Level-2.5 large deviations and uncertainty relations for self-interacting jump processes: Tilting constructions and the emergence of timescale separation

Francesco Coghi* and Juan P. Garrahan

School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, United Kingdom
and Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, Nottingham NG7 2RD, United Kingdom

(Received 25 March 2026; accepted 21 October 2026; published 6 November 2026)

Self-interacting jump processes (SIJPs) describe systems with non-Markovian stochastic dynamics in which transition rates depend on empirical observables of the process, which gives rise to long-range memory and feedback. We derive the "level-2.5" large deviation (LD) principle governing the joint fluctuations of empirical occupation measure and the flux matrix for a broad class of SIJPs with general functional dependence on an empirical observable. The derivation is based on an exponential tilting construction and reveals a separation between a faster timescale of the microscopic dynamics and a slower timescale of the memory-driven evolution of transition rates, which is expressed through an exponentially discounted LD rate functional. Using this variational framework, we derive kinetic and thermodynamic uncertainty relations that extend classical Markovian bounds to non-Markovian systems and illustrate their performance with simple examples.

DOI: 10.1103/vhv8-tcrs

This is a companion paper to Ref. [1]. Here we present the full derivations of the results reported in an abridged manner in Ref. [1], together with additional general results and illustrative examples, in a self-contained form.

I. INTRODUCTION

Communicating information about past experiences is a distinctive feature of intelligent systems. While many organisms store and process information internally via a nervous system, numerous others are too small to possess such structures. Over evolutionary timescales, these organisms have instead developed forms of *external spatial memory*, whereby information is encoded in the surrounding environment rather than stored internally. Several microorganisms, including *Escherichia coli*, *Pseudomonas aeruginosa*, and *Pharaoh's ants*, achieve this by depositing chemical traces [2–7], while others, such as slime molds [8], encode information through persistent structural modifications of their environment. These environmental changes act as cues that can be sensed either by conspecifics or by the same organism at later times. In the latter case, the resulting dynamics are described as *autochemotactic* or *self-interacting*, whereby an organism's future behavior is influenced by its interaction with an environment it has itself modified. Remarkably, analogous behaviors have recently been realized in artificial systems, including autonomous agents such as self-propelled particles [9–12] and self-morphic active matter [13].

Approximate models of the dynamics of self-interacting autochemotactic agents exist. Reinforced random walks [14–21], both on edges and vertices, and field-theoretic descriptions [22–25], also in combination with Langevin-type dynamics [26–29], can replicate the long-time behavior of self-interacting agents, which reveal rich dynamical features. These models are extremely valuable in mimicking self-attracting and self-repelling behavior and in replicating the formation of specific trails and patterns. A rigorous description is offered by self-interacting processes (SIPs, see Ref. [17] for an introduction), a class of reinforced processes, such as chains [30–32], jump processes [33], and diffusion [34], whereby the self-interaction is mediated by a functional of the *empirical occupation measure*, a random field that models the chemical, or structural, changes generated by autochemotactic agents.

Methods such as stochastic approximation [34,35] and hydrodynamic limits are often effective in studying the asymptotic behavior of the models above, including stationary dynamics and typical fluctuations (see, e.g., Refs. [34,36–38]). By contrast, finite-time fluctuations and rare events have only recently begun to attract attention [1,32,33,38–42], despite their known importance in shaping future dynamics. This difficulty arises because SIPs, as well as other reinforced random walks and related field-theoretic models, are inherently non-Markovian when restricted to the configuration space accessible to the agent. This structure often obstructs straightforward analytical treatments and gives rise to complex phenomena, such as dynamical ergodicity breaking [34,38,43] and modified first-passage dynamics [42,44].

To address this gap, here we derive the large deviation asymptotics of the empirical occupation and flux, that is,

*Contact author: francesco.coghi@leicester.ac.uk

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

the dynamical large deviations (LDs) at "level 2.5" (see Refs. [45,46] for the Markovian case), for a broad class of self-interacting jump processes (SIJPs), in which the self-interaction depends functionally on both the empirical occupation measure and the empirical flux. (Our main results here are presented in shortened form, and without detailed proofs, in Ref. [1].)

