $r$  occurs in the next  $\Delta t$ . For each  $r = 1, \dots, R$  implement one reaction of this type with probability  $p_r(t)$ . With probability  $1 - p_r(t)$  no reaction of type  $r$  occurs. Executing all reactions that fire, one obtains  $\mathbf{n}(t + \Delta t)$ .

(4) Increment time by  $\Delta t$ , and go to step 1.

Step 3 disregards the possibility that a particular reaction fires multiple times during one time step. This is a valid approximation, provided that the  $p_r(t) = \Delta t \times R_{r,\sigma(t)}$  are much smaller than one. As an alternative step 3 could be replaced by a conventional  $\tau$ -leaping step. The number of reactions of type  $r$  that fire is then a Poissonian random variable with parameter  $p_r(t)$ .

#### 3. Thinning algorithm by Lewis

A population subject to a dynamic external environment with continuous state space can also be simulated using the so-called thinning algorithm by Lewis [52]. This algorithm generates a statistically faithful ensemble of sample paths for Markovian systems with discrete states and transition rates with explicit external time dependence.

In the context of our model the population is such a system. If the environmental dynamics is independent of the population, then realizations  $\sigma(t)$  for the environment can be generated in advance independently from the population. For instance, sample solutions of the Ornstein–Uhlenbeck process in Eq. (26) could be generated. Each such realization  $\sigma(t)$  then determines a realization of time-dependent rates  $R_r(\mathbf{n}, t) \equiv R_{r,\sigma(t)}(\mathbf{n})$  for the population. The Lewis algorithm can then be used to produce sample paths for the population dynamics.

In practice, numerical approximation schemes are required to generate realizations for the environment. For example, Eq. (26) can be solved numerically using the Euler–Maruyama method, with time step  $\Delta t$ . As discussed above this time step needs to be sufficiently small ( $\Delta t \lesssim \lambda^{-1}$ ) to resolve the short-time features of the environmental process. The Lewis algorithm then uses this as an input and generates sample paths for the population in continuous time.

We stress that none of these methods are exact. For the GADE algorithm the environmental state is discretized, DEED operates in discretized time, and the thinning algorithm requires paths of the environmental process, which are generated by discretizing the corresponding stochastic differential equation for the environmental state. As discussed in Sec. V A there is therefore no exact benchmark to calibrate the  $\tau$ FE algorithm against. Among the approaches listed above,

we think GADE algorithm is probably most suitable as a benchmark, as it does not introduce any discretization of time. Comparison against the outcome of GADE simulations can therefore be useful to confirm the validity of the  $\tau$ FE approach for genuinely new problems, and the choice of time step. The  $\tau$ FE algorithm (which we will demonstrate to be computationally more efficient for fast environments) can then be used in parameter regimes which are not easily accessible within reasonably computing time by the other methods.

## VI. APPLICATION OF THE $\tau$ FE TO CONTINUOUS-ENVIRONMENTAL MODELS

In this section we test the  $\tau$ FE algorithm on a number of different examples of models with continuous environmental states. Simulation outcomes are compared against those from the algorithms described in Sec. V C.

$$\langle \bar{R}_b(n)\bar{R}_d(n) \rangle - R_b^*(n)R_d^*(n) = \frac{\beta\delta v^2 e^{-2\lambda\Delta t}}{\lambda^2 \Delta t^2} \{8m^2 e^{\lambda\Delta t} + e^{2\lambda\Delta t} [8m^2(\lambda\Delta t - 1) + v^2(2\lambda\Delta t - 1)] + v^2\}. \quad (33)$$

for the covariance. The expressions for the variances are similar, with suitable replacements  $\beta\delta \rightarrow \beta^2$  and  $\beta\delta \rightarrow \delta^2$  in the prefactor in Eq. (33). This covariance matrix and the means in Eq. (32) are then used in the  $\tau$ FE algorithm.

Figure 5 shows simulation results from the  $\tau$ FE algorithm, as well as from the GADE and DEED schemes (Secs. V C 1 and V C 2, respectively). Figure 5(a) shows that all simulation methods result in linear growth (parameters are such that  $\beta > \delta$ , i.e., the growth rate is always larger than the death rate). Figure 5(b) confirms that GADE and DEED both generate the correct statistics for the stationary distribution of the environmental process [the solid line is the Gaussian distribution in Eq. (28)]. In Fig. 5(c) we focus on a fixed time  $t = 10$ , and show that all three simulation methods results in very similar distributions for the number of individuals in the population  $n$  at that time. Figure 5(d) finally shows a dynamic quantity, the Fourier spectrum  $S(\omega)$  of the time series  $n(t)$ , or equivalently the Fourier transform of the correlation function of  $n$ . Again, all three simulation methods produce very similar results.

In Table II we compare the average computing time required by the different algorithms to generate a trajectory up to time  $t = 10^3$ . We show data for varying values of the typical timescale  $\lambda^{-1}$  of the environmental process. GADE does not require any discretization of time. For the DEED approach we use  $\Delta t = 1/(100\lambda)$ . For the  $\tau$ FE method we choose  $\Delta t = 10/\lambda$ . This is in-line with the requirements  $\Delta t \lesssim \lambda^{-1}$  for DEED, and  $\Delta t \gtrsim \lambda^{-1}$  for  $\tau$ FE. The choice of time steps will be discussed in further detail below.

The data in the table indicates that the simulation timescales approximately linearly with  $\lambda$  for all three algorithms tested, provided  $\lambda$  is sufficiently large. This is to be expected: The rates for the environmental events in the GADE simulations (Sec. V C 1) scale as  $\lambda$ , and therefore dominate the events in the population for  $\lambda \gg 1$ . Each typical Gillespie

### A. Toy model: Population dynamics with production and removal rates proportional to $\sigma^2$

We first consider a production-removal process for a single species. The environmental state  $\sigma(t)$  follows the Ornstein–Uhlenbeck process in Eq. (26). The corresponding transition kernel  $q_{\sigma \rightarrow \sigma'}(\tau)$  is given in Eq. (27), and the stationary distribution  $\rho_\sigma^*$  in Eq. (28). The production rate in the population is assumed to be  $R_{b,\sigma} = \beta\sigma^2$ , and the removal rate  $R_{d,\sigma} = \delta\sigma^2$ . These are not chosen with any particular natural system in mind, instead this example serves as an illustration (see also Appendix F for similar calculations for two related examples).

From Eq. (29) we obtain

$$\begin{aligned} R_b^* &= \beta(m^2 + v^2), \\ R_d^* &= \delta(m^2 + v^2). \end{aligned} \quad (32)$$

The second moments of the rates  $\bar{R}_b(n)$  and  $\bar{R}_d(n)$  can be calculated from Eq. (30). We find

step then advances time by an amount proportional to  $\lambda^{-1}$ , and  $\mathcal{O}(\lambda)$  such steps are required to reach the designated end time. A similar argument applies to the DEED algorithm (Sec. V C 2) and for the  $\tau$ FE algorithm: For both of these we use time steps  $\Delta t \propto \lambda^{-1}$ , so again the number of iteration steps required scales as  $\lambda$ .

