

Improved diffusive approximation of Markov jump processes close to equilibrium

David Roberts , Trevor McCourt*, Geremia Massarelli , Jeremy Rothschild , and Nahuel Freitas 
Extropic Corporation, Cambridge, Massachusetts, USA

(Received 4 July 2025; accepted 9 January 2026; published 11 February 2026)

Diffusive approximations of Markov jump processes often fail to accurately capture large fluctuations. This is confounding, as the rare events triggered by these large fluctuations, such as the failure of electronic memories, are often the object of interest. In this paper, we present an improved diffusive approximation, extending a method previously limited to equilibrium systems. Using recently developed tools from stochastic thermodynamics, we prove its validity to linear order in departures from equilibrium and demonstrate its superior accuracy over the Kramers-Moyal expansion in predicting both steady state and transient properties, including the error rate of a nonequilibrium electronic memory.

DOI: [10.1103/dw4d-yj8c](https://doi.org/10.1103/dw4d-yj8c)

I. INTRODUCTION

Markov jump processes (MJPs) in continuous time are ubiquitous across many different disciplines, being employed to model stochastic phenomena in biology, chemistry, physics, and finance. For example, continuous-time MJPs are the basis for the stochastic description of elementary chemical reactions in cells, and also for effective models of their gene regulatory networks [1,2]. In physics, MJPs are at the core of important models of nonequilibrium systems like lattice-gas models [3,4] and, among other examples, are employed to describe the transport of electrons across both nanoscopic [5,6] and mesoscopic [7–9] electronic circuits.

While exact algorithms to simulate MJPs exist [10,11], it is often convenient to approximate MJPs by diffusive processes typically described by Langevin dynamics at the trajectory level or Fokker-Planck equations at the ensemble level. A continuous representation is often advantageous—analytically (e.g., closed-form asymptotics, perturbation, and spectral methods) and computationally. Examples in which diffusion approximations are particularly effective include chemical Langevin equations [12,13] in biochemistry or stochastic Lotka-Volterra models [14,15] in population dynamics.

Thus, the diffusive approximation of MJPs is an important problem in statistical physics with a long “confused history” [12,17,18], for which different methods have been developed. The truncations of the Kramers-Moyal expansion [17,19] and the van Kampen linear noise approximation are the most popular [20].

However, it is well known that popular diffusive approximations are only guaranteed to be accurate for small, typical

fluctuations and fail to capture the probability of large, rare fluctuations [21–23]. This is relevant in problems displaying metastability, where rare fluctuations are involved in the switching events between different dynamical attractors. Prominent examples are given by extinction events in population dynamics models [24–26], metastable gene expression patterns in gene regulatory networks [2], or the occurrence of thermally induced soft errors in bistable electronic circuits [27,28].

To address this limitation, we introduce an improved diffusive approximation for Markov jump processes that remains accurate for arbitrarily large fluctuations, even in nonequilibrium systems. Our central contribution is to generalize a method, previously restricted to systems at detailed balance, by proving that it correctly captures macroscopic fluctuations to linear order in the departure from equilibrium.

Our work provides a more robust and systematically improvable framework for studying rare events using diffusion approximations. We validate our approach on models of stochastic electronic circuits, demonstrating that it more accurately predicts both steady-state distributions and key transient properties, such as the error rate of a bistable memory, when compared to standard approximations. This improved accuracy is not merely a theoretical curiosity: The modeling framework for stochastic electronic circuits that is featured in the examples studied here has been successfully applied to the design and experimental validation of probabilistic circuits in the companion articles [9,29].

The key to our improved approximation is a modification of the diffusion tensor in the Fokker-Planck equation. While the standard Kramers-Moyal approximation uses a simple sum of forward and reverse jump rates, our formulation employs their logarithmic mean. This seemingly minor change has profound consequences, ensuring the resulting diffusive process correctly captures large fluctuations in the steady state for systems that are close to being detailed balanced. While primarily tailored to reproduce steady-state fluctuations, our method also yields more precise estimates for key dynamical properties. Crucially, by better capturing the behavior in

*Contact author: trevor@extropic.ai

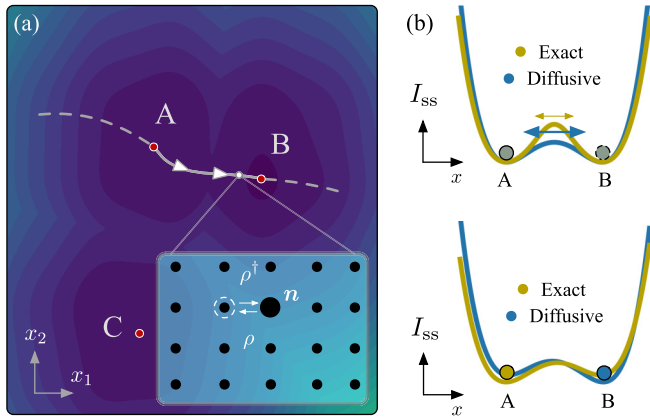


FIG. 1. Diffusive approximation of Markov jump processes. (a) The steady-state probability distribution P_{ss} for a two-dimensional stochastic process with three metastable states (A, B, and C). The white line indicates a transition path from A to B, driven by stochastic fluctuations. Inset: A “zoom” on the transition path reveals the underlying discrete Markov jump process. (b) A comparison of the effective energy landscape along the path from A to B, given by the steady-state self-information $I_{ss} := -\log P_{ss}$ for the exact process (yellow) and a typical diffusive approximation (blue). The following panels highlight common failure modes of standard diffusive approximations that our improved method is designed to address. Top panel: The approximation underestimates the barrier height between the two metastable states. Bottom panel: While standard diffusive approximations correctly capture local behavior around stable states, their inaccuracy for large fluctuations can lead to qualitatively wrong predictions, such as identifying the wrong most stable state (the global maximum of P_{ss}) [16].

the tails of higher-order eigenvectors of the generator of the jump process, our approximation provides a more faithful description of large fluctuations not just at steady state, but throughout the system’s transient evolution.

This article is organized as follows. Section II introduces both the standard diffusive approximation of Markov jump processes, namely, the Kramers-Moyal approximation, as well as the unconventional diffusion approximation that we study in this work. Section III then presents its derivation. In Sec. IV, we prove fundamental accuracy guarantees for the improved diffusion approximation using recent results in macroscopic stochastic thermodynamics. In Sec. V, we compare the performance of the usual Kramers-Moyal approximation and the improved approximation by applying them to several models of stochastic nonlinear electronic circuits. Finally, we summarize our results in Sec. VI and mention possible connections with recent work by other authors.

II. MARKOV JUMP PROCESSES AND DIFFUSIVE APPROXIMATIONS

In this section, we present our main theoretical result: an improved diffusive approximation for a general Markov jump process over states $\mathbf{n} \in \mathbb{Z}^d$ (see Fig. 1). For example, each state can represent the copy numbers of different species in chemical reaction networks [30,31] or the number of charges in each node of an electronic circuit [7,8]. The state of the

system evolves through transitions $\mathbf{n} \rightarrow \mathbf{n} + \Delta_\rho$, which have associated jump vector Δ_ρ and jump rate $\chi_\rho(\mathbf{n})$. Thus, transitions are indexed by $\rho \in \mathbb{Z}^d \setminus \{0\}$. Given any transition with $\rho > 0$, the reverse transition is indexed by $\rho^\dagger = -\rho$ and has jump vector $\Delta_{-\rho} = -\Delta_\rho$. For some initial condition, let $P(\mathbf{n}, t)$ be the probability of the system being in state $\mathbf{n}$ at time t . This probability evolves according to the master equation

$$\partial_t P(\mathbf{n}, t) = \sum_\rho [\chi_\rho(\mathbf{n} - \Delta_\rho) P(\mathbf{n} - \Delta_\rho, t) - \chi_\rho(\mathbf{n}) P(\mathbf{n}, t)]. \quad (1)$$

In chemical reaction network theory, Eq. (1) is the chemical master equation, the standard stochastic description in the low-copy-number regime where discreteness and intrinsic noise are essential. For systems in which the jump rates depend on some scale parameter Ω , we can define the density $\mathbf{x} = \mathbf{n}/\Omega$ and the density-dependent rates $\lambda_\rho(\mathbf{x}) \equiv \chi_\rho(\Omega\mathbf{x})$ [12,20]. Then, defining the rescaled jump vectors $\mathbf{x}_\rho \equiv \Delta_\rho/\Omega$, the master equation for the probability $P(\mathbf{x}, t)$ is

$$\partial_t P(\mathbf{x}, t) = \sum_\rho [\lambda_\rho(\mathbf{x} - \mathbf{x}_\rho) P(\mathbf{x} - \mathbf{x}_\rho, t) - \lambda_\rho(\mathbf{x}) P(\mathbf{x}, t)]. \quad (2)$$