Our approach is based on a formal tilting procedure and, under suitable assumptions, clarifies the intrinsic multiscale dynamics characteristic of these systems. We then study generalized scalar flow- and currentlike observables and derive variational bounds on their fluctuations, yielding SIJP generalizations of kinetic and thermodynamic uncertainty relations [47–53]. We thus extend the fluctuation-theoretic tools developed for Markov processes to non-Markovian dynamics. Finally, through representative examples, we show how self-interaction can enhance fluctuations relative to purely Markovian systems and assess how exact or numerical large deviation results compare with the SIJP fluctuation bounds.

The paper is structured as follows. In Sec. II we give a nonexhaustive overview of the statistical physics literature on fluctuations of non-Markov processes, in Sec. III we introduce the class of non-Markov SIJPs we focus on, and in Sec. IV we derive their level-2.5 LDs. By restricting to specific classes of SIJPs, in Sec. V we derive the probability bounds for flux observables: Specifically, in Sec. V A we derive a SIJP version of the kinetic uncertainty relation (KUR) [49,50], which we call SIJP-KUR, and in Sec. V B a SIJP version of the "ultimate KUR" (UKUR) [52,53], which we call SIJP-UKUR. We then look at the case of generalized current observables in Sec. VI and derive a SIJP version of the thermodynamic uncertainty relation (TUR) [47,48], which we call SIJP-TUR. In Sec. VII we illustrate our results with simple examples. Finally, in Sec. VIII we conclude the paper with a summary and an overview of interesting research directions.

II. RELATED WORK ON FLUCTUATIONS OF NON-MARKOV PROCESSES

Long-range memory effects, characteristic of non-Markovian models, are known to strongly influence fluctuation behavior, leading to departures from the large deviation theory established for Markovian systems [45,46,54–62]. In SIPs (and SIJPs), such memory effects originate from the explicit dependence of the dynamics on the empirical occupation measure (and empirical flux).

Non-Markovian random walks are known to exhibit memory-induced phenomena [63–66], and a "temporal additivity principle" has been proposed to partially account for fluctuations of time-additive observables [67]. These observables—including the empirical occupation measure—were argued to become slow variables with vanishing noise in the long-time limit, leading to reduced fluctuations and reinforced dynamics that progressively constrain exploration. Within this framework, mechanisms generating rare events in non-Markovian random walks, such as the elephant random walk and its Gaussian variant [66], have been identified.

Despite these advances, a comprehensive understanding of fluctuation dynamics in SIPs—where the evolution intrinsically depends on slowly varying variables—remained elusive until very recently. Large-deviation principles for self-interacting Markov chains [40,41] and jump processes [1,33] have now been established. In particular, Ref. [33] rigorously proves a level-2.5 large deviation principle for SIJPs under a self-interaction affine in the empirical occupation measure. In Ref. [1] we extend this result to general functionals of empirical occupation measure and flux, showing that fluctuations of a SIJP are *optimally* sampled by a nonhomogeneous Markov process. Together with related large deviation results for self-interacting Markov chains [40,41], generalized Pólya urn models [39,68], nonhomogeneous (periodically driven) Markov processes [69], and semi-Markov processes [52,70–77], these works provide the foundations for a coherent framework to study fluctuations in non-Markovian systems.

The loss of Markovianity introduces substantial technical challenges in the analysis of fluctuations. Alongside large deviation theory, fluctuation-dissipation theorems (FDTs) and uncertainty relations play a central role, as they provide general principles linking fundamental macroscopic observables to their typical fluctuations. As in the Markov case [47–52], these results can often be derived variationally from level-2.5 LDs. Despite sustained theoretical interest, progress in this direction remains limited for non-Markovian dynamics, with most existing results confined to generalized Langevin equations (GLEs) and semi-Markov processes.

Work on FDTs in non-Markov systems has mostly proceeded by embedding memory into GLEs or semi-Markov descriptions and then asking which parts of the Markovian FDT survive. For diffusions with memory kernels, it was shown that memory generally enhances the *frenetic*, or time-symmetric dynamical activity, of the linear response [78]. For general time dependence, the *second* FDT associated to GLEs can be derived via Mori-Zwanzig projections [79–81], with notable limitations when coarse-graining is in place [82–84], even in the case nonequilibrium baths are introduced [85]. Beyond diffusions, related fluctuation-response structures have been derived for semi-Markov networks [86].