The key message from Table II is that, for the choice of time steps made in the table, the computing time required by the  $\tau$ FE algorithm is substantially lower than that for the other two simulation methods. Given the linear dependence on  $\lambda$ , this increase in efficiency can be extrapolated to environments operating on timescales faster than the smallest timescale shown in the table (i.e., to the range  $\lambda > 10^4$ ). We note that, due to the smaller time step, DEED produces a finer resolution of sample paths in time than  $\tau$ FE. When we make our comparison we have average macroscopic quantities in mind (such as those in Fig. 5), and not necessarily the generation of individual paths with the highest possible resolution in time.

TABLE II. Mean computing time (in seconds) required for one simulation run of the model described in Sec. VI A until  $t = 10^3$ . Data is from ten independent runs, parameters are as in Figure 5, i.e.,  $\beta = 1.1$ ,  $\delta = 1.0$ ,  $m = 1$ , and  $v^2 = 5 \times 10^{-4}$ . For GADE we set  $\Delta\sigma = 10^{-3}$ ; for the DEED approach we set  $\lambda\Delta t = 1/100$ ; for the  $\tau$ FE algorithm we set  $\lambda\Delta t = 10$ .

$\lambda^{-1}$	GADE	DEED	$\tau$ FE
$1 \times 10^{-2}$	28.47	3.84	$0.79 \times 10^{-2}$
$5 \times 10^{-3}$	53.20	7.88	$0.16 \times 10^{-1}$
$1 \times 10^{-3}$	288.30	40.91	0.08
$5 \times 10^{-4}$	576.69	82.78	0.15
$1 \times 10^{-4}$	3022.47	397.63	0.79



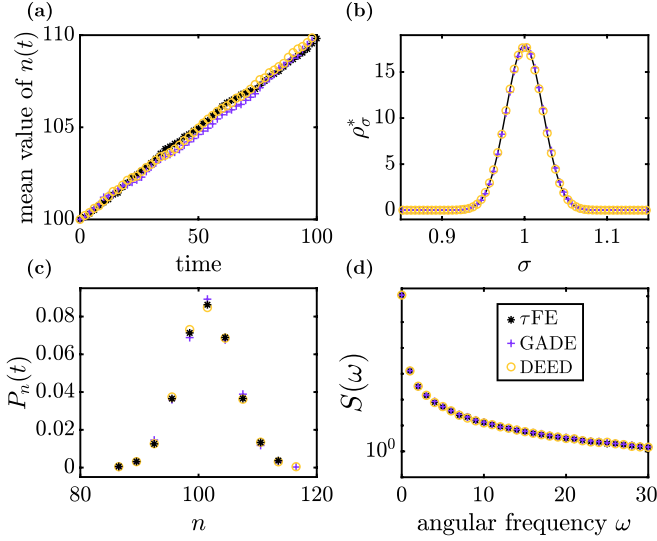


FIG. 5. Simulation results for a production-removal process with rates  $b = \beta\sigma^2$  and  $d = \delta\sigma^2$  (Sec. VI A), for the different algorithms described in Sec. V. Parameters used:  $\beta = 1.1$  and  $\delta = 1.0$ ,  $m = 1$ ,  $\lambda = 10^3$ , and  $v^2 = 5 \times 10^{-4}$ . Panel (a): mean value of the number of individuals as function of time, obtained from  $10^3$  runs. Panel (b): stationary distribution of the environmental state,  $\rho_\sigma^*$ , for the GADE method and the DEED approach (the  $\tau$ FE algorithm does not simulate the environment). The solid line in panel (b) is the analytical solution from Eq. (28). Panel (c): distribution of the number of particles  $n$  in the population at time  $t = 10$ . Panel (d): spectral density [Eq. (21)] obtained from  $10^3$  runs. We use  $\Delta\sigma = 10^{-3}$  for the GADE simulations, and  $\Delta t = 1/(100\lambda) = 10^{-5}$  for DEED. For the  $\tau$ FE algorithm, we set  $\Delta t = 10/\lambda = 10^{-2}$ .

We now briefly discuss the choice of time steps for the  $\tau$ FE method and for DEED. In principle, we could have increased or decreased the step for either method. This would then reduce or increase the computing time required to reach the designated end point. It might also affect the accuracy of the outcome. Our choice of  $\Delta t = 10/\lambda$  for  $\tau$ FE is motivated by the good agreement with GADE in Fig. 5, noting that GADE does not require any discretization of time. Similarly, for the example discussed below in Sec. VI B good agreement with analytical predictions is found for this choice, see the regime of small  $\lambda^{-1}$  in Fig. 6. Our conclusion is therefore that the  $\tau$ FE algorithm is able to produce results of the accuracy as in Fig. 5 with computing times as reported in Table II.

The DEED algorithm requires  $\Delta t \lesssim \lambda^{-1}$  to be able to resolve the environmental dynamics. Our choice  $\Delta t = 1/(100\lambda)$  in Table II is well below this requirement, and the algorithm can in principle be speed up by choosing a larger time step. If we were to exhaust the limit and used  $\Delta t = \lambda^{-1}$  for DEED, then this would reduce the computing time by about a factor of one hundred in Table II. For  $\lambda^{-1} = 10^{-3}$  this would mean a reduction from approximately 40 to 0.4 s per sample path. Using this larger time step also results in noticeable deviations in measurements of the quantities in Fig. 5 from continuous-time GADE simulations. But even if we accept this and use the hundredfold larger time step for DEED, the  $\tau$ FE algorithm would remain approximately five

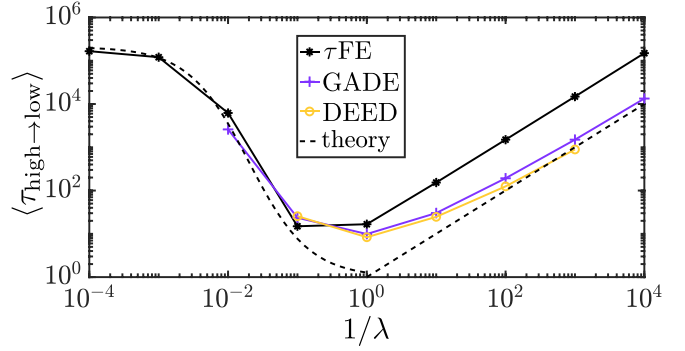


FIG. 6. Mean switching time (MST) from the high to the low state in the model described in Sec. VI B. For the  $\tau$ FE algorithm we used  $\Delta t = 10/\lambda$ , for GADE  $\Delta\sigma = 0.01$ , and for DEED  $\Delta t = 10^{-4}$ . Theory curves are from Eqs. (8) and (10) in Ref. [20]. Model parameters are as in the top right panel of Fig. 2 in Ref. [20] ( $\Omega = 5000$ ,  $v = 0.1$ ,  $\alpha_0 = 0.01$ ,  $x_0 = 0.93$ ).

times faster, requiring 0.08 s per sample path at  $\lambda^{-1} = 10^{-3}$ ; see Table II.

We have also conducted tests with Lewis' thinning algorithm. To do this we have first generated sample paths of the Ornstein–Uhlenbeck process for the environment [Eq. (26)] using an Euler–Maruyama scheme. This is then fed into the Lewis' algorithm for systems with time dependent rates. Given that the typical timescale of the environment is  $\lambda^{-1}$ , the largest sensible time step for the Euler–Maruyama scheme is  $\Delta t = \lambda^{-1}$ , similar to DEED. This choice minimises the computing time for the Lewis' approach. We therefore use this time step to compare the efficiency of the Lewis' approach with that of  $\tau$ FE. We find that the thinning algorithm is considerably slower than the  $\tau$ FE approach. For  $\lambda^{-1} = 10^{-3}$ , for example, we obtain a simulation time of approximately 13 s per run up to  $t = 10^3$  compared to 0.08 s for  $\tau$ FE (see Table II).