At the ensemble level, general diffusive processes are described by Fokker-Planck equations of the form

$$\partial_t P(\mathbf{x}, t) = -\partial_n [\mu_n(\mathbf{x}) P(\mathbf{x}, t) - \partial_m (D_{nm}(\mathbf{x}) P(\mathbf{x}, t))], \quad (3)$$

in terms of a drift vector field $\mu_n(\mathbf{x})$ and a diffusion tensor field $D_{nm}(\mathbf{x})$. In the previous equation, Einstein summation is assumed and $\partial_n \equiv \partial/\partial x_n$. A generic method to obtain a diffusive approximation of a given MJP is given by the second-order truncation of the Kramers-Moyal expansion (Ref. [20], Sec. VIII.5). According to this method, the drift and diffusion fields are constructed as

$$\mu_n(\mathbf{x}) \equiv \frac{1}{\Omega} \sum_{\rho>0} (\lambda_\rho(\mathbf{x}) - \lambda_{-\rho}(\mathbf{x})) \Delta_\rho^n \quad (4)$$

and

$$D_{nm}^{\text{KM}}(\mathbf{x}) \equiv \frac{1}{2\Omega^2} \sum_{\rho>0} (\lambda_\rho(\mathbf{x}) + \lambda_{-\rho}(\mathbf{x})) \Delta_\rho^n \Delta_\rho^m, \quad (5)$$

where Δ_ρ^n is the n th component of Δ_ρ . As mentioned above, this approach fails to capture large fluctuations accurately. To make this statement more precise, let us define $P_{ss}^{\text{ME}}(\mathbf{x})$ as the steady-state distribution according to the master equation in Eq. (2), and $P_{ss}^{\text{KM}}(\mathbf{x})$ as the steady-state distribution according to Eqs. (3)–(5). Under the assumption that the transition rates $\lambda_\rho(\mathbf{x})$ are extensive, it can be seen that both $P_{ss}^{\text{ME}}(\mathbf{x})$ and $P_{ss}^{\text{KM}}(\mathbf{x})$ decrease exponentially in the limit $\Omega \rightarrow \infty$. This means that the rate functions $I_{ss}^{\text{ME/KM}}(\mathbf{x}) \equiv \lim_{\Omega \rightarrow \infty} -\log(P_{ss}^{\text{ME/KM}}(\mathbf{x}))/\Omega$ are well defined and non-negative. Furthermore, it can be shown [23] that the functions $I_{ss}^{\text{ME}}(\mathbf{x})$ and $I_{ss}^{\text{KM}}(\mathbf{x})$ share the same local minima (the most probable values of $\mathbf{x}$) and actually match each other to second order around those minima (small Gaussian fluctuations, see Fig. 1, top panel). However, it is also possible to see that $I_{ss}^{\text{ME}}(\mathbf{x}) \neq I_{ss}^{\text{KM}}(\mathbf{x})$ in general, which means that the approximation given by Eqs. (4) and (5) fails to capture the

probability of large steady-state fluctuations, even to dominant order in the macroscopic limit. In Fig. 1(b), we illustrate an example in which the height of the barrier separating two minima of $I_{ss}(\mathbf{x})$, which is related to the rate of switching transitions, is underestimated by the Kramers-Moyal approximation. More seriously, even if both functions share the same local minima, it is known that their global minima can be different, which leads to incompatible predictions in the macroscopic limit [see the bottom panel of Fig. 1(b)].

It is useful to contrast our setting with two related lines of work. First, in studies of stochastic transport with tilted periodic potentials (e.g., molecular motors), the underlying dynamics is taken to be continuous (Smoluchowski/Langevin), and discrete Kramers-type descriptions are used as surrogates [32–34]. Our problem is the complementary direction: Given an MJP, we construct a diffusion that preserves its large-deviation structure (exactly at equilibrium and to first order out of equilibrium). Second, there is a body of work starting from generalized (often non-Markovian) Langevin dynamics that introduces effective temperatures or chemical potentials to extend fluctuation-dissipation relations far from equilibrium [35]. In contrast, we start from a discrete, locally detailed-balanced MJP and, in the large-system-size limit, derive a modified diffusion tensor that preserves its large-deviation structure (exactly at equilibrium and to first order out of equilibrium).

The main result in this article is that, for a large class of MJPs, one must instead consider the following diffusion tensor [36]:

$$D_{nm}(\mathbf{x}) \equiv \frac{1}{\Omega^2} \sum_{\rho>0} \frac{\lambda_{\rho}(\mathbf{x}) - \lambda_{-\rho}(\mathbf{x})}{\log(\lambda_{\rho}(\mathbf{x})/\lambda_{-\rho}(\mathbf{x}))} \Delta_{\rho}^n \Delta_{\rho}^m. \quad (6)$$

Equation (6) is the generalization to arbitrary MJPs of a result first obtained in Refs. [21–23] for one-dimensional (1D) systems with nearest-neighbor jumps. There, it was also shown that the resulting diffusive dynamics has exactly the same macroscopic steady-state fluctuations as the original MJP under strict detailed balance (which is always satisfied in such 1D systems). In this article, we leverage a connection between the rate function of non-equilibrium steady states (NESSs) and the entropy produced along deterministic trajectories [37–40] to prove that Eq. (6) also reproduces steady-state macroscopic fluctuations to linear order in departures from strict detailed balance. Finally, we also note that the quantity $\frac{\lambda_{\rho}(\mathbf{x}) - \lambda_{-\rho}(\mathbf{x})}{\log(\lambda_{\rho}(\mathbf{x})/\lambda_{-\rho}(\mathbf{x}))}$ that enters in Eq. (6) has recently been identified in Refs. [41,42] to play a crucial role in the connection between stochastic thermodynamics and optimal transport theory. As explained in those works, it can be interpreted as a microscopic and nonlinear version of the usual Onsager response coefficients.

The theory behind Eq. (6) is tailored to accurately reproduce steady-state fluctuations, i.e., to improve the correspondence between the zero modes of the generators of the master equation in Eq. (2) and the Fokker-Planck dynamics in Eq. (3). However, as we show, it also improves the correspondence between higher-order eigenmodes and eigenvalues of both dynamics, compared to the regular diffusive approximation based on Eq. (5). In particular, we show that our theory improves the estimation of the error rate of a bistable

electronic memory, which is given by the first nonzero eigenvalue of the MJP generator [28].

III. DERIVATION OF THE MODIFIED DIFFUSION APPROXIMATION

A. Effective Fokker-Planck equation

In order to derive Eq. (6), our strategy will be to construct a drift vector $\mu_n(\mathbf{x})$ and a diffusion tensor $D_{nm}(\mathbf{x})$ such that the dynamics in Eq. (3) mimics that of Eq. (2). The first step in our derivation is to recast the master equation in Eq. (2) in the form of a Fokker-Planck equation, but for which the diffusion tensor depends on the probability distribution itself. Indeed, with some manipulation of Eq. (2), we obtain the exact equation of motion

$$\partial_t P(\mathbf{x}, t) = \sum_{\rho} \left[\sum_{k=1}^{\infty} \frac{1}{k!} \left(\frac{-\Delta_{\rho} \cdot \nabla}{\Omega} \right)^k (\lambda_{\rho} P) \right] \quad (7)$$

$$= -\partial_n [\mu_n(\mathbf{x}) P(\mathbf{x}, t) - \partial_m (\tilde{D}_{nm}(\mathbf{x}, t) P(\mathbf{x}, t))], \quad (8)$$

where (see Ref. [43] for a brief derivation)

$$\mu_n(\mathbf{x}) \equiv \frac{1}{\Omega} \sum_{\rho} \lambda_{\rho}(\mathbf{x}) \Delta_{\rho}^n, \quad (9)$$

$$\tilde{D}_{nm}(\mathbf{x}, t) \equiv \frac{1}{\Omega^2} \sum_{\rho} F_{\rho}(\mathbf{x}, t) \Delta_{\rho}^n \Delta_{\rho}^m \quad (10)$$

with

$$F_{\rho}(\mathbf{x}, t) \equiv e^{\mathcal{I}(\mathbf{x}, t)} \sum_{k \geq 2} \frac{1}{k!} \left(\frac{-\Delta_{\rho} \cdot \nabla}{\Omega} \right)^{k-2} [\lambda_{\rho}(\mathbf{x}) e^{-\mathcal{I}(\mathbf{x}, t)}]. \quad (11)$$

In the previous equation, we have introduced the self-information or surprisal associated with the distribution $P(\mathbf{x}, t)$:

$$\mathcal{I}(\mathbf{x}, t) \equiv -\log(P(\mathbf{x}, t)). \quad (12)$$

The prefactor $e^{\mathcal{I}(\mathbf{x}, t)} \equiv P(\mathbf{x}, t)^{-1}$ in Eq. (11) is included to balance the multiplication of $\tilde{D}_{nm}(\mathbf{x}, t)$ by $P(\mathbf{x}, t)$ in Eq. (8), thereby ensuring the effective Fokker-Planck equation remains an exact representation of the master equation. Note that Eq. (7) cannot be interpreted as a regular Fokker-Planck equation since the function $\tilde{D}_{nm}(\mathbf{x})$ depends implicitly on the probability distribution itself. We now discuss different approximations that can be performed on Eq. (11) in order to reduce Eq. (7) to a regular Fokker-Planck dynamics.