For uncertainty relations, the classical TUR [47,48] for Markov jump and diffusion processes has been generalized along three main non-Markov directions: delay, semi-Markov kernels, and quantum or non-Markovian baths. For time-delayed Langevin systems an uncertainty relation was derived by bounding fluctuations in terms of the Kullback-Leibler divergence between a path measure and its time-reversed counterpart [87–89]. For antisymmetric observables this yields a bound that reduces to the standard TUR in the Markov limit. For semi-Markov jump processes, TURs that relate entropy production to the current asymptotic variance were derived using renewal theory [72] and an alternative derivation [74] that does not use the time-direction independence [70,90].

Notably, in deriving TURs, it is widely recognized that non-Markovian dynamics complicates the definition of entropy production, since irreversibility cannot in general be assigned from an instantaneous state alone. Nevertheless, several works have shown that local or instantaneous entropy-production rates can still be obtained once the memory structure is properly accounted for. This is the case, for example, for active Ornstein-Uhlenbeck particles [91] and GLE

dynamics with memory [92]. In both cases, irreversibility is quantified through a comparison of path weights, or equivalently through a dynamical cost associated with an auxiliary representation of the process.

For the case of KURs in the same semi-Markov and renewal settings, they have been obtained bounding current fluctuations in terms of effective activities depending on the full waiting-time statistics rather than just mean jump counts [49,50] or mean waiting times [52], providing natural non-Markov analogs of Markovian KURs [72,74]. Related kinetic-type bounds also appear in delayed and quantum transport models, where fluctuation-theorem or full-counting-statistics approaches yield constraints on current Fano factors with explicit dependence on delay kernels or bath correlation functions [87,88,93]. Additionally, stricter TURs that hold if the observed semi-Markov process arises from coarse-graining an underlying Markov chain have also been established, and their violation can be used as diagnostic of genuine memory [74].

Finally, in Ref. [1], we extended this body of work by deriving uncertainty relations for general SIJPs. In the following, we present the detailed derivations underlying the SIJP-KUR and SIJP-TUR reported in Ref. [1] and introduce the SIJP-UKUR together with its derivation.

III. SELF-INTERACTING JUMP PROCESSES

We define a SIJP as the continuous-time chain $(X_t)_{0\leqslant t\leqslant T}$ of discrete configurations, $x \in \mathcal{S} := \{1, 2, \dots, d\}$, with $\mathcal{E} := \{(x, y) \in \mathcal{S}\times\mathcal{S} : x \neq y\}$ denoting the set of all possible jumps between them. At time t, the SIJP takes a jump from the current state x to $y \neq x$ with evolving rate matrix (stochastic generator) $Q_{xy}(A_t) \in \mathbb{R}_+$, that depends on a general state-dependent *empirical observable* of the process

$$A_t = t^{-1}\int_0^t f_{X_{t'}}\, dt' + t^{-1}\sum_{\substack{0\leqslant t'\leqslant t\\ (X_{t'_-},X_{t'})\in\mathcal{E}}} g_{X_{t'_-}X_{t'}}, \tag{1}$$

where t' are jump times and $f : \mathcal{S} \to \mathbb{R}$ and $g : \mathcal{E} \to \mathbb{R}$ bounded functions. We assume the mapping $Q(a)$ to be at least of class C^1 on $\mathbb{R}$ as this guarantees enough regularity for the existence of a stationary state. Furthermore, for the block-averaging argument that comes below, we assume $Q_{xy}(a)$ to be uniformly positive on the admissible domain of a. This means

$$Q_{xy}(a) \geqslant q_- > 0, \tag{2}$$

for all $(x, y) \in \mathcal{E}$ and relevant a, with q_- some fixed constant independent of a. This means that irrespective of the argument of $Q(\cdot)$, the transition probabilities for the directed links in $\mathcal{E}$ are never vanishing and always at least q_-. If $g = 0$ in Eq. (1), then such a condition simply follows from continuity of the map $Q(a)$ as well as f being bounded; otherwise, it must be imposed.

By conservation of probability, the diagonal terms of the rate matrix are the negative escape rates

$$Q_{xx}(A_t) = -\sum_{y\in\mathcal{S},y\neq x} Q_{xy}(A_t). \tag{3}$$

In contrast with a homogeneous Markov process, the jump rates are not constant over time and they change as A_t evolves, making the process non-Markovian.