## B. Genetic switch with Hill-like regulatory function

As a final example we consider a model of protein production subject to a continuous environment discussed in Ref. [20]. The model entails positive feedback, in that the presence of protein has the potential to increase production of protein. There is one single species in the model (protein), we write the number of protein molecules as  $n$ . We also define  $x = n/\Omega$ , where  $\Omega$  is again a model parameter setting the typical size of the system. The production rate of protein is given by

$$f(x, \sigma) = \alpha_0 + (1 - \alpha_0 + \sigma)\Theta(x - x_0), \quad (34)$$

where  $0 < \alpha_0 < 1$  and  $x_0 > 0$  are constants, and where  $\Theta(x)$  is the Heaviside function. Protein molecules also decay with unit rate. In the absence of environmental influence ( $\sigma \equiv 0$ ), the production rate is thus unity when  $x > x_0$ , and  $\alpha_0 < 1$  when  $x < x_0$ . For  $\sigma \equiv 0$  the mean re-scaled number of protein follows the rate equation

$$\dot{\bar{x}} = f(\bar{x}) - \bar{x}, \quad (35)$$

where time is measured in units of generations. We choose  $\alpha_0 < x_0 < 1$ . Eq. (35) has three fixed points  $x_1^* < x_2^* < x_3^*$ , where  $x_1^* = \alpha_0$  and  $x_3^* = 1$  are attractors, and  $x_2^* = x_0$  is a repeller. Similar to Ref. [20], we refer to  $x_1^*$  and  $x_3^*$  as the “low” and “high” states, respectively.

The environmental process  $\sigma(t)$  modulates the production rate when  $x > x_0$ . As in Ref. [20] we assume that  $\sigma$  follows an Ornstein-Uhlenbeck process of the form given in Eq. (26). The noisy system has the potential to switch between the “high” and “low” states. To test the performance of the  $\tau$ FE algorithm, we focus on the mean switching time (MST) to transit from the high state to the low state. This time is studied and calculated in Ref. [20], we denote it by  $\langle \tau_{\text{high} \rightarrow \text{low}} \rangle$ . In simulations we start the system in the high state, and measure the first time the system reaches the low state.

Only the production of protein is affected by the state  $\sigma$  of the environment, we write  $R_{\text{prod},\sigma}(\mathbf{n}) = f(x, \sigma)$ , with  $f$  as in Eq. (34). Inserting this in Eqs. (29) and (30), and after straightforward calculations, we obtain

$$R_{\text{prod}}^*(n) = \alpha_0 + (1 - \alpha_0)\Theta(n/\Omega - x_0), \quad (36)$$

and the second moment

$$\begin{aligned} & \langle (\bar{R}_{\text{prod}}(n))^2 \rangle - [R_{\text{prod}}^*(n)]^2 \\ &= \frac{2v^2}{\lambda^2 \Delta t^2} [\lambda \Delta t + (e^{-\lambda \Delta t} - 1)] \Theta(n/\Omega - x_0). \end{aligned} \quad (37)$$

In Fig. 6 we show the MST measured in simulations using the different approaches described in in Sec. V. Assaf *et al.* [20] report nonmonotonous behavior of the MST as a function of  $\tau_c = \lambda^{-1}$ . As seen in Fig. 6 the  $\tau$ FE algorithm reproduces this behavior. For fast environmental dynamics (low  $\lambda^{-1}$ ) the MST obtained from the  $\tau$ FE algorithm is in good agreement with measurements obtained from the other simulation methods, and with the analytical approximations from Ref. [20]. The agreement extends over several decades of values of  $\tau_c = \lambda^{-1}$ .

At the same time we observe that the  $\tau$ FE algorithm requires significantly less computing time than the GADE or DEED approaches. For  $\lambda^{-1} = 10^{-1}$  for example, we measured an average computing time of  $2 \times 10^{-3}$  s to generate one run of the system up to time  $10^3$  with the  $\tau$ FE algorithm ( $\Delta t = 10/\lambda$ ). GADE required 0.674 s, and DEED 4.2 s (for a time step  $\Delta t = 10^{-4}$ ).

We note that we have implemented DEED as described in Sec. V C 2. In particular at most one reaction of each type can fire in each time step (step 3 of the algorithm). This requires a sufficiently small time step  $\Delta t$  to ensure  $p_r(t) < 1$  for all  $r$ . This is achieved by our choice  $\Delta t = 10^{-4}$ . Alternatively step 3 of the DEED algorithm could be replaced by a (conventional)  $\tau$ -leaping step. Larger choices of the time step  $\Delta t$  are then possible, up to the limit of  $\Delta t \approx \lambda^{-1}$  to ensure that the environmental dynamics are captured appropriately. Focusing on  $\lambda^{-1} = 10^{-1}$  we expect that increasing the time step by a factor of a thousand (from  $10^{-4}$  to  $10^{-1}$ ) would reduce the simulation time by at most a factor of a thousand for a  $\tau$ -leaping version of DEED. This would result in a computing time of approximately  $4 \times 10^{-3}$  for one simulation run up to  $t = 10^3$  instead of the 4.2 s reported for DEED in the previous paragraph. This is comparable with the CPU

time required by the  $\tau$ FE algorithm ( $2 \times 10^{-3}$  s), but would resolve environmental fluctuations with lower accuracy. For example one observes systematic deviations for the stationary distribution of the environment in Fig. 5(b).

## VII. DISCUSSION AND CONCLUSIONS

In summary, we have presented  $\tau$ FE, a variant of the  $\tau$ -leaping stochastic simulation algorithm for systems subject to fast environmental dynamics. Just like conventional  $\tau$ -leaping the algorithm operates in discrete time. The rates of the reactions in the system proper are treated as constant during each time step, and the numbers of different reactions firing in one step have Poissonian statistics.

The key difference compared to conventional  $\tau$ -leaping is the external environment. In the full continuous-time model reaction rates which depend on the environmental state fluctuate in time even when the state of the population does not change. An adiabatic approximation would consist of assuming an infinitely fast environment and of replacing the reaction rates by their means with respect to the stationary distribution of the environmental process. This is justified if the relaxation timescale of the environmental process is infinitely shorter than the time step of the simulation.

The  $\tau$ FE algorithm goes beyond this approximation, and is based on time averages of reaction rates over the *finite* time step. For finite speeds of the environment these average rates are random variables. If the environmental dynamics is fast, then we can make a Gaussian approximation. The rates feeding into the  $\tau$ -leaping step are clipped Gaussian random numbers designed to retain the first and second moments of the actual environmental dynamics. It is important to note that this not the same as drawing an environmental state  $\sigma$  from the stationary distribution  $\rho_\sigma^*$ , and then using the rates  $R_{r,\sigma}(\mathbf{n})$  for the next  $\tau$ -leaping step. Instead, the covariance matrix of the rates  $\bar{R}_r(\mathbf{n})$  in Eq. (8) is calculated as described in Eqs. (10) for discrete environments, and in Eq. (30) for continuous environmental states.