1. Kramers-Moyal truncation

In the first place, we notice that if only the $k = 2$ term in the summation of Eq. (11) is kept, then the diffusion tensor $\tilde{D}_{nm}(\mathbf{x})$ becomes independent of $P(\mathbf{x}, t)$ and reduces to

$$D_{nm}^{\text{KM}}(\mathbf{x}) \equiv \frac{1}{2} \sum_{\rho} \lambda_{\rho}(\mathbf{x}) \frac{\Delta_{\rho}^n \Delta_{\rho}^m}{\Omega^2}. \quad (13)$$

In this case, Eq. (7) coincides with the Fokker-Planck equation obtained by the usual truncation to second order of the Kramers-Moyal expansion [20,44]. As mentioned before, this

method only captures small fluctuations around the most probable values of $\mathbf{x}$.

2. Steady-state matching

We now consider situations in which the master equation in Eq. (2) has a unique steady state $P_{ss}(\mathbf{x})$. Then, one possible approximation is to replace $\mathcal{I}(\mathbf{x}, t)$ in Eq. (11) with the steady-state self-information $\mathcal{I}_{ss}(\mathbf{x}) = -\log(P_{ss}(\mathbf{x}))$, obtaining

$$F_{\rho}^{SS}(\mathbf{x}) \equiv e^{\mathcal{I}_{ss}(\mathbf{x})} \sum_{k \geq 2} \frac{1}{k!} \left(\frac{-\Delta_{\rho} \cdot \nabla}{\Omega} \right)^{k-2} [\lambda_{\rho}(\mathbf{x}) e^{-\mathcal{I}_{ss}(\mathbf{x})}]. \quad (14)$$

We can then treat this approximation as a time-independent ansatz for the diffusion tensor, $\tilde{D}(\mathbf{x}, t) \equiv \tilde{D}(\mathbf{x})$ appearing in Eq. (8). The regular Fokker-Planck dynamics that results from this approximation is guaranteed to have exactly the same steady-state distribution as the original Markov jump process. However, this ansatz is hardly useful, since the steady-state distribution is unknown in most cases. Some exceptions are one-dimensional jump processes with nearest-neighbor transitions [23] or the kind of detailed-balanced processes discussed next.

3. Strict detailed balance

If the transition rates $\lambda_{\rho}(\mathbf{x})$ satisfy the detailed-balance conditions,

$$\log \frac{\lambda_{\rho}(\mathbf{x})}{\lambda_{-\rho}(\mathbf{x} + \mathbf{x}_{\rho})} = -(\Phi(\mathbf{x} + \mathbf{x}_{\rho}) - \Phi(\mathbf{x})), \quad (15)$$

in terms of some state function $\Phi(\mathbf{x})$, then the steady state of the Markov jump process is given by the Gibbs equilibrium distribution

$$P_{eq}(\mathbf{x}) = \frac{e^{-\Phi(\mathbf{x})}}{Z}, \quad (16)$$

with $Z = \sum_{\mathbf{x}} e^{-\Phi(\mathbf{x})}$. Therefore, $\mathcal{I}_{ss}(\mathbf{x}) = \Phi(\mathbf{x}) + \log(Z)$. Introducing this in Eq. (14), we obtain the detailed-balance result

$$F_{\rho}^{DB}(\mathbf{x}) \equiv F^{SS}(\mathbf{x})|_{\mathcal{I}_{ss}=\Phi+\log(Z)}. \quad (17)$$

Having constrained the parameters to yield identical steady-state distributions, we can now probe the deeper similarities and differences between the models. To this end, we turn in Sec. IV A to the linear response framework, which provides the tools to analyze both the dynamical response and key statistical features of the steady state itself, including the self-information.

B. Scaling limit

The expression in Eq. (11) is exact. We now discuss how Eq. (11) is simplified by taking the limit $\Omega \rightarrow \infty$. We assume that the jump rates λ_{ρ} are extensive quantities and expand $\lambda_{\rho}(\mathbf{x}) \equiv \Omega \omega_{\rho}(\mathbf{x}) + o(\Omega)$. In that case, it can be shown that the distribution $P(\mathbf{x}, t)$ satisfies a large-deviation principle, which means that the self-information $\mathcal{I}(\mathbf{x}, t)$ is also extensive [45]. Thus, we consider the expansion

$$\mathcal{I}(\mathbf{x}, t) = \Omega I(\mathbf{x}, t) + o(\Omega). \quad (18)$$

One can then show that the function $F_{\rho}(\mathbf{x})$ involved in Eq. (10) admits an identical expansion

$$F_{\rho}(\mathbf{x}) = \Omega f_{\rho}(\mathbf{x}) + o(\Omega). \quad (19)$$

As shown in Sec. III of the Supplemental Material [46], $f_{\rho}(\mathbf{x})$ then takes the form

$$\begin{aligned} f_{\rho}(\mathbf{x}) &= \omega_{\rho}(\mathbf{x}) \sum_{k \geq 2} \frac{1}{k!} (\Delta_{\rho} \cdot \nabla I(\mathbf{x}, t))^{k-2} \\ &= \omega_{\rho}(\mathbf{x}) \frac{e^{\Delta_{\rho} \cdot \nabla I(\mathbf{x}, t)} - 1 - \Delta_{\rho} \cdot \nabla I(\mathbf{x}, t)}{(\Delta_{\rho} \cdot \nabla I(\mathbf{x}, t))^2}. \end{aligned} \quad (20)$$

It is instructive to specialize the last expression for the case of equilibrium matching. For this, we expand $\Phi(\mathbf{x}) = \Omega \phi(\mathbf{x}) + o(\Omega)$. Then, in the scaling limit $\Omega \rightarrow \infty$, the detailed-balance conditions in Eq. (15) read as

$$\log \frac{\omega_{\rho}(\mathbf{x})}{\omega_{-\rho}(\mathbf{x})} = -\Delta_{\rho} \cdot \nabla \phi(\mathbf{x}). \quad (21)$$

Therefore, we have

$$\Delta_{\rho} \cdot \nabla \mathcal{I}_{ss}(\mathbf{x}) = \Delta_{\rho} \cdot \nabla \phi(\mathbf{x}) = -\log \frac{\omega_{\rho}(\mathbf{x})}{\omega_{-\rho}(\mathbf{x})}. \quad (22)$$

Using this, it is easy to show that

$$[f_{\rho}(\mathbf{x}) + f_{-\rho}(\mathbf{x})]^{DB} = \frac{\omega_{\rho}(\mathbf{x}) - \omega_{-\rho}(\mathbf{x})}{\log(\omega_{\rho}(\mathbf{x})/\omega_{-\rho}(\mathbf{x}))}. \quad (23)$$

This result implies that employing the diffusion tensor

$$D_{nm}^{DB}(\mathbf{x}) \equiv \frac{1}{\Omega^2} \sum_{\rho > 0} \frac{\lambda_{\rho}(\mathbf{x}) - \lambda_{-\rho}(\mathbf{x})}{\log(\lambda_{\rho}(\mathbf{x})/\lambda_{-\rho}(\mathbf{x}))} \Delta_{\rho}^n \Delta_{\rho}^m \quad (24)$$

in Eq. (3) corresponds to a diffusive process that is guaranteed to have the same macroscopic fluctuations as the original Markov jump process.

The previous result can also be stated in connection to the Einstein-Smoluchowski relation for general diffusive dynamics. This relation just expresses the fact that for a Fokker-Planck dynamics like the one in Eq. (3) to relax to the Gibbs distribution corresponding to a potential $\Phi(\mathbf{x})$, the drift vector $\mu_n(\mathbf{x})$ and diffusion tensor $D_{nm}(\mathbf{x})$ must be related to each other by the identity $\mu_n(\mathbf{x}) = -D_{nm}(\mathbf{x}) \partial_m \Phi(\mathbf{x}) + \partial_m D_{mn}(\mathbf{x})$. Indeed, from Eq. (24) it follows that

$$\begin{aligned} \lim_{\Omega \rightarrow \infty} (\mu_n(\mathbf{x}) + D_{nm}^{DB}(\mathbf{x}) \partial_m \Phi(\mathbf{x}) - \partial_m D_{mn}^{DB}(\mathbf{x})) \\ = \sum_{\rho > 0} \left[(\omega_{\rho} - \omega_{-\rho}) + \frac{\omega_{\rho} - \omega_{-\rho}}{\log(\omega_{\rho}/\omega_{-\rho})} (\Delta_{\rho} \cdot \nabla \phi) \right] \Delta_{\rho}^n = 0, \end{aligned} \quad (25)$$

where in the last line we used the macroscopic limit of the detailed-balance conditions [Eq. (21)]. We note that the usual Kramers-Moyal diffusion tensor in Eq. (13) fails to satisfy the previous condition. In particular, it is well documented that the Kramers-Moyal diffusion tensor can yield spurious steady currents and hence a Fokker-Planck dynamics that *violates* detailed balance even when the underlying MJP is at equilibrium [47,48]. In contrast, Eq. (25) demonstrates that our construction maps equilibrium MJPs to equilibrium diffusions, a property that is highly desirable for any diffusion approximation.