We define

$$L_x(t) = t^{-1}\int_0^t \mathbf{1}_x(X'_t)\, dt', \tag{4}$$

as the *empirical occupation measure* of the process, viz. the fraction of time that the process has spent at each and every configuration of the state space up to time t, with $\mathbf{1}_x(X_{t'})$ the indicator function of configuration x. Such a time-evolving object lives in a subset of the whole probability space on $\mathcal{S}$, i.e., $L(t) \in \mathcal{P}_0(\mathcal{S}) \subseteq \mathcal{P}(\mathcal{S})$. Along with it, we introduce the empirical flux

$$\Phi_{xy}(t) = t^{-1}\sum_{\substack{0\leqslant t'\leqslant t\\ (X_{t'_-},X_{t'})\in\mathcal{E}}} \mathbf{1}_x(X_{t'_-})\mathbf{1}_y(X_{t'}), \tag{5}$$

viz. the non-negative $d\times d$ matrix (with zero diagonal elements) of time-averaged number of jumps in the interval $[0, t]$ between any two different configurations in $\mathcal{S}$.

Given the two objects above, it is possible to cast A_t in Eq. (1) as a function of them as follows:

$$A_t = \sum_{x\in\mathcal{S}} f_x L_x(t) + \sum_{(x,y)\in\mathcal{E}} g_{xy}\Phi_{xy}(t). \tag{6}$$

Such a rewriting highlights the dependence on the specific stochastic features of the SIJP that we consider in this work. As usual, if one moves to an extended space where the state is $(X_t, L(t), \Phi(t))_{0\leqslant t\leqslant T}$, then a Markovian description of the dynamics can be constructed recursively from one jump to the next by a standard procedure.

In the main text of Ref. [1], the observable A_t is restricted to the case $g_{xy} = 0$ for all $(x, y) \in \mathcal{E}$ (and only at the end we briefly mention the extension to the case with a nonzero g dependence). Here we consider from the outset the most general class of SIJPs as the derivation of the level-2.5 LDs appearing in Sec. IV follows the same lines.

To proceed with the derivation, we need to make a few other assumptions. First, we assume that, for every value a in the range relevant to the observable, the rate matrix $Q(a)$ is irreducible—and therefore positively recurrent, since $\mathcal{S}$ is finite. For each frozen value of a, this guarantees the existence of a unique *accompanying distribution* $\gamma(a)$ satisfying

$$\gamma(a)Q(a) = 0, \quad \sum_{x\in\mathcal{S}} \gamma_x(a) = 1. \tag{7}$$

The C^1 regularity of the map $a \mapsto Q(a)$ then implies that $a \mapsto \gamma(a)$ is continuous. Although irreducibility and regularity together imply the existence of at least a limiting stationary state via Brouwer's fixed-point theorem, they do not imply that the self-interacting dynamics converges. We therefore need to further assume that both the empirical measure and flux converge in the long-time limit to finite typical values, i.e.,

$$L(t) \to \pi \quad \Phi(t) \to \varphi, \tag{8}$$

for some pair (π, φ), which may depend on the trajectory and is not necessarily unique. These assumptions can be checked

on a case-by-case basis and in Appendix A we discuss them more in detail and provide simple examples in which they fail.

We remark that for a Markov process with fixed generator Q, the stationary state π obeys $\pi Q = 0$, the empirical measure converges to π, and the empirical flux to $\pi \circ Q$ (where $\circ$ denotes the Hadamard product). In the SIJP case things are different: One needs the long-time limit of the observable, i.e.,

$$A_t \to \bar{a} = \sum_{x\in\mathcal{S}} f_x \pi_x + \sum_{(x,y)\in\mathcal{E}} g_{xy}\varphi_{xy}, \tag{9}$$

and the stationary density π solution of the nonlinear self-consistency equation

$$\pi Q(\bar{a}) = 0. \tag{10}$$

Also, $\varphi \neq \pi \circ Q(\bar{a})$ in general; it will be given an exact form in the next section. In particular, the limiting empirical flux has to satisfy the stationarity constraint

$$\sum_{y\neq x} \varphi_{xy} = \sum_{y\neq x} \varphi_{yx}, \quad x \in \mathcal{S}, \tag{11}$$

but it need not be fixed *a priori* by the typical Markov form $\pi_x Q_{xy}(\bar{a})$ when considering the more general stationary pairs appearing in the level-2.5 variational problem that follows.