The choice of time step for the  $\tau$ FE algorithm requires careful consideration. On the one hand, the time step must be long enough to justify the averaging procedure over the environmental dynamics and the Gaussian assumption for the reaction rates in the  $\tau$ -leaping step. Broadly speaking  $\lambda \Delta t$  must be sufficiently large ( $\lambda \Delta t \gg 1$ ). At the same time the so-called leap condition for the  $\tau$ -leaping part of the algorithm must be fulfilled [37]. This means that the state of the system must not change significantly in each iteration step, as a constant state  $\mathbf{n}$  of the population is an assumption made in setting up the  $\tau$ -leaping. Mathematically, this means that the change of the number of particles in the system in a time step must be much smaller than the typical number of particles in the system. Assuming that the stoichiometric coefficients do not scale with the system size  $\Omega$  this means that  $\Delta t \times R_{r,\sigma}(\mathbf{n})$  must be much smaller than  $\Omega$ . Noting that  $R_{r,\sigma}(\mathbf{n})$  is of order  $\Omega$  in many applications we thus require that  $\Delta t$  is much smaller than one. For  $\lambda \gg 1$  and  $\Delta t$  proportional to  $\lambda^{-1}$  this condition is often relatively easy to meet in practice.

We have tested the  $\tau$ FE algorithm on a number of systems with discrete and continuous environments. This includes examples of systems which can be addressed analytically and

models motivated by applications in biology. Our tests focus on stationary distributions, but also dynamic features such as Fourier spectra of fluctuations or first-passage time distributions. In all cases we have tested the  $\tau$ FE method produces good agreement with results from conventional simulation methods in the regime of fast environmental dynamics. This is the regime for which  $\tau$ FE is designed. Naturally, quantitative deviations are found when the timescales of the environmental dynamics and system proper are insufficiently separated.

We stress that  $\tau$ FE goes beyond simulations in the adiabatic limit, and is able to capture the dependence of macroscopic observables on the timescale separation, provided this dependence is sufficiently strong [see, e.g., Figs. 4(d) and 6]. At the same time our analysis also reveals limitations of the algorithm. If the dependence of observables on the timescale separation is weak, such as in Fig. 4(c), then  $\tau$ FE may not be able to fully resolve these dependencies. When the environment is fast the quantitative agreement with simulations of the full system is however still within approximately 10% in the example in Fig. 4(c).

The computing time required for the  $\tau$ FE algorithm to generate sample paths up to a designated end time is proportional to the inverse time step. The time step on the other hand is typically a multiple of the characteristic timescale  $\lambda^{-1}$  of the environmental dynamics. This means that the computational effort scales approximately linearly in the timescale separation.

In all cases we have tested we found that  $\tau$ FE is considerably more efficient for the measurement of macroscopic quantities than alternative simulation algorithms.

In summary, we think the  $\tau$ FE algorithm has passed the initial selection of tests presented in this paper. It provides an promising approach to probing the regime of fast environmental dynamics, and captures effects induced by extrinsic noise beyond the adiabatic limit. The algorithm is particularly valuable for systems in which the regime of intermediate timescale separation can be accessed with conventional simulation methods. The accuracy of the  $\tau$ FE algorithm can then be assessed in this regime (an example can be found in Fig. 6). If the comparison is favourable, then it is justified to use  $\tau$ FE in the regime of increasing timescale separation.

### ACKNOWLEDGMENTS

We thank Yen Ting Lin (Los Alamos) for useful discussions and feedback on earlier versions of the manuscript. E.B.C. acknowledges a President's Doctoral Scholarship (The University of Manchester). T.G. acknowledges funding from the Spanish Ministry of Science, Innovation and Universities, the Agency AEI and FEDER (EU) under PACSS Grant No. RTI2018-093732-B-C22, and the Maria de Maeztu program for Units of Excellence in R&D (Grant No. MDM-2017-0711).

### APPENDIX A: SECOND MOMENTS OF RATES

In this Appendix we calculate the second moments of the quantities  $\bar{R}_r(\mathbf{n})$  ( $r = 1, \dots, R$ ) defined in Eq. (8). Without loss of generality we assume that the time interval in question starts at  $t = 0$ , the end point is then  $\Delta t$ . Assuming the space of environmental states is discrete, we have

$$\begin{aligned}
\langle \bar{R}_r(\mathbf{n}) \bar{R}_s(\mathbf{n}) \rangle &= \frac{1}{\Delta t^2} \int_0^{\Delta t} dt_1 \int_0^{\Delta t} dt_2 \langle R_{r,\sigma(t_1)}(\mathbf{n}) R_{s,\sigma(t_2)}(\mathbf{n}) \rangle \\
&= \frac{1}{\Delta t^2} \int_0^{\Delta t} dt_1 \int_{t_1}^{\Delta t} dt_2 \langle R_{r,\sigma(t_1)}(\mathbf{n}) R_{s,\sigma(t_2)}(\mathbf{n}) \rangle + \frac{1}{\Delta t^2} \int_0^{\Delta t} dt_2 \int_{t_2}^{\Delta t} dt_1 \langle R_{r,\sigma(t_1)}(\mathbf{n}) R_{s,\sigma(t_2)}(\mathbf{n}) \rangle \\
&= \frac{1}{\Delta t^2} \sum_{\sigma\sigma'} \int_0^{\Delta t} dt_1 \int_{t_1}^{\Delta t} dt_2 \rho_{\sigma}^* q_{\sigma \rightarrow \sigma'}(t_2 - t_1) R_{r,\sigma}(\mathbf{n}) R_{s,\sigma'}(\mathbf{n}) \\
&\quad + \frac{1}{\Delta t^2} \sum_{\sigma\sigma'} \int_0^{\Delta t} dt_2 \int_{t_2}^{\Delta t} dt_1 \rho_{\sigma'}^* q_{\sigma' \rightarrow \sigma}(t_1 - t_2) R_{r,\sigma}(\mathbf{n}) R_{s,\sigma'}(\mathbf{n}) \\
&= \frac{1}{\Delta t^2} \sum_{\sigma\sigma'} \int_0^{\Delta t} dt_1 \int_{t_1}^{\Delta t} dt_2 \rho_{\sigma}^* q_{\sigma \rightarrow \sigma'}(t_2 - t_1) R_{r,\sigma}(\mathbf{n}) R_{s,\sigma'}(\mathbf{n}) \\
&\quad + \frac{1}{\Delta t^2} \sum_{\sigma\sigma'} \int_0^{\Delta t} dt_1 \int_{t_1}^{\Delta t} dt_2 \rho_{\sigma'}^* q_{\sigma \rightarrow \sigma'}(t_2 - t_1) R_{r,\sigma'}(\mathbf{n}) R_{s,\sigma}(\mathbf{n}). \tag{A1}
\end{aligned}$$

In the first step we have applied the definition of the over-bar average [Eq. (8)]. In the third step we have carried out the average over realizations of the environmental process. In the last step we have renamed  $t_1 \leftrightarrow t_2$  and  $\sigma \leftrightarrow \sigma'$  in the second term. Therefore,

$$\langle \bar{R}_r(\mathbf{n}) \bar{R}_s(\mathbf{n}) \rangle = \frac{1}{\Delta t^2} \sum_{\sigma\sigma'} \int_0^{\Delta t} dt_1 \int_{t_1}^{\Delta t} dt_2 \rho_{\sigma}^* q_{\sigma \rightarrow \sigma'}(t_2 - t_1) [R_{r,\sigma}(\mathbf{n}) R_{s,\sigma'}(\mathbf{n}) + R_{r,\sigma'}(\mathbf{n}) R_{s,\sigma}(\mathbf{n})]. \tag{A2}$$

Up to a shift of the start point of the time step, this is identical to Eq. (10).