We now show in what follows that the diffusion tensor in Eq. (24) also leads to the correct macroscopic fluctuations for first-order deviations from perfect detailed-balance conditions. This is a crucial result of our theory since it significantly extends its scope of validity.

IV. DYNAMICAL LARGE-DEVIATION THEORY AND LINEAR-RESPONSE REGIME

In this section, we consider the linear-response regime around detailed-balance conditions and sketch the main ingredients of the proof showing that the previous results are still valid in that regime. We first collect relevant results concerning the equations of motion satisfied by the large-deviation rate functions $I(\mathbf{x}, t)$, introduced in Eq. (18). For Markov jump processes, the rate function evolves according to the equation of motion [45]

$$\partial_t I(\mathbf{x}, t) = \sum_{\rho} \omega_{\rho}(\mathbf{x}) (1 - e^{\Delta_{\rho} \cdot \nabla I(\mathbf{x}, t)}). \quad (26)$$

On the other hand, the rate function for a drift-diffusion process evolves according to the equation of motion [40]

$$\partial_t I(\mathbf{x}, t) = u_n(\mathbf{x}) \partial_n I(\mathbf{x}, t) + d_{nm}(\mathbf{x}) \partial_n I(\mathbf{x}, t) \partial_m I(\mathbf{x}, t), \quad (27)$$

which is written in terms of the following macroscopic versions of the drift vector and diffusion tensor [40]:

$$u_n(\mathbf{x}) \equiv \lim_{\Omega \rightarrow \infty} \mu_n(\mathbf{x}), \quad (28)$$

$$d_{nm}(\mathbf{x}) \equiv \lim_{\Omega \rightarrow \infty} \Omega D_{nm}(\mathbf{x}). \quad (29)$$

To dominant order in Ω and for detailed-balance conditions [Eq. (21)], the Einstein-Smoluchowski relation in Eq. (25) reads as the following relation between the above macroscopic versions of the drift vector and diffusion tensor:

$$u(\mathbf{x}) = -d^{\text{EQ}}(\mathbf{x}) \cdot \nabla \phi(\mathbf{x}). \quad (30)$$

In particular, we can substitute the above into Eq. (27) and confirm that $I_{\text{eq}}(\mathbf{x}) = \phi(\mathbf{x})$ is a steady-state solution of both equations of motion. Therefore, confirming the previous results, our diffusion approximation yields the correct macroscopic steady-state potential at equilibrium.

A. Linear response

In the linear-response regime, we consider small departures from the global detailed-balance condition. In particular, the detailed-balance conditions for the underlying MJP in Eq. (21) are generalized to the *local* detailed-balance condition

$$\log \frac{\omega_{\rho}(\mathbf{x})}{\omega_{-\rho}(\mathbf{x})} = -\Delta_{\rho} \cdot \nabla \phi(\mathbf{x}) + \epsilon w_{\rho}(\mathbf{x}), \quad (31)$$

for $w_{\rho}(\mathbf{x})$ some function of both the state and the transition, and ϵ a perturbative parameter that allows us to continuously tune to the detailed-balanced limit $\epsilon \rightarrow 0$. In Sec. I of the Supplemental Material [46], we show that, under mild conditions on $u(\mathbf{x})$, the steady-state rate function $I_{\text{ss}}(\mathbf{x})$ for the diffusive approximation of Eq. (6) agrees with the corresponding rate function of the underlying Markov jump process to first order in ϵ . The proof exploits recent results on macroscopic stochastic thermodynamics [39] and can be intuitively

understood as follows. As shown in Refs. [37,38] for MJPs and in Ref. [40] for diffusive processes, the steady-state rate function can be directly related to the macroscopic entropy production rate in the linear-response regime. In both cases, this relation reads

$$d_t I_{\text{ss}}(\mathbf{x}_t) = -\dot{\sigma}(\mathbf{x}_t) + \mathcal{O}(\epsilon^2), \quad (32)$$

where $\dot{\sigma}(\mathbf{x})$ is the scaled macroscopic entropy production rate and $\mathbf{x}_t$ is any trajectory satisfying the deterministic equation of motion $d_t \mathbf{x}_t = u(\mathbf{x}_t)$. The difference between MJPs and diffusive processes lies in their respective expressions for the entropy production rate $\dot{\sigma}(\mathbf{x})$. For MJPs, we have

$$\dot{\sigma}_{\text{MJP}}(\mathbf{x}) = \sum_{\rho>0} (\omega_{\rho}(\mathbf{x}) - \omega_{-\rho}(\mathbf{x})) \log \frac{\omega_{\rho}(\mathbf{x})}{\omega_{-\rho}(\mathbf{x})}, \quad (33)$$

while for diffusive processes

$$\dot{\sigma}_{\text{diff}}(\mathbf{x}) = u_n(\mathbf{x}) d_{nm}(\mathbf{x}) u_m(\mathbf{x}). \quad (34)$$

It is possible to show that if $d_{nm}(\mathbf{x})$ follows from Eq. (6), then $\int dt \sigma_{\text{MJP}}(\mathbf{x}_t) = \int dt \sigma_{\text{diff}}(\mathbf{x}_t) + \mathcal{O}(\epsilon^2)$ for any deterministic trajectory $\mathbf{x}_t$, which then implies that $I_{\text{ss}}^{\text{MJP}}(\mathbf{x}) = I_{\text{ss}}^{\text{diff}}(\mathbf{x}) + \mathcal{O}(\epsilon^2)$ (see Sec. I of the Supplemental Material for an explicit perturbative proof [46]). Therefore, we conclude that $I_{\text{ss}}^{\text{MJP}}(\mathbf{x})$ and $I_{\text{ss}}^{\text{diff}}(\mathbf{x})$ match each other to first order in ϵ . Of course, this is not the case for a diffusive process based on the usual Kramers-Moyal approximation of Eq. (5), since in that case both rate functions differ even for $\epsilon = 0$.

V. EXAMPLES

We now focus on example systems with which to compare our diffusive approximation with the standard Kramers-Moyal expansion. We will see in each example that the improved diffusive approximation more accurately captures phenomena associated with rare events/large steady-state fluctuations. Throughout this section, we will focus on examples of Markov jump processes coming from electronic circuits that commonly appear in complementary metal-oxide semiconductor (CMOS) technology, but our formalism is more general and applies to generic Markov jump processes, e.g., chemical reaction networks, random walks, etc.

A. One-dimensional jump processes

Consider a one-dimensional Markov jump process with nearest-neighbor jumps $\Delta_{\pm} = \pm 1$. In this simple case, we can analyze the performance of both diffusive approximations analytically. Indeed, with only two distinct jump vectors, the sum in Eq. (26) has only two terms, and we immediately obtain the solution $I(x) = -\int dx \log \frac{\omega_{+}}{\omega_{-}}$. On the other hand, for a one-dimensional drift-diffusion process, Eq. (27) admits the steady-state solution $I(x) = -\int dx \frac{u}{d}$. The diffusion tensor for the improved diffusive approximation satisfies $d = -u \log(\omega_{+}/\omega_{-})$ and thus [23]

$$\frac{d}{dx} I_{\text{IMP}}(x) = -\log \frac{\omega_{+}(x)}{\omega_{-}(x)}, \quad (35)$$

which implies that, in this simple case, the steady-state large-deviation function obtained by the improved diffusion approximation is exact. The reason is that this kind of 1D

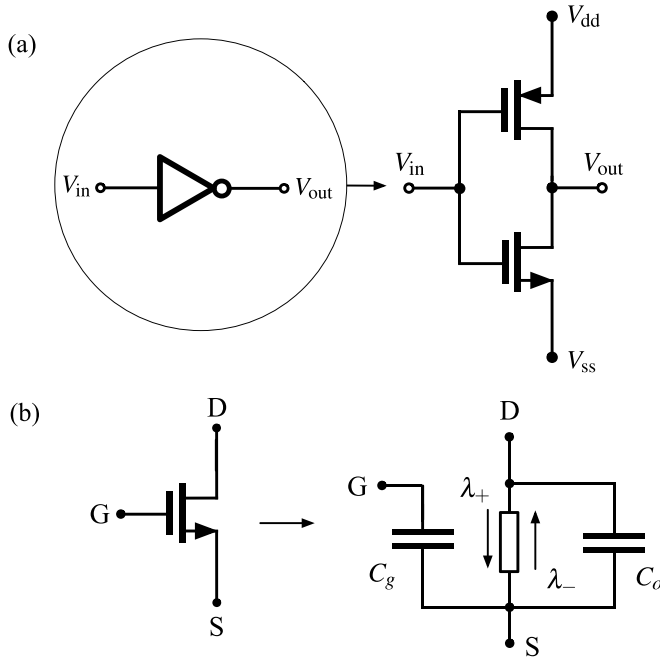


FIG. 2. Stochastic CMOS circuits. (a) In the context of digital electronics, a logical not gate (left) is typically implemented as a CMOS inverter (right), which employs two field-effect transistors. In the examples that follow, $V_{in} = 0$ and $V_{dd} = -V_{ss} = V_T$, with $V_T \equiv k_B T / e$ the thermal voltage. (b) Each transistor is modeled in the subthreshold regime as a stochastic conduction channel between its source (S) and drain (D) terminals, with voltage-dependent Poisson rates $\lambda_{+/-}$, cf. Ref. [8] for a detailed discussion. For the circuit examples used in this paper, $C_o = 2C_T$ and $C_g = 4C_T$, where $C_T \equiv e^2 / V_T$.