In the following, we denote by $\mathcal{A}_{\text{stat}}$ the set of stationary limiting pairs (π, φ) satisfying (9), (10), and (11). We also highlight that, differently from the more usual time-homogeneous case, the selection of the specific stationary pair (π, φ) is not ruled by a breaking of the irreducibility of the process over the state space—we assume above that irreducibility always holds—but it is a dynamical feature emerging from the interplay of noise-induced fluctuations and long-range memory.

Notation

The notation we use is as follows. We use uppercase for random variables (or functions and sets) and lowercase for their realizations, while reserving P for probability. $\mathcal{P}$ is used to refer to a probability space and, generally, P for probability distributions, with $\mathbb{P}$ specially reserved for path probability measures. We use a compact notation for multidimensional objects: Given vector $\pi := (\pi_x)_{x\in\mathcal{S}}$ and matrix $Q := (Q_{xy})_{(x,y)\in\mathcal{S}\times\mathcal{S}}$, πQ denotes their internal product, $(\pi Q)_y = \sum_{x\in\mathcal{S}} \pi_x Q_{xy}$, and use $\circ$ for the Hadamard product, with $\pi \circ Q$ having elements $\pi_x Q_{xy}$. For scalar quantities we write the time dependence with a subscript, e.g., A_t, for compactness, whereas for higher-dimensional objects, say, $\mu := (\mu_x)_{x\in\mathcal{S}}$ as $\mu(t)$. $\mathbf{1}_x(\cdot)$ is the indicator function, which gives one if the argument coincides with x or zero otherwise. We use overbar to mark typical values, e.g., $\lim_{t\to\infty} A_t = \bar{a}$. The notation $a_T \sim b_T$ as $T \to \infty$ means asymptotic equivalence, namely $\frac{a_T}{b_T} \to 1$. In particular, $P_T(a) \sim 1$ means that $P_T(a) \to 1$ as $T \to \infty$. Similarly, the symbol $\asymp$ indicates equality up to subexponential factors in a large parameter (often time), so that $P_T(\cdot) \asymp e^{-TI(\cdot)}$ is equivalent to $I(\cdot) = -\lim_{T\to\infty} T^{-1} \log P_T(\cdot)$. We define the asymptotic variance (or susceptibility) of an (intensive) trajectory observable such as B_T as $\text{var}(b) := \lim_{T\to\infty} T\,\text{var}(B_T)$, where $\text{var}(B_T) = \mathbb{E}[(B_T - \bar{b})^2]$. Finally, we use the row convention: Probabilities are row vectors, stochastic generators have rows adding up to zero, and time-propagation is left to right, e.g., $\partial_t \mu(t) = \mu(t)Q$.

IV. LEVEL-2.5 LARGE DEVIATIONS: SLOW-FAST DYNAMICS

A. Probability path measures and tilting

Our focus is on the long, but finite, time dynamics of the empirical occupation measure in Eq. (4) and the empirical flux in Eq. (5). By going beyond their typical behavior, we formally derive the asymptotic exponential scaling of the joint probability $P(L(t) = \ell, \Phi_t = \phi) =: P_t(\ell, \phi)$, namely the level-2.5 large deviation rate function for SIJPs.

To do so, we use exponential tilting, a well-known technique in statistical physics. To our knowledge, this method has hitherto been applied to derive the level-2.5 LDs of Markov processes only, including jump processes [45,46,58,94,95] and diffusions [58,95,96]. [Note that for discrete-time Markov chains, to study the statistics of kth order empirical occupation measures, such as the $(k = 2)$ number of jumps between any two states, it is enough to *magnify* the state space, viz. replace original states with states obtained by k steps in the chain.] In what follows, by properly modifying the assumptions in the derivation, we will extend the use of tilting to non-Markov jump processes.