As explained in Section VB, the sums over  $\sigma$  become integrals when the environment takes continuous states. We then find Eq. (30).

When the environmental space is discrete, we can use Eq. (7) and find

$$\begin{aligned}
 \langle \bar{R}_r(\mathbf{n}) \bar{R}_s(\mathbf{n}) \rangle &= \frac{1}{\Delta t^2} \sum_{\sigma\sigma'} \int_0^{\Delta t} dt_1 \int_{t_1}^{\Delta t} dt_2 \rho_{\sigma}^* \rho_{\sigma'}^* R_{r,\sigma}(\mathbf{n}) R_{s,\sigma'}(\mathbf{n}) + R_{r,\sigma'}(\mathbf{n}) R_{s,\sigma}(\mathbf{n}) \\
 &+ \frac{1}{\Delta t^2} \sum_{\sigma\sigma'} \sum_{\ell=2}^M \int_0^{\Delta t} dt_1 \int_{t_1}^{\Delta t} dt_2 \rho_{\sigma}^* c_{\ell,\sigma} v_{\ell,\sigma'} e^{-\lambda \mu_{\ell}(t_2-t_1)} [R_{r,\sigma}(\mathbf{n}) R_{s,\sigma'}(\mathbf{n}) + R_{r,\sigma'}(\mathbf{n}) R_{s,\sigma}(\mathbf{n})] \\
 &= R_{r,\text{avg}}(\mathbf{n}) R_{s,\text{avg}}(\mathbf{n}) \\
 &+ \frac{1}{\Delta t^2} \sum_{\sigma\sigma'} \sum_{\ell=2}^M \rho_{\sigma}^* c_{\ell,\sigma} v_{\ell,\sigma'} [R_{r,\sigma}(\mathbf{n}) R_{s,\sigma'}(\mathbf{n}) + R_{r,\sigma'}(\mathbf{n}) R_{s,\sigma}(\mathbf{n})] \int_0^{\Delta t} dt_1 \int_{t_1}^{\Delta t} dt_2 e^{\lambda \mu_{\ell}(t_2-t_1)}. \quad (\text{A3})
 \end{aligned}$$

## APPENDIX B: FURTHER DETAILS FOR SYSTEMS WITH TWO SPECIES AND TWO ENVIRONMENTAL STATES

The case of two species and two environmental states ( $S = 2, M = 2$ ) was studied in Ref. [38], and a simple version of the  $\tau$ FE algorithm was presented for this restricted case. We assume  $\sigma$  switches from state 0 to state 1 with rate  $\lambda k_1$ , and from 1 to 0 with rate  $\lambda k_0$ . The environmental transition matrix then becomes

$$\mathbf{A} = \begin{pmatrix} -k_1 & k_0 \\ k_1 & -k_0 \end{pmatrix}, \quad (\text{B1})$$

whose eigenvalues are  $\mu_1 = 0$  and  $\mu_2 = -(k_0 + k_1)$ . The respective eigenvectors take the form

$$\mathbf{v}_1 = \boldsymbol{\rho}^* = \frac{1}{k_0 + k_1} \begin{pmatrix} k_0 \\ k_1 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad (\text{B2})$$

where  $\boldsymbol{\rho}^*$  has been normalized to represent the stationary distribution for  $\sigma$ . The coefficients  $c_{2,0}$  and  $c_{2,1}$  are obtained from Eq. (6), for the initial conditions  $\boldsymbol{\rho}(0) = (1, 0)$  and  $\boldsymbol{\rho}(0) = (0, 1)$ . We find

$$c_{2,0} = \frac{k_1}{k_0 + k_1} \quad \text{and} \quad c_{2,1} = \frac{-k_0}{k_0 + k_1}. \quad (\text{B3})$$

Putting all together in Eq. (13), and after straightforward calculations we arrive at

$$\Xi_{rs} \equiv \langle \bar{R}_r(\mathbf{n}) \bar{R}_s(\mathbf{n}) \rangle - R_r^*(\mathbf{n}) R_s^*(\mathbf{n}) = \frac{\theta^2}{\lambda \Delta t} [R_{r,1}(\mathbf{n}) - R_{r,0}(\mathbf{n})][R_{s,1}(\mathbf{n}) - R_{s,0}(\mathbf{n})], \quad (\text{B4})$$

where  $\theta^2 = 2k_0k_1/(k_0 + k_1)^3$ . The indices  $r$  and  $s$  stand for reactions affected by the environment. As explained in Sec. III C, to simulate the  $\tau$ FE algorithm we need to draw correlated Gaussian random numbers  $\bar{R}_r$  with means

$$R_r^*(\mathbf{n}) = \frac{k_0 R_{r,0} + k_1 R_{r,1}}{k_0 + k_1}, \quad (\text{B5})$$

for  $r = 1, 2$ , and covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \Xi_{11} & \Xi_{12} \\ \Xi_{21} & \Xi_{22} \end{pmatrix}. \quad (\text{B6})$$

One way to do this is by drawing independent Gaussian random numbers  $z_1$  and  $z_2$  with mean zero and unit variance, and then to set

$$\begin{pmatrix} \bar{R}_1(\mathbf{n}) \\ \bar{R}_2(\mathbf{n}) \end{pmatrix} = \begin{pmatrix} R_1^*(\mathbf{n}) \\ R_2^*(\mathbf{n}) \end{pmatrix} + \mathbf{C} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \quad (\text{B7})$$

with a matrix  $\mathbf{C}$  that fulfils  $\mathbf{C}\mathbf{C}^T = \boldsymbol{\Sigma}$ , where  $T$  denotes the transpose. This matrix is not unique. We use

$$\mathbf{C} = \frac{\boldsymbol{\Sigma}}{\sqrt{\theta^2/(\lambda \Delta t) \{ [R_{1,1}(\mathbf{n}) - R_{1,0}(\mathbf{n})]^2 + [R_{2,1}(\mathbf{n}) - R_{2,0}(\mathbf{n})]^2 \}}}. \quad (\text{B8})$$

## APPENDIX C: BIRTH-DEATH PROCESS WITH TWO SPECIES AND THREE ENVIRONMENTAL STATES

In the example in Sec. IV B we have the following transition matrix for the environmental process

$$\mathbf{A} = \begin{pmatrix} -k_1 & 0 & k_0 \\ k_1 & -k_2 & 0 \\ 0 & k_2 & -k_0 \end{pmatrix}. \quad (\text{C1})$$