system can always be considered to be detailed balance. In contrast, repeating the same exercise for the Kramers-Moyal diffusion approximation yields

$$\frac{d}{dx} I_{KM}(x) = -2 \frac{\omega_+(x) - \omega_-(x)}{\omega_+(x) + \omega_-(x)}. \quad (36)$$

These results are analogous to those obtained in Refs. [22,23]. To quantify precisely by how much the Kramers-Moyal approximation breaks down, we expand both rate functions near a deterministic fixed point x^* , for which $\omega_+(x^*) = \omega_-(x^*)$:

$$\frac{dI_{IMP}}{dx} - \frac{dI_{KM}}{dx} = \sum_{k>0} \left(\frac{1}{k} - \frac{1}{2^{k-1}} \right) \left(1 - \frac{\omega_+}{\omega_-} \right)^k. \quad (37)$$

Because the above difference is of cubic order in $1 - \omega_+/\omega_-$, to see a difference between the two diffusive approximations we must evaluate the steady-state distribution far away from any deterministic fixed point. This means that we must look at the tails of the distribution to see a difference between the approximations. We will see this behavior in the example discussed below.

1. CMOS inverter

As a basic example of a one-dimensional jump process, we consider a CMOS inverter [cf. Fig. 2(a)]. In the inverter circuit, n - and p -type field-effect transistors are connected in series and a common input voltage V_{in} is applied to the gate of

each transistor (in this case $V_{in} = 0$). When both transistors are operated in the subthreshold regime, the dimensionless voltage $x \equiv V_{out}/V_T$ at the output node (in units of the thermal voltage $V_T \equiv k_B T / e$) can be modeled as a Markov jump process with nearest-neighbor jumps $\Delta_{\pm} = \pm 1$ and jump rates

$$\begin{aligned} \lambda_+(x) &= \frac{1}{\tau} \left(e^{-(x+\Delta x)} e^{-(1+1/n)V_{dd}/V_T} + e^{+V_{dd}/nV_T} \right), \\ \lambda_-(x) &= \frac{1}{\tau} \left(e^{+(x-\Delta x)} e^{-(1-1/n)V_{dd}/V_T} + e^{-V_{dd}/nV_T} \right). \end{aligned} \quad (38)$$

Here, $\Delta x \equiv (2\Omega)^{-1}$ is half the change in the dimensionless voltage x after a single jump and n is a constant called the subthreshold slope. Furthermore, V_{dd} is the powering voltage of the logic gate [cf. Fig. 2(a)] and τ a time constant. In this context of stochastic circuits, Ω is the smallest eigenvalue of the dimensionless matrix C/C_T , where C is the Maxwell capacitance matrix for the circuit, with $C_T \equiv e^2 / k_B T$ an elementary unit of capacitance.

Issues with the Kramers-Moyal approximation in accurately capturing large stationary voltage fluctuations in this circuit were noted in Ref. [49]. However, this observation was limited to the steady state and did not cover quantities relevant to transient dynamics, such as higher-order eigenmodes. In our numerical analysis of this example (as well as all following examples), we analyze the full spectral properties of the Liouvillian $\mathcal{L}$ for our Markov jump process and its diffusive approximations. In terms of the spectral decomposition of $\mathcal{L}$, the evolution of an initial distribution $P(\mathbf{x}, 0) = \sum_k c_k v_k(\mathbf{x})$ takes the form

$$P(\mathbf{x}, t) = \sum_k c_k e^{-\lambda_k t} v_k(\mathbf{x}), \quad (39)$$

where λ_k and $v_k(\mathbf{x})$ are the eigenvalues and right eigenmodes of $\mathcal{L}$. To solve this problem numerically, we truncate the state space to a finite box $x \in [-L, L]^d$, where L is some large cutoff voltage, and impose reflecting boundary conditions on $\mathcal{L}$. For diffusive dynamics, we also discretize the state space using a sufficiently fine grid (see Sec. IV of the Supplemental Material for more details [46]). We then use a sparse solver to obtain the low-lying eigenmodes and eigenvalues v_k, λ_k for the resulting finite-dimensional problem. We then compare the results with the eigenvectors and eigenvalues $\lambda_k^{\text{exact}}, v_k^{\text{exact}}$ of the Liouvillian for the original Markov jump process.

In Fig. 3(a), we plot the three slowest decaying eigenmodes v_0, v_1, v_2 of the Liouvillian for the inverter circuit. In confirmation of our analytic arguments, we observe that the improved diffusive approximation more accurately captures the tails of the steady-state distribution far away from the deterministic fixed point $x^* = 0$. However, from the numerics, we also observe improvements in the tails of fast-decaying eigenmodes above the steady state, e.g., v_{10}, v_{15}, v_{20} [cf. Fig. 3(b)], a result that goes beyond the predictions of our analytic arguments.

B. Probabilistic bit (CMOS implementation)

We now consider an example of a two-dimensional system that, at variance with the 1D examples above and in Refs. [22,23], is genuinely nondetailed balanced. It consists of a bistable circuit formed by two CMOS inverters

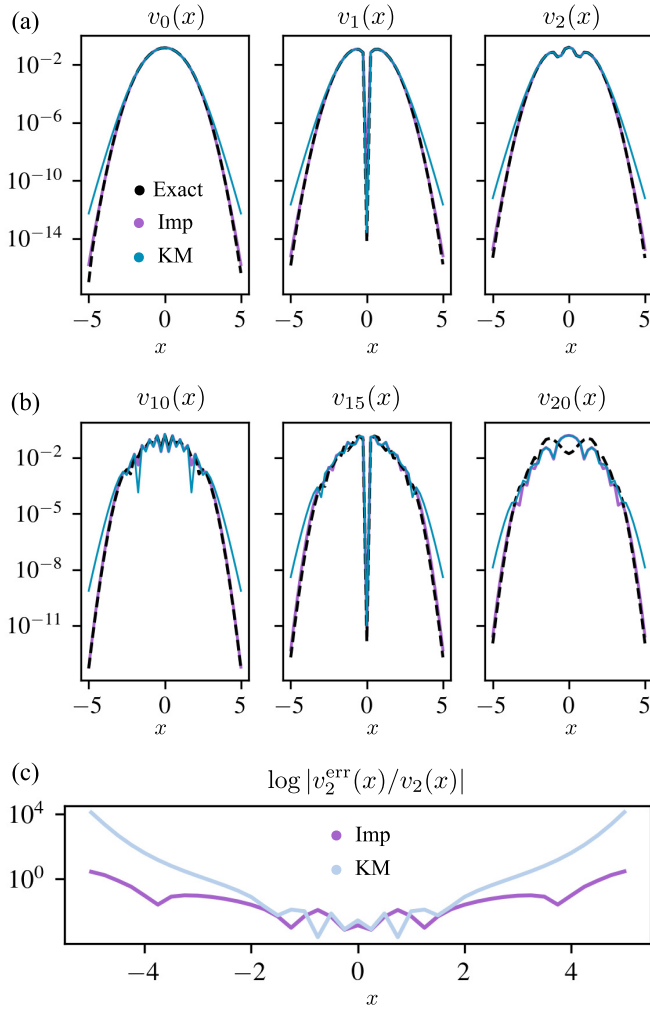


FIG. 3. Liouvillian eigenmodes of a CMOS inverter under different diffusive approximations. (a) Three slowest-decaying eigenvectors for the Liouvillian of the inverter circuit in Fig. 2(a), plotted together with the three slowest-decaying eigenmodes of the Liouvillian for both the Kramers-Moyal approximation and the improved diffusion approximation. (b) Representative midspectrum eigenmodes (11th, 16th, and 21st slowest) for the same system. (c) Fractional eigenmode error $v_2^{\text{err}}(x) \equiv v_2(x) - v_2^{\text{exact}}(x)$, for both the Kramers-Moyal and improved diffusive approximations. Parameters are the same as in panel (a).

connected in feedback [cf. Fig. 4(a)]. Again, we use the model developed in Ref. [8], under which the normalized voltages $\mathbf{x} \equiv (V_0/V_T, V_1/V_T)$ at the output terminals of the circuit evolve according to a Markov jump process with jump vectors $\Delta_{j,\pm}^n = \pm \delta_{j,n}$, with corresponding jump rates [28]

$$\lambda_{0,\pm}(\mathbf{x}) = \frac{1}{\tau} (e^{V_{\text{dd}}/nV_T \mp x_1/n} + e^{\pm x_1/n} e^{\mp(x_0 \pm \Delta x)}), \quad (40)$$

$$\lambda_{1,\pm}(\mathbf{x}) = \frac{1}{\tau} (e^{V_{\text{dd}}/nV_T \mp x_0/n} + e^{\pm x_0/n} e^{\mp(x_1 \pm \Delta x)}). \quad (41)$$

Here, V_{dd} is the powering voltage of the inverters in the memory cell (cf. Fig. 4), and as before τ is a time constant.