With the definitions above we can write the path probability measure: Given a path $\omega = \{(x_0, t_0'), (x_1, t_1'), \ldots, (x_n, T)\}$ that is a specific realization of the SIJP $(X_t)_{0\leqslant t\leqslant T}$, where t_{i-1}' is the ith jump time, the corresponding path measure reads

$$\mathbb{P}_T(\omega) = \mu_0(x_0) \exp\left(\int_0^{t_0'} Q_{x_0x_0}(A_s)\,ds\right) Q_{x_0x_1}(A_{t_0'}) dt_0' \times \cdots \times \exp\left(\int_{t_{n-1}'}^{T} Q_{x_nx_n}(A_s) ds\right), \tag{12}$$

with μ_0 the probability distribution of the initial state. Notice that the SIJP spends an exponentially distributed time at each state in $\mathcal{S}$. However, differently from the time-homogeneous case, the overall jump rate out of a state is random for a SIJP as rates change with time. As detailed in Ref. [97], the overall jump rate out of x is

$$\Lambda_x(t, t+u) := -\int_t^{t+u} Q_{xx}(A_s)\,ds, \tag{13}$$

where t is the time at which the SIJP jumps onto x and u is the random time at which it jumps out of it. (In Appendix C, we describe two ways to simulate a SIJP.)

Thanks to the stationarity condition (10), the law of large numbers implies that, in the long-time limit, the empirical measure and flux converge with probability tending to one to the set of stationary limiting pairs. Accordingly, rather than concentrating on a single pair (π, φ), one has

$$P((L(T), \Phi(T)) \in \mathcal{A}_{\text{stat}}) \overset{T\to\infty}{\sim} 1, \tag{14}$$

where the probability density above is a marginal of the path measure $\mathbb{P}_T$ in Eq. (12) and the dependence on the trajectory is implicit in the definitions of $L(t)$ and $\Phi(t)$ in Eqs. (4) and (5), respectively.

Along with the SIJP $(X_t)_{0\leqslant t\leqslant T}$ we assume that a time-inhomogeneous Markov process $(\tilde{X}_t)_{0\leqslant t\leqslant T}$ with path probability measure $\tilde{\mathbb{P}}_T$ and stationary state $\ell \neq \pi$ exists, such that

$$\tilde{P}_T(\ell,\phi) = \int \mathbf{1}_\ell(L(T))\mathbf{1}_\phi(\Phi(T))\, d\tilde{\mathbb{P}}_T(\omega) \overset{T\to\infty}{\sim} 1, \quad (15)$$

for $\phi \neq \varphi$, in probability, and 0 otherwise. We assume the dynamics of such a process to be generated by a time nonhomogeneous rate matrix $\tilde{H}(t) := (\tilde{H}_{xy}(t))_{(x,y)} \in \mathcal{S}\times\mathcal{S}$ with diagonal elements the negative escape rates

$$\tilde{H}_{xx}(t) = -\sum_{y\in\mathcal{S}:y\neq x} \tilde{H}_{xy}(t), \quad (16)$$

as well as uniformly positive jump rates, viz.

$$\tilde{H}_{xy}(t) \geqslant h_- > 0, \quad (17)$$

for all $(x,y)\in\mathcal{E}$ and $t>0$, with h_- some fixed constant independent of t. Additionally, we assume that this time-inhomogeneous Markov process evolves on timescale comparable or slower than the timescale of the original SIJP (this requirement comes from the block-averaging derivation below and we postpone the detailed explained to then) and that an accompanying distribution $\tilde{\rho}(t)$ exists, is unique, and satisfies

$$\tilde{\rho}(t)\tilde{H}(t) = 0, \quad (18)$$

making the Markov process ergodic for all times $t\in\mathbb{R}_+$. For the trajectory ω and the initial condition μ_0, we write the path probability measure of $(\tilde{X}_t)_{0\leqslant t\leqslant T}$ as

$$\tilde{\mathbb{P}}_T(\omega) = \mu_0(x_0)\exp\left(\int_0^{t_0'} \tilde{H}_{x_0x_0}(t)\,dt\right)\tilde{H}_{x_0x_1}(t_0')dt_0' \times\cdots\times \exp\left(\int_{t_{n-1}'}^{T} \tilde{H}_{x_nx_n}(t)\,dt\right), \quad (19)$$

and assume it to be absolutely continuous with respect to the path measure of the original SIJP in Eq. (12) for every $T\in\mathbb{R}_+$. This assumption is supported by the fact that $\mathbb{P}_T$ and $\tilde{\mathbb{P}}_T$ have the same support as no breaking of irreducibility is allowed for the SIJPs here considered as mentioned at the end of Sec. III.