The eigenvalues of this matrix are

$$\mu_1 = 0, \quad \mu_2 = -\frac{1}{2}(k_0 + k_1 + k_2 + \Gamma), \quad \text{and}, \quad \mu_3 = -\frac{1}{2}(k_0 + k_1 + k_2 - \Gamma), \quad (\text{C2})$$

with  $\Gamma = \sqrt{k_0^2 + k_1^2 + k_2^2 - 2(k_0k_1 + k_1k_2 + k_2k_0)}$ . The associated eigenvectors take the form

$$\mathbf{v}_1 = \boldsymbol{\rho}^* = \frac{1}{k_0k_1 + k_1k_2 + k_2k_0} \begin{pmatrix} k_2k_0 \\ k_0k_1 \\ k_1k_2 \end{pmatrix}, \quad (\text{C3})$$

and

$$\mathbf{v}_2 = \begin{pmatrix} (-k_0 + k_1 - k_2 + \Gamma)/(2k_2) \\ (k_0 - k_1 - k_2 - \Gamma)/(2k_2) \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} (-k_0 + k_1 - k_2 - \Gamma)/(2k_2) \\ (k_0 - k_1 - k_2 + \Gamma)/(2k_2) \\ 1 \end{pmatrix}. \quad (\text{C4})$$

Using Eq. (6) and three sets of initial conditions (each concentrated on one environmental state) we find

$$c_{2,0} = \frac{k_1k_2(k_0 + k_1 + k_2 - \Gamma)}{2\Gamma(k_0k_1 + k_1k_2 + k_2k_0)}, \quad c_{3,0} = -\frac{k_1k_2(k_0 + k_1 + k_2 + \Gamma)}{2\Gamma(k_0k_1 + k_1k_2 + k_2k_0)}, \quad (\text{C5})$$

as well as

$$c_{2,1} = -\frac{k_2(k_0(k_1 + 2k_2) + k_1(-k_1 + k_2 + \Gamma))}{2\Gamma(k_0k_1 + k_1k_2 + k_2k_0)}, \quad c_{3,1} = \frac{k_2(k_0(k_1 + 2k_2) + k_1(-k_1 + k_2 - \Gamma))}{2\Gamma(k_0k_1 + k_1k_2 + k_2k_0)}, \quad (\text{C6})$$

and finally

$$c_{2,2} = \frac{k_0(-k_1^2 - k_2^2 + k_0(k_1 + k_2) + k_1\Gamma + k_2\Gamma)}{2\Gamma(k_0k_1 + k_1k_2 + k_2k_0)}, \quad c_{3,2} = \frac{k_0(k_1^2 + k_2^2 - k_0(k_1 + k_2) + k_1\Gamma + k_2\Gamma)}{2\Gamma(k_0k_1 + k_1k_2 + k_2k_0)}. \quad (\text{C7})$$

Putting all together in Eqs. (2) and (13) and after further tedious but straightforward calculations, we arrive at the expressions in Eqs. (18) and (19).

To draw the correlated Gaussian random numbers  $\bar{\alpha}$  and  $\bar{\beta}$  required for the  $\tau$ -leaping step, we proceed as in Appendix B. We construct the covariance matrix  $\boldsymbol{\Sigma}$  [Eq. (B6)] and then find a matrix  $\mathbf{C}$  such that  $\mathbf{C}\mathbf{C}^T = \boldsymbol{\Sigma}$ . We then draw independent Gaussian random numbers  $z_1$  and  $z_2$  with mean zero and unit variance, and use an expression analogous to that in Eq. (B7) to obtain  $\bar{\alpha}$  and  $\bar{\beta}$ . The matrix  $\mathbf{C}$  we use is

$$\mathbf{C} = A \begin{pmatrix} \frac{\sigma_{\alpha\alpha} + B}{\sigma_{\alpha\beta}} & 1 \\ 1 & \frac{\sigma_{\beta\beta} + B}{\sigma_{\alpha\beta}} \end{pmatrix}, \quad (\text{C8})$$

with  $\sigma_{\alpha\alpha}$  and  $\sigma_{\alpha\beta}$  as given in Eq. (19), and

$$A = \frac{\sigma_{\alpha\beta}}{\sqrt{\sigma_{\alpha\alpha} + \sigma_{\beta\beta} + B}}, \quad (\text{C9})$$

and

$$B = \frac{\sqrt{3}k_0k_1k_2}{\lambda\Delta t(k_0k_1 + k_1k_2 + k_2k_0)^2} |\alpha_0(\beta_2 - \beta_1) + \alpha_1(\beta_0 - \beta_2) + \alpha_2(\beta_1 - \beta_0)|. \quad (\text{C10})$$

#### APPENDIX D: BIMODAL GENETIC SWITCH

For the model in Sec. IV C the rates of the environmental transitions depend on the number of proteins  $n_P$  in the population. We assume that  $n_P$  remains constant during each  $\tau$ -leaping step. The environmental transition matrix then becomes

$$\mathbf{A} = \begin{pmatrix} -\tilde{k}_+ & k_- & 0 \\ \tilde{k}_+ & -\tilde{k}_+ - k_- & k_- \\ 0 & \tilde{k}_+ & -k_- \end{pmatrix}, \quad (\text{D1})$$

with  $\tilde{k}_+ = k_+n_P/\Omega$ . The eigenvalues of this matrix are

$$\mu_1 = 0, \quad \mu_2 = -k_- - \tilde{k}_+ - \sqrt{k_- \tilde{k}_+}, \quad \mu_3 = -k_- - \tilde{k}_+ + \sqrt{k_- \tilde{k}_+}, \quad (\text{D2})$$

while the associated eigenvectors take the form

$$\mathbf{v}_1 = \boldsymbol{\rho}^* = \frac{1}{k_-^2 + k_- \tilde{k}_+ + \tilde{k}_+^2} \begin{pmatrix} k_-^2 \\ k_- \tilde{k}_+ \\ \tilde{k}_+^2 \end{pmatrix},$$

$$\mathbf{v}_2 = \begin{pmatrix} \sqrt{k_-/\tilde{k}_+} \\ (-\sqrt{k_-} - \sqrt{\tilde{k}_+})/\sqrt{\tilde{k}_+} \\ 1 \end{pmatrix}, \quad \text{and} \quad \mathbf{v}_3 = \begin{pmatrix} -\sqrt{k_-/\tilde{k}_+} \\ (\sqrt{k_-} - \sqrt{\tilde{k}_+})/\sqrt{\tilde{k}_+} \\ 1 \end{pmatrix}. \quad (\text{D3})$$

Applying Eq. (6) for different initial conditions as above, we obtain

$$c_{2,0} = \frac{\tilde{k}_+^{3/2}}{2\sqrt{k_-}(k_- + \sqrt{k_- \tilde{k}_+} + \tilde{k}_+)}, \quad c_{3,0} = -\frac{\tilde{k}_+^{3/2}}{2\sqrt{k_-}(k_- - \sqrt{k_- \tilde{k}_+} + \tilde{k}_+)}, \quad (\text{D4})$$

as well as

$$c_{2,1} = -\frac{\tilde{k}_+ + \sqrt{k_- \tilde{k}_+}}{2(k_- + \sqrt{k_- \tilde{k}_+} + \tilde{k}_+)}, \quad c_{3,1} = -\frac{\tilde{k}_+ - \sqrt{k_- \tilde{k}_+}}{2(k_- - \sqrt{k_- \tilde{k}_+} + \tilde{k}_+)}, \quad (\text{D5})$$

and finally

$$c_{2,2} = \frac{k_-}{2(k_- + \sqrt{k_- \tilde{k}_+} + \tilde{k}_+)}, \quad c_{3,2} = \frac{k_-}{2(k_- - \sqrt{k_- \tilde{k}_+} + \tilde{k}_+)}. \quad (\text{D6})$$

Putting this together in Eqs. (2) and (13) and after straightforward calculations, we arrive at the expressions in Eqs. (23) and (24).