This model has a transition from a monostable to a bistable steady state at the critical powering bias $V_{\text{dd}} = \log(2)V_T$ [28]. In Fig. 5(a) we plot, as a function of the powering voltage

and for different scale parameters Ω , the Kullback-Leibler (KL) divergence of the steady-state distribution of each diffusive approximation with respect to the exact steady state of the MJP. We see that the improved diffusion approximation results in a lower KL divergence for all powering voltages and scale parameters. Notably, in the bistable regime $V_{\text{dd}} > \log(2)V_T$, the KL divergence increases with the powering voltage for the Kramers-Moyal approximation, while it levels off for the improved diffusion approximation. Thus, the improved diffusion approximation leads to a better description of steady-state fluctuations even for non-detailed-balance conditions.

We now turn our attention to the behavior of the smallest nonzero eigenvalue λ_1 of the Liouvillian of both diffusive approximations. In the bistable phase, this eigenvalue gives the rate of thermally induced switching transitions between the two metastable states, i.e., the rate of logical errors in the memory. In Figs. 5(b) and 5(c), we show the relative difference of λ_1 with respect to the corresponding eigenvalue of the MJP. While the Kramers-Moyal approximation shows a better performance in the monostable regime $V_{\text{dd}} < \log(2)V_T$, the improved diffusion approximation shows a much slower increase of the error in the bistable regime $V_{\text{dd}} > \log(2)V_T$. Furthermore, as we show in more detail in Fig. 6, the performance of the improved diffusion approximation becomes insensitive to the scale. In contrast, the performance of the Kramers-Moyal approximation degrades as one goes deeper in the macroscopic limit. As we show below, this can be understood in terms of a generalization of the Eyring-Kramers formula to NESSs [50].

1. Reduced rate function and metastable lifetime of the memory

The bistable voltage states for this circuit lie along the line $x_1 = -x_0$ [8]. The rate function for this circuit for CMOS memory can be restricted to this line and admits the solution [28]

$$I(x) = -2 \int dx \log \frac{\omega(x)}{\omega(-x)}, \quad (42)$$

where $I(x) \equiv I(x, -x)$ and $\omega(x) \equiv \omega_{0,+}(x, -x)$. In deriving this identity, one uses the symmetries that relate the jump rates for this circuit, as well as the fact that the restricted rate function $f(y) \equiv I(y+x, y-x)$ always attains a global minimum at $y=0$.

We can apply the same assumptions to the rate function of any diffusive approximation for the dynamics of the circuit, in which case we find that Eq. (27) simplifies to

$$u_{\text{eff}} \frac{dI}{dx} + d_{\text{eff}} \left(\frac{dI}{dx} \right)^2 = 0, \quad (43)$$

where $u_{\text{eff}} = (u_0 - u_1)/2$ and $d_{\text{eff}} = (d_{00} + d_{11})/4$ are effective drift and diffusion coefficients along the bistable axis $y \equiv 0$ of the memory. Substituting in the objects from the improved diffusion approximation, we obtain the solution

$$\frac{dI_{\text{IMP}}}{dx} = -2 \log \frac{\omega(x)}{\omega(-x)}, \quad (44)$$

which means that I_{IMP} and I exactly coincide along the bistable line $x_1 = -x_0$. Repeating this calculation

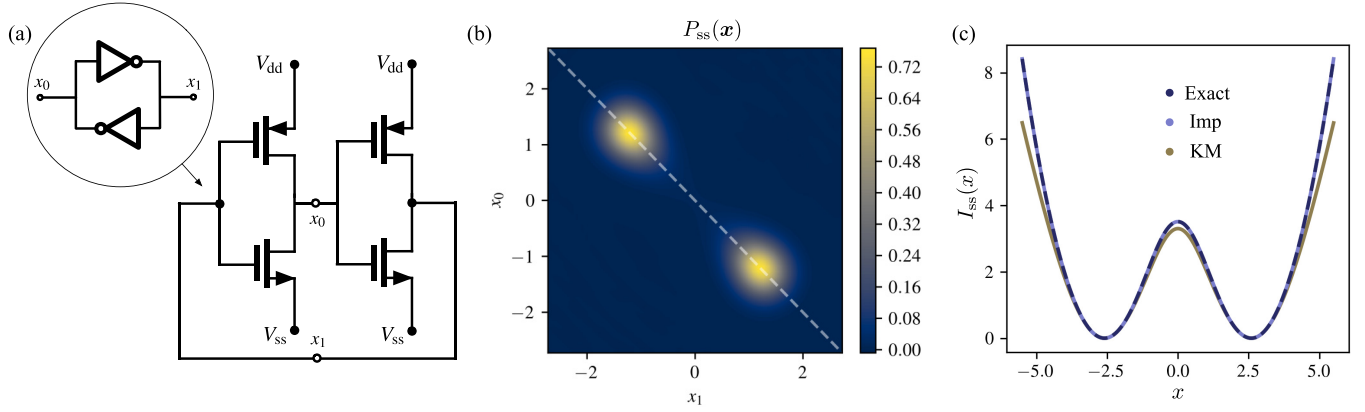


FIG. 4. (a) Circuit diagram for a random-access memory (RAM) cell formed by two CMOS inverters connected in a feedback loop (left inset). The stochastic degrees of freedom here are the voltages x_0, x_1 at the two output terminals of the circuit. We consider $V_{dd} = -V_{ss}$ for the powering voltages. (b) Steady-state probability density for the circuit in panel (a), under the powering configuration $V_{dd} = -V_{ss} = 1.3V_T$, and with $\Omega = 11$. The bistable axis $x_1 \equiv -x_0$ for the memory is marked with a dashed line. (c) Cut along the bistable axis of the large-deviation function $I_{ss}(x)$ for the circuit. Here, $V_{dd} = -V_{ss} = 2.6V_T$. We show results obtained from both the Kramers-Moyal approximation as well as the improved diffusion approximation.

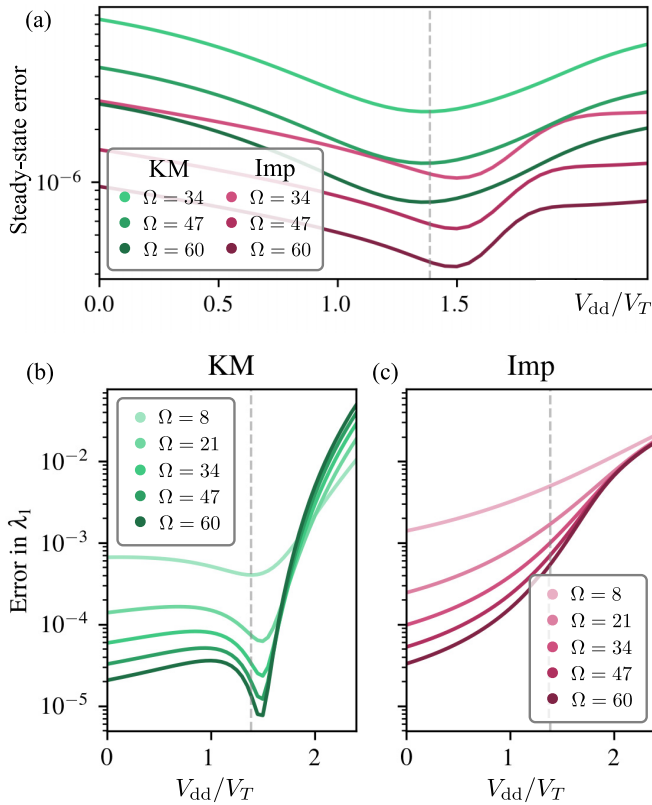


FIG. 5. Performance of improved diffusion approximation applied to a CMOS memory. (a) Error (as measured by the Kullback-Leibler divergence) in estimating the nonequilibrium steady state of a CMOS RAM cell [cf. Fig. 4(a)], under the Kramers-Moyal and improved diffusion approximations, as a function of the powering voltage V_{dd} for the circuit, in units of $V_T \equiv k_B T/e$. Here, Ω corresponds to the smallest eigenvalue of the Maxwell capacitance matrix of the underlying circuit. (b) Fractional error in λ_1 , the first Liouvillian eigenvalue, for both approximations. At $V_{dd} = \log(2)V_T$ (vertical dashed line), the system undergoes a phase transition and becomes bistable.

for the Kramers-Moyal approximation, we obtain $\frac{dI_{KM}}{dx} = -4 \frac{\omega(x) - \omega(-x)}{\omega(x) + \omega(-x)}$. In particular, the rate function given by the Kramers-Moyal diffusion approximation is exact up to second order in the expansion parameter $1 - \omega(x)/\omega(-x)$ (as was seen in the example of the CMOS inverter). However, at high powering voltages, i.e., $V_{dd} \gtrsim 2V_T$, the discrepancy is enough to produce an observable difference in the value of the rate function at the center point $x = 0$ [cf. Fig. 4(c)].