Having introduced the fundamental ingredients, we can now formally calculate the probability of hitting a rare event in the SIJP by an exponential change of measure. We write

$$P_T(\ell,\phi) = \mathbb{E}_{\tilde{\mathbb{P}}_T}\left[\mathbf{1}_\ell(L(T))\mathbf{1}_\phi(\Phi(T))\frac{d\mathbb{P}_T}{d\tilde{\mathbb{P}}_T}(\omega)\right] \asymp \exp[-TI_{2.5}(\ell,\phi)], \quad (20)$$

where the first equality is simply the definition of probability multiplied and divided by the path measure (19). By construction, the nonhomogeneous Markov process typically returns the pair (ℓ,ϕ), and the Radon-Nikodym derivative between path measures exists by absolute continuity and takes the asymptotic form in the second line. This latter expression we use to define the level-2.5 rate function $I_{2.5}(\ell,\phi)$. Such an exponential change of measure, or tilting, is widely adopted in importance sampling techniques whereby to sample an atypical event of a Markov process one introduces a new biased distribution which typically returns the event of interest [98–104] and corrects the quantity averaged with an appropriate "umbrella".

B. Timescale separation, time rescaling, and reversal

We now consider the calculation of the exponential factor in Eq. (20). Using the definitions in Eqs. (12) and (19) yields

$$\log\frac{d\mathbb{P}_T}{d\tilde{\mathbb{P}}_T}(\omega) = \sum_{i=0}^{n}\left(\int_{t_{i-1}'}^{t_i'} \left(Q_{x_ix_i}(A_t) - \tilde{H}_{x_ix_i}(t)\right)dt\right) + \sum_{i=1}^{n}\log\frac{Q_{x_{i-1}x_i}\left(A_{t_{i-1}'}\right)}{\tilde{H}_{x_{i-1}x_i}(t_{i-1}')}, \quad (21)$$

where $t'_{-1} := 0$ and $t'_n := T$. By introducing the identities $\sum_{x\in\mathcal{S}}\mathbf{1}_x(\tilde{X}_t)$ and $\sum_{(x,y)\in\mathcal{E}}\mathbf{1}_x(\tilde{X}_{t_-})\mathbf{1}_y(\tilde{X}_t)$, and using conservation of probability in the forms of Eqs. (3) and (16), Eq. (21) simplifies to

$$\log\frac{d\mathbb{P}_T}{d\tilde{\mathbb{P}}_T}(\omega) = -\sum_{(x,y)\in\mathcal{E}}\left[\int_0^T \mathbf{1}_x(\tilde{X}_t)\left(Q_{xy}(A_t) - \tilde{H}_{xy}(t)\right)dt + \sum_{\substack{0\leqslant t\leqslant T\\ \Delta\tilde{X}_t\neq 0}} \mathbf{1}_x(\tilde{X}_{t_-})\mathbf{1}_y(\tilde{X}_t)\log\frac{\tilde{H}_{xy}(t)}{Q_{xy}(A_t)}\right], \quad (22)$$

where we also used the fact that for the specific trajectory ω if the process is at x at time t'_{i-1}, then it will stay there at all times $t\in(t'_{i-1},t'_i]$. We remark at this stage that due to the time dependence of the rate matrices it is not possible to write Eq. (22) explicitly as a function of $L(T)$ and $\Phi(T)$, which is instead possible for time-homogeneous Markov processes. However, it is still possible to simplify Eq. (22) arguing, as we do in the following, a separation of timescales.

On the one hand, the SIJP's inherent timescale is set by the lowest jump rate in $Q(A_t)$. On the other hand, the rate matrix itself changes over time due to the dependence on the empirical observable in Eq. (1) [or Eq. (6)]. At short times, the jump rates may change rapidly due to the t^{-1} appearing in A_t. However, as time progresses, A_t, $L(t)$, and $\Phi(t)$ tend to stiffen: This progressive slowness is also due to the factor t^{-1}, which implies that to accumulate a change of a $O(1)$ a time of $O(t)$ is needed. Therefore, in the long-time limit, we identify a fast mode, inherent of the microscopic dynamics of the SIJP, and a slower mode consequence of the dependence of the rate matrix over an empirical occupation measure and flux that change slowly.