Since only one reaction is affected by the environmental state, it is only necessary to draw one Gaussian random number with mean  $b^*$  and variance  $\sigma_{bb}$  in each step of the  $\tau$ FE algorithm, with  $b^*$  and  $\sigma_{bb}$  given in Eqs. (23) and (24), respectively.

## APPENDIX E: GILLESPIE ALGORITHM WITH DISCRETIZED ENVIRONMENTAL DYNAMICS (GADE)

In this Appendix we briefly describe the constructions of the rates given in Eq. (31). They define a continuous-time dynamics on a discrete state space approximating the Ornstein–Uhlenbeck process in Eq. (26).

*Matching the first moments of movements.* We first look at the mean drift of  $\sigma$ , i.e., the mean change of  $\sigma$  per unit time. Suppose the environment is in a given state  $\sigma$ . The mean drift in the Ornstein–Uhlenbeck process [Eq. (26)] is then  $\lambda(m - \sigma)$ .

Suppose now the above discrete- $\sigma$  process is in state  $\sigma = k\Delta\sigma$ . Then  $\sigma$  increases to  $\sigma + \Delta\sigma$  with rate  $T_k^+$  and decreases to  $\sigma - \Delta\sigma$  with rate  $T_k^-$ . The expected change (per unit time) is therefore  $\Delta\sigma \times (T_k^+ - T_k^-)$ .

We conclude that we need to impose

$$\Delta\sigma \times (T_k^+ - T_k^-) = \lambda(m - k\Delta\sigma). \quad (\text{E1})$$

### 1. Matching the variance of movements

Next we look at the variance of movements of  $\sigma$ . For the Ornstein–Uhlenbeck process in Eq. (26) the second moment of movements (per unit time) is given by  $2\lambda v^2$ . In the discrete- $\sigma$  process, the second moment of movements is  $(\Delta\sigma)^2 \times (T_k^+ + T_k^-)$ . To match the Ornstein–Uhlenbeck process, we then need to impose

$$(\Delta\sigma)^2 \times (T_k^+ + T_k^-) = 2\lambda v^2. \quad (\text{E2})$$

### 2. Overall solution

Simultaneously solving Eqs. (E1) and (E2) for  $T_k^+$  and  $T_k^-$  we arrive at Eq. (31).

## APPENDIX F: ADDITIONAL EXAMPLES OF PRODUCTION-REMOVAL PROCESSES IN CONTINUOUS ENVIRONMENTS

In this Appendix we include results for the variances and covariances  $\langle \bar{R}_r(n)\bar{R}_s(n) \rangle - R_r^*(\mathbf{n})R_s^*(\mathbf{n})$  for two further exemplar systems in which the environment follows the Ornstein–Uhlenbeck process in Eq. (26). We set  $m = 0$  for both examples. Both systems describe production and removal dynamics of a single species. In the first example, production and removal rates are proportional to  $\sigma$  when  $\sigma > 0$  and zero otherwise. In the second example the rates are each proportional to  $|\sigma|$ . These examples are not used in the main paper, we report them here for completeness, as they may prove useful for future applications of the  $\tau$ FE algorithm.

### 1. Rates $R_{r,\sigma}(n) = \alpha_r \sigma \Theta(\sigma)$

We look at the example  $R_{r,\sigma}(n) = \alpha_r \sigma \Theta(\sigma)$ , where  $\Theta(\sigma)$  is the Heaviside function,  $\Theta(\sigma) = 1$  for  $\sigma > 0$  and  $\Theta(\sigma) = 0$  otherwise. For  $m = 0$ , we find

$$R_r^*(n) = \alpha_r \frac{v}{\sqrt{\pi}}, \quad (\text{F1})$$

and

$$\begin{aligned} \langle \bar{R}_r(n) \bar{R}_s(n) \rangle - R_r^*(n) R_s^*(n) &= \frac{\alpha_r \alpha_s v^2}{24\pi \Delta t^2} \left\{ \frac{1}{\lambda^2} [24\pi(\lambda \Delta t - 1) - \pi^2 + 12 \log^2(2)] \right. \\ &+ \frac{4e^{-\lambda \Delta t}}{\lambda^2} \left[ 3 \left( 6\sqrt{e^{2\lambda \Delta t} - 1} + \pi \right) - 4e^{\lambda \Delta t} \log \left( \sqrt{e^{2\lambda \Delta t} - 1} + e^{\lambda \Delta t} \right) + 6 \tan^{-1} \left( \frac{1}{\sqrt{e^{2\lambda \Delta t} - 1}} \right) \right] \\ &- \frac{32}{\lambda^2} \text{Re}(i \sin^{-1}(e^{\lambda \Delta t})) - 6 \left[ \frac{1}{\lambda^2} \log^2 \left( \sqrt{1 - e^{-2\lambda \Delta t}} + 1 \right) + \frac{4\Delta t}{\lambda} \log \left( \sqrt{1 - e^{-2\lambda \Delta t}} + 1 \right) \right. \\ &- \frac{4\Delta t \log(2)}{\lambda} - \frac{\log(4)}{\lambda^2} \log \left( \sqrt{1 - e^{-2\lambda \Delta t}} + 1 \right) - \frac{4\Delta t}{\lambda} \tanh^{-1} \left( e^{-\lambda \Delta t} \sqrt{e^{2\lambda \Delta t} - 1} \right) \\ &\left. \left. - \frac{2}{\lambda^2} \text{Li}_2 \left( \frac{1}{2} \left( 1 - \sqrt{1 - e^{-2\lambda \Delta t}} \right) \right) + \frac{\log^2(2)}{\lambda^2} + 2\Delta t^2 \right] - 24 \right\}, \quad (\text{F2}) \end{aligned}$$

where  $\text{Re}(\cdot)$  denotes the real part, and  $\text{Li}_2(\cdot)$  is the polylogarithm of order 2.

### 2. Rates $R_{r,\sigma}(n) = \alpha_r |\sigma|$

For this case (and setting again  $m = 0$ ), we find

$$R_r^*(n) = \alpha_r \frac{2v}{\sqrt{\pi}}, \quad (\text{F3})$$

and

$$\begin{aligned} \langle \bar{R}_r(n) \bar{R}_s(n) \rangle - R_r^*(n) R_s^*(n) &= \frac{\alpha_r \alpha_s v^2}{6\pi \Delta t^2} \left\{ \frac{12\Delta t}{\lambda} \left[ -2 \log \left( \sqrt{1 - e^{-2\lambda \Delta t}} + 1 \right) + 2 \tanh^{-1} \left( \sqrt{1 - e^{-2\lambda \Delta t}} \right) + \pi + \log(4) \right] \right. \\ &+ \frac{1}{\lambda^2} \left[ 72e^{-\lambda \Delta t} \sqrt{e^{2\lambda \Delta t} - 1} - 6 \log^2 \left( \frac{1}{2} \left( \sqrt{1 - e^{-2\lambda \Delta t}} + 1 \right) \right) + 24e^{-\lambda \Delta t} \tan^{-1} \left( \frac{1}{\sqrt{e^{2\lambda \Delta t} - 1}} \right) \right. \\ &\left. \left. - 48 \tanh^{-1} \left( e^{-\lambda \Delta t} \sqrt{e^{2\lambda \Delta t} - 1} \right) + 12 \text{Li}_2 \left( \frac{1}{2} \left( 1 - \sqrt{1 - e^{-2\lambda \Delta t}} \right) \right) - \pi^2 - 12\pi + 12 \log^2(2) \right] - 12(\Delta t^2 + 2) \right\}. \quad (\text{F4}) \end{aligned}$$