The first Liouvillian eigenvalue for our circuit can be estimated from the stationary rate function I via the following expression [50]:

$$\lim_{\Omega \rightarrow \infty} \Omega^{-1} \log \lambda_1 = \Delta I, \quad (45)$$

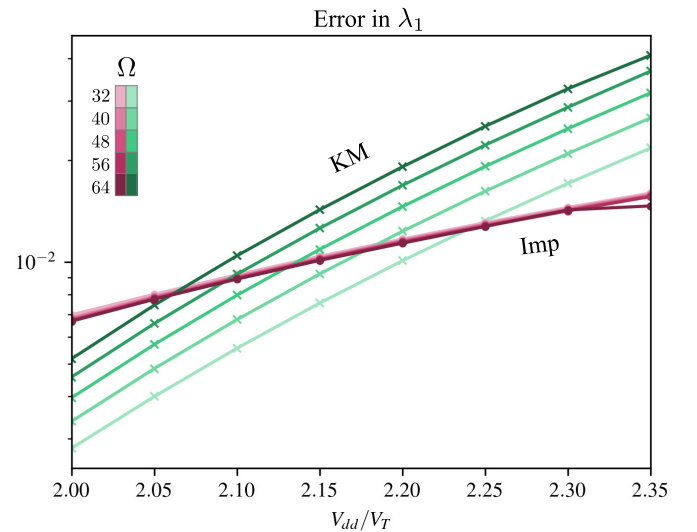


FIG. 6. Error in estimating the error rate of a CMOS memory cell. Same data as in Fig. 5(b), except with a slightly higher range for Ω and V_{dd} .

where ΔI is the difference of the stationary rate function between the center point $x = 0$ and any of its two minima. This result is the nonequilibrium generalization of the well-known Eyring-Kramers formula for metastable transitions in detailed-balanced systems [50]. Because $I_{\text{KM}} \neq I$ whereas $I_{\text{IMP}} = I$, we expect the improved diffusion approximation to provide a better estimate for λ_1 in the large system size limit, whenever ΔI_{KM} is appreciably different from ΔI . ΔI_{KM} is appreciably different from ΔI only at powering voltages $V_{\text{dd}} \gtrsim 2V_T$, and accordingly we see the improved approximation outperform the Kramers-Moyal approximation when the memory is operated at these higher powering voltages [cf. Fig. 5(b)].

C. Fixed points versus limit cycles

Although the improved diffusion approximation is theoretically expected to work well only in the linear-response regime close to detailed-balance conditions, the previous example shows that it can outperform the usual Kramers-Moyal approximation even beyond linear order in the thermodynamic forces that break detailed balance. This can be understood as follows. The deterministic dynamics of the CMOS memory has fixed-point attractors $\mathbf{x}^*$ at which the drift vanishes: $\mu_n(\mathbf{x}^*) = 0$. For cases in which the different jump vectors Δ_ρ are linearly independent, which is a condition satisfied in the examples considered, $\mu_n(\mathbf{x}^*) = 0$ implies that $\lambda_\rho(\mathbf{x}^*) - \lambda_{-\rho}(\mathbf{x}^*) = 0$ for all ρ . This in turn implies that the entropy production of the original MJP vanishes at the fixed points [see Eq. (33)]. Thus, near the fixed points, where most of the probability is concentrated in the macroscopic limit, the system operates close to detailed-balanced conditions, even if detailed balance is significantly broken at a global level. For this reason, we expect the better performance of the improved diffusion approximation to carry on to other systems satisfying the above conditions.

However, for systems that have more complex dynamical attractors at which entropy is produced, we do not expect the advantages of the improved diffusion approximation to extend beyond the linear-response regime. To test this case, we have performed simulations on the CMOS ring oscillator shown in Fig. 7(a). This circuit can be described via a MJP in the same way as the examples considered above. It features a Hopf bifurcation from a monostable phase into a limit cycle at $V_{\text{dd}} = -V_{\text{ss}} = \log(3)V_T$. As with the CMOS memory, we numerically compute the steady-state distributions according to both diffusive approximations and measure their KL divergence with respect to the exact steady state of the MJP. The results are shown in Fig. 7(b). We indeed see that, as expected from the theory, the improved diffusion approximation only performs better than the usual Kramers-Moyal approximation for moderate powering voltages, below the transition into a limit cycle.

VI. DISCUSSION

In this work, we have considered an alternative diffusive approximation of Markov jump processes that, under

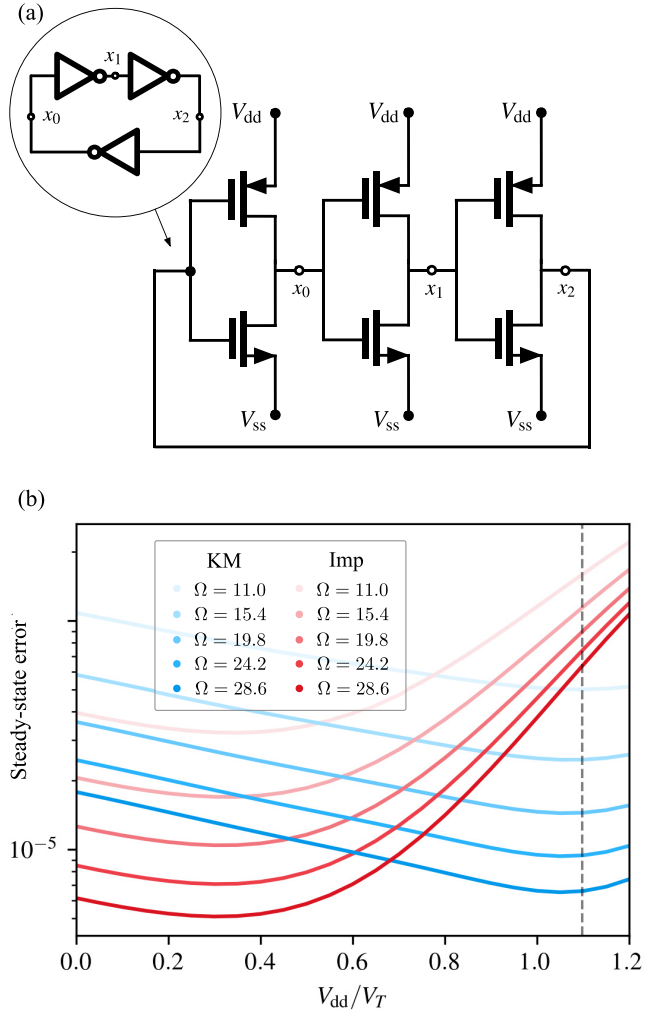


FIG. 7. Error in estimating the nonequilibrium steady state of a CMOS ring oscillator. (a) A ring oscillator consists of an odd number of inverters connected in feedback. Here, we show a CMOS implementation of a ring oscillator with three inverters. (b) Error (as measured by the Kullback-Leibler divergence) in estimating the nonequilibrium steady state of the ring oscillator as a function of the powering voltage of the oscillator (darker curves indicate increasing Ω). Here, $V_{\text{ss}} = -V_{\text{dd}}$. Above the threshold voltage $V_{\text{dd}}/V_T = \log 3$ (dashed vertical line), the system undergoes a phase transition and a limit cycle emerges in the deterministic dynamics for the oscillator.

strict detailed-balance conditions, was known to outperform the usual diffusive approximation based on the truncation to second order of the Kramers-Moyal expansion. We have extended the validity of this method by proving that its main advantage, namely, the accurate description of arbitrarily large macroscopic steady-state fluctuations, is preserved in the linear-response regime around strict detailed balance. By comparing the performance of both diffusive approximations on stochastic models of realistic electronic circuits, we have shown that the improved diffusion approximation is also more accurate for the description of dynamical features beyond steady-state fluctuations, in particular higher-order dynamical

modes and relaxation times. The theoretical arguments and numerical experiments supporting our results indicate that these advantages can be expected to hold significantly beyond the linear-response regime for systems with fixed-point attractors, while not for systems with more complex dynamical attractors like limit cycles.

DATA AVAILABILITY

The source code and precomputed simulation datasets required to reproduce the figures in this manuscript, as well as the data that support the findings of this article, are openly available on Zenodo at Ref. [51].

- [1] A. Arkin, J. Ross, and H. H. McAdams, Stochastic kinetic analysis of developmental pathway bifurcation in phage λ -infected *Escherichia coli* cells, *Genetics* **149**, 1633 (1998).
- [2] P. C. Bressloff, Stochastic switching in biology: From genotype to phenotype, *J. Phys. A: Math. Theor.* **50**, 133001 (2017).
- [3] H. Spohn, Long range correlations for stochastic lattice gases in a non-equilibrium steady state, *J. Phys. A: Math. Gen.* **16**, 4275 (1983).
- [4] J. L. Lebowitz and H. Spohn, A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics, *J. Stat. Phys.* **95**, 333 (1999).
- [5] C. Wasshuber, *Computational Single-Electronics* (Springer, Vienna, Wien, Austria, 2001).
- [6] S. V. Moreira, P. Samuelsson, and P. P. Potts, Stochastic thermodynamics of a quantum dot coupled to a finite-size reservoir, *Phys. Rev. Lett.* **131**, 220405 (2023).
- [7] C. Y. Gao and D. T. Limmer, Principles of low dissipation computing from a stochastic circuit model, *Phys. Rev. Res.* **3**, 033169 (2021).
- [8] N. Freitas, J.-C. Delvenne, and M. Esposito, Stochastic thermodynamics of nonlinear electronic circuits: A realistic framework for computing around kT , *Phys. Rev. X* **11**, 031064 (2021).
- [9] N. Freitas, G. Massarelli, J. Rothschild, D. Keane, E. Dawe, S. Hwang, A. Garlapati, and T. McCourt, Taming non-equilibrium thermal fluctuations in subthreshold CMOS circuits (unpublished).
- [10] D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *J. Comput. Phys.* **22**, 403 (1976).
- [11] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.* **81**, 2340 (1977).
- [12] D. T. Gillespie, The chemical Langevin equation, *J. Chem. Phys.* **113**, 297 (2000).
- [13] J. M. Horowitz, Diffusion approximations to the chemical master equation only have a consistent stochastic thermodynamics at chemical equilibrium, *J. Chem. Phys.* **143**, 044111 (2015).
- [14] A. J. McKane and T. J. Newman, Predator-prey cycles from resonant amplification of demographic stochasticity, *Phys. Rev. Lett.* **94**, 218102 (2005).
- [15] C. K. Fisher and P. Mehta, The transition between the niche and neutral regimes in ecology, *Proc. Natl. Acad. Sci. USA* **111**, 13111 (2014).
- [16] M. Vellela and H. Qian, Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: The Schlögl model revisited, *J. R. Soc. Interface* **6**, 925 (2009).
- [17] C. Gardiner, *Stochastic Methods* (Springer, Berlin, 2009).
- [18] M. I. Dykman, E. Mori, J. Ross, and P. M. Hunt, Large fluctuations and optimal paths in chemical kinetics, *J. Chem. Phys.* **100**, 5735 (1994).
- [19] L. E. Reichl, *A Modern Course in Statistical Physics*, 2nd ed. (John Wiley & Sons, New York, NY, 1998).
- [20] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007).
- [21] P. Hanggi, Nonlinear fluctuations: The problem of deterministic limit and reconstruction of stochastic dynamics, *Phys. Rev. A* **25**, 1130 (1982).
- [22] P. Hanggi, H. Grabert, P. Talkner, and H. Thomas, Bistable systems: Master equation versus Fokker-Planck modeling, *Phys. Rev. A* **29**, 371 (1984).
- [23] P. Hanggi and P. Jung, Bistability in active circuits: Application of a novel Fokker-Planck approach, *IBM J. Res. Dev.* **32**, 119 (1988).
- [24] T. Reichenbach, M. Mobilia, and E. Frey, Coexistence versus extinction in the stochastic cyclic Lotka-Volterra model, *Phys. Rev. E* **74**, 051907 (2006).
- [25] O. Ovaskainen and B. Meerson, Stochastic models of population extinction, *Trends Ecol. Evol.* **25**, 643 (2010).
- [26] S. Panahi, U. Feudel, K. C. Abbott, A. Hastings, and Y.-C. Lai, Generalized paradox of enrichment: Noise-driven rare rarity in degraded ecological systems, *J. R. Soc. Interface* **22**, 20250087 (2025).
- [27] E. Rezaei, M. Donato, W. R. Patterson, A. Zaslavsky, and R. I. Bahar, Fundamental thermal limits on data retention in low-voltage CMOS latches and SRAM, *IEEE Trans. Device Mater. Reliab.* **20**, 488 (2020).
- [28] N. Freitas, K. Proesmans, and M. Esposito, Reliability and entropy production in nonequilibrium electronic memories, *Phys. Rev. E* **105**, 034107 (2022).
- [29] A. Jelinčič, O. Lockwood, A. Garlapati, P. Schillinger, I. Chuang, G. Verdon, and T. McCourt, An efficient probabilistic hardware architecture for diffusion-like models, [arXiv:2510.23972](https://arxiv.org/abs/2510.23972).
- [30] T. Schmiedl and U. Seifert, Stochastic thermodynamics of chemical reaction networks, *J. Chem. Phys.* **126**, 044101 (2007).
- [31] A. Lazarescu, T. Cossetto, G. Falasco, and M. Esposito, Large deviations and dynamical phase transitions in stochastic chemical networks, *J. Chem. Phys.* **151**, 064117 (2019).
- [32] N. J. López-Alamilla, M. W. Jack, and K. J. Challis, Enhanced diffusion and the eigenvalue band structure of Brownian motion in tilted periodic potentials, *Phys. Rev. E* **102**, 042405 (2020).
- [33] K. J. Challis, Numerical study of the tight-binding approach to overdamped Brownian motion on a tilted periodic potential, *Phys. Rev. E* **94**, 062123 (2016).
- [34] R. Hayashi, K. Sasaki, S. Nakamura, S. Kudo, Y. Inoue, H. Noji, and K. Hayashi, Giant acceleration of diffusion observed in a single-molecule experiment on F₁-ATPase, *Phys. Rev. Lett.* **114**, 248101 (2015).
- [35] I. Santamaría-Holek and A. Pérez-Madrid, Eyring equation and fluctuation–dissipation far away from equilibrium, *J. Chem. Phys.* **153**, 244116 (2020).

- [36] It is useful to point out that, provided that the jump vectors $\{\Delta_\rho\}_\rho$ are linearly independent, near a deterministic fixed point $\mathbf{x} \approx \mathbf{x}_*$, the logarithmic mean appearing in Eq. (6) reduces to the arithmetic mean, so the above diffusion tensor and Eq. (5) asymptotically coincide near a deterministic fixed point.
- [37] N. Freitas, G. Falasco, and M. Esposito, Linear response in large deviations theory: A method to compute non-equilibrium distributions, *New J. Phys.* **23**, 093003 (2021).
- [38] J. N. Freitas and M. Esposito, Emergent second law for non-equilibrium steady states, *Nat. Commun.* **13**, 5084 (2022).
- [39] G. Falasco and M. Esposito, Macroscopic stochastic thermodynamics, *Rev. Mod. Phys.* **97**, 015002 (2025).
- [40] D. Santolin, N. Freitas, M. Esposito, and G. Falasco, Bridging Freidlin-Wentzell large deviations theory and stochastic thermodynamics, *Phys. Rev. E* **111**, 024106 (2025).
- [41] T. Van Vu and K. Saito, Thermodynamic unification of optimal transport: Thermodynamic uncertainty relation, minimum dissipation, and thermodynamic speed limits, *Phys. Rev. X* **13**, 011013 (2023).
- [42] T. Van Vu, V. T. Vo, and K. Saito, Dissipation, quantum coherence, and asymmetry of finite-time cross-correlations, *Phys. Rev. Res.* **6**, 013273 (2024).
- [43] To see this, note that, defining $\mathbf{x}_\rho \equiv \Delta_\rho/\Omega$, we can write
- $$\partial_t P = \sum_\rho \left(-\partial_n x_\rho^n (\lambda_\rho P) + \partial_n \partial_m x_\rho^m x_\rho^n \sum_{k=2}^{\infty} \frac{(-\partial_n x_\rho^n)^{k-2}}{k!} (\lambda_\rho P) \right).$$
- We then note that the above is in the form Eq. (8), with $\mu_n(\mathbf{x}) = \sum_\rho x_\rho^n \lambda_\rho(\mathbf{x})$ and $\tilde{D}(\mathbf{x}, t)$ given by Eq. (10).
- [44] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed., edited by H. Haken, Springer Series in Synergetics Vol. 18 (Springer-Verlag, Berlin, 1989).
- [45] R. Kubo, K. Matsuo, and K. Kitahara, Fluctuation and relaxation of macrovariables, *J. Stat. Phys.* **9**, 51 (1973).
- [46] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/dw4d-yj8c> for more details, which includes Refs. [16,37,44,52].
- [47] M. Mendler and B. Drossel, Predicting properties of the stationary probability currents for two-species reaction systems without solving the Fokker-Planck equation, *Phys. Rev. E* **102**, 022208 (2020).
- [48] A. Ceccato and D. Frezzato, Remarks on the chemical Fokker-Planck and Langevin equations: Nonphysical currents at equilibrium, *J. Chem. Phys.* **148**, 064114 (2018).
- [49] A. Gopal, M. Esposito, and N. Freitas, Large deviations theory for noisy nonlinear electronics: CMOS inverter as a case study, *Phys. Rev. B* **106**, 155303 (2022).
- [50] F. Bouchet and J. Reygner, Generalisation of the Eyring-Kramers transition rate formula to irreversible diffusion processes, *Ann. Henri Poincaré* **17**, 3499 (2016).
- [51] D. Roberts, T. McCourt, G. Massarelli, J. Rothschild, and N. Freitas, Replication package: Improved diffusive approximation of Markov jump processes, Zenodo (2026), doi:10.5281/zenodo.18204136.
- [52] J. Chang and G. Cooper, A practical difference scheme for Fokker-Planck equations, *J. Comput. Phys.* **6**, 1 (1970).