- 
- [1] J. D. Murray, *Mathematical Biology I. An Introduction*, 3rd ed., Interdisciplinary Applied Mathematics (Springer, New York, 2002), Vol. 17.
- [2] J. D. Murray, *Mathematical Biology II: Spatial Models and Biomedical Applications*, Interdisciplinary Applied Mathematics (Springer, New York, 2003), Vol. 18.
- [3] N. S. Goel and N. Richter-Dyn, *Stochastic Models in Biology* (Blackburn Press, Caldwell, NJ, 2004).
- [4] W. Ewens, *Mathematical Population Genetics I* (Springer-Verlag, New York, 2004).
- [5] A. Traulsen and C. Hauert, Stochastic evolutionary game dynamics, in *Reviews of Nonlinear Dynamics and Complexity* (Wiley-VCH Verlag GmbH and Co. KGaA, Berlin, 2010), pp. 25–61.
- [6] C. Castellano, S. Fortunato, and V. Loreto, *Rev. Mod. Phys.* **81**, 591 (2009).
- [7] M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, Princeton, NJ, 2008).
- [8] N. V. Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 2007).
- [9] M. Acar, J. T. Mettetal, and A. Van Oudenaarden, *Nat. Genet.* **40**, 471 (2008).
- [10] P. Patra and S. Klumpp, *Phys. Biol.* **12**, 046004 (2015).
- [11] K. Wienand, E. Frey, and M. Mobilia, *Phys. Rev. Lett.* **119**, 158301 (2017).
- [12] K. Wienand, E. Frey, and M. Mobilia, *J. R. Soc., Interface* **15**, 20180343 (2018).

- [13] A. Taitelbaum, R. West, M. Assaf, and M. Mabilia, *Phys. Rev. Lett.* **125**, 048105 (2020).
- [14] A. J. Black and A. J. McKane, *J. Theor. Biol.* **267**, 85 (2010).
- [15] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 3rd ed., Springer Series in Synergetics (Springer-Verlag, Berlin, 2004), Vol. 13.
- [16] E. Kussell and S. Leibler, *Science* **309**, 2075 (2005).
- [17] T. B. Kepler and T. C. Elston, *Biophys. J.* **81**, 3116 (2001).
- [18] M. Thattai and A. Van Oudenaarden, *Genetics* **167**, 523 (2004).
- [19] P. S. Swain, M. B. Elowitz, and E. D. Siggia, *Proc. Natl. Acad. Sci. USA* **99**, 12795 (2002).
- [20] M. Assaf, E. Roberts, Z. Luthey-Schulten, and N. Goldenfeld, *Phys. Rev. Lett.* **111**, 058102 (2013).
- [21] A. Duncan, S. Liao, T. Vejchodský, R. Erban, and R. Grima, *Phys. Rev. E* **91**, 042111 (2015).
- [22] M. Assaf, M. Mabilia, and E. Roberts, *Phys. Rev. Lett.* **111**, 238101 (2013).
- [23] P. Ashcroft, P. M. Altrock, and T. Galla, *J. R. Soc., Interface* **11**, 20140663 (2014).
- [24] R. West, M. Mabilia, and A. M. Rucklidge, *Phys. Rev. E* **97**, 022406 (2018).
- [25] M. Assaf, A. Kamenev, and B. Meerson, *Phys. Rev. E* **78**, 041123 (2008).
- [26] P. G. Hufton, Y. T. Lin, and T. Galla, *Phys. Rev. E* **99**, 032122 (2019).
- [27] A. Eldar and M. Elowitz, *Nature (London)* **467**, 167 (2010).
- [28] D. T. Gillespie, *J. Comput. Phys.* **22**, 403 (1976).
- [29] D. T. Gillespie, *J. Phys. Chem.* **81**, 2340 (1977).
- [30] J. Bowen, A. Acrivos, and A. Oppenheim, *Chem. Eng. Sci.* **18**, 177 (1963).
- [31] L. A. Segel and M. Slemrod, *SIAM Rev.* **31**, 446 (1989).
- [32] Y. T. Lin and N. E. Buchler, *J. R. Soc., Interface* **15**, 20170804 (2018).
- [33] J. M. Newby and P. C. Bressloff, *Bull. Math. Biol.* **72**, 1840 (2010).
- [34] P. C. Bressloff, *Phys. Rev. E* **94**, 042129 (2016).
- [35] P. C. Bressloff, *Phys. Rev. E* **95**, 012124 (2017).
- [36] P. C. Bressloff, *Phys. Rev. E* **95**, 012138 (2017).
- [37] D. T. Gillespie, *J. Chem. Phys.* **115**, 1716 (2001).
- [38] P. G. Hufton, Y. T. Lin, and T. Galla, *Phys. Rev. E* **99**, 032121 (2019).
- [39] P. C. Bressloff and J. M. Newby, *Phys. Rev. E* **89**, 042701 (2014).
- [40] Y. Suhov and M. Kelbert, *Probability and Statistics by Example* (Cambridge University Press, Cambridge, UK, 2008), Vol. 2.
- [41] P. G. Hufton, Y. T. Lin, T. Galla, and A. J. McKane, *Phys. Rev. E* **93**, 052119 (2016).
- [42] J. Gunawardena, *FEBS J.* **281**, 473 (2014).
- [43] N. E. Buchler, U. Gerland, and T. Hwa, *Proc. Natl. Acad. Sci. USA* **100**, 5136 (2003).
- [44] B. Fuglede and F. Topsøe, in *Proceedings of the International Symposium on Information Theory* (IEEE, Piscataway, NJ, 2004), p. 31.
- [45] J. Lin, *IEEE Trans. Info. Theory* **37**, 145 (1991).
- [46] For completeness, we add that simulations were performed on a MacBook Pro (Mid 2014), with processor 2.6 GHz Dual-Core Inter Core i5, and memory 8 GB 1600 MHz DDR3.
- [47] Y. T. Lin, P. G. Hufton, E. J. Lee, and D. A. Potoyan, *PLoS Comput. Biol.* **14**, 1 (2018).
- [48] E. Roberts, S. Be'er, C. Bohrer, R. Sharma, and M. Assaf, *Phys. Rev. E* **92**, 062717 (2015).
- [49] F. C. Klebaner, *Introduction to Stochastic Calculus with Applications* (World Scientific, Singapore, 2005).
- [50] H. Risken, in *The Fokker-Planck Equation* (Springer, Berlin, 1996), pp. 63–95.
- [51] G. Maruyama, *Rendiconti del Circolo Matematico di Palermo* **4**, 48 (1955).
- [52] P. W. Lewis and G. S. Shedler, *Naval Res. Logist. Quart.* **26**, 403 (1979).