In the following, we make this multiscale dynamics argument more rigorous, introducing sufficient assumptions that support the derivation. We set $T\gg 1$ and split the integral over time into K integrals spanning the time intervals $I_k = [t_k, t_{k+1})$ with $\Delta t = t_{k+1} - t_k$. The following sequence of inequalities holds:

$$\tau_\mu := \max\left\{Q_{xy}^{-1}(A_t)\right\}_{(x,y)\in\mathcal{E}} \approx \max\left\{\tilde{H}_{xy}^{-1}(t)\right\}_{(x,y)\in\mathcal{E}} \ll \Delta t \ll t =: \tau_{\text{slow}}, \quad (23)$$

where τ_μ identifies the timescale associated to the microscopic SIJP as well as to the auxiliary time-inhomogeneous Markov process. The timescale separation above justifies that

within a time block of size Δt, $L(t)$, $\Phi(t)$, and A_t can be treated as constant, while the microscopic processes—which are themselves generated by frozen generators—mix on a characteristic timescale τ_μ which is much shorter than Δt. The sequence of inequalities above gives the conditions under which we can carry out the block-averaging calculation below. The left inequality requires the frozen auxiliary chain to equilibrate within a block, while the right inequality requires both the original SIJP generator and the auxiliary generator to remain approximately constant over the same block. For the SIJP, the variation time of $Q(A_t)$ is of order t as argued above. For the auxiliary generator, it is thus sufficient that its variation time is of order t or larger. Auxiliary generators varying on a faster timescale are not ruled out as stochastic processes, but they are outside the frozen-block derivation used here.

Block-averaging is consequence of mixing over the time window Δt. We write

$$\int_{t_k}^{t_{k+1}} \mathbf{1}_x(\tilde{X}_t)\, dt \approx \Delta t\ \tilde{\rho}_x(t_k), \tag{24}$$

$$\sum_{\substack{t_k \leqslant t \leqslant t_{k+1} \\ (\tilde{X}_{t_-}, \tilde{X}_t) \in \mathcal{E} \neq 0}} \mathbf{1}_x(\tilde{X}_{t_-}) \mathbf{1}_y(\tilde{X}_t) \approx \Delta t\ \tilde{\rho}_x(t_k) \tilde{H}_{xy}(t_k), \tag{25}$$

while the generators $Q(A_{t_k})$ and $\tilde{H}(t_k)$ are frozen and, therefore, approximately constant. We also stress that, as outlined in Eq. (23), the block-averaging argument requires the frozen SIJP generator and the frozen auxiliary generator to relax to their accompanying distributions within the time block Δt. The block size Δt can be chosen *a posteriori*. Since the state space $\mathcal{S}$ is finite and strongly connected, and since the rates of both the SIJP and the auxiliary process are assumed to be uniformly positive on the allowed transitions, the corresponding frozen generators have spectral gaps uniformly bounded away from zero. In particular, one may assume

$$\inf_{t>0} \operatorname{gap} \tilde{H}(t) > 0, \tag{26}$$

and analogously for $Q(A_t)$. This gives a uniform upper bound on the mixing times of the frozen dynamics. One can then choose Δt larger than this bound, while still satisfying the remaining scale-separation inequalities in Eq. (23).

We can therefore write

$$\log \frac{d\mathbb{P}_T}{d\tilde{\mathbb{P}}_T}(\omega) = -\sum_{(x,y)\in\mathcal{E}} \Bigg[\sum_{k=0}^{K-1} \tilde{\rho}_x(t_k) \big(Q_{xy}(A_{t_k}) - \tilde{H}_{xy}(t_k) \big)\, \Delta t + \sum_{k=0}^{K-1} \tilde{\rho}_x(t_k) \tilde{H}_{xy}(t_k) \log \frac{\tilde{H}_{xy}(t_k)}{Q_{xy}(A_{t_k})}\, \Delta t \Bigg] + \mathrm{Err}_T(\Delta t), \tag{27}$$

where the cumulative error $\mathrm{Err}_T(\Delta t)$ made in the timescale separation approximation comes from two approximations: freezing the SIJP and auxiliary generators on each block and replacing block averages by the corresponding frozen accompanying distribution. We can reformulate the timescale separation in Eq. (23) as requesting that the frozen cost density varies on the slow timescale, namely

$$\left| \frac{d}{dt} \sum_{(x,y)\in\mathcal{E}} \Bigg[\tilde{\rho}_x(t) \big(Q_{xy}(A_t) - \tilde{H}_{xy}(t) \big) + \tilde{\rho}_x(t) \tilde{H}_{xy}(t) \log \frac{\tilde{H}_{xy}(t)}{Q_{xy}(A_t)} \Bigg] \right| = O(t^{-1}). \tag{28}$$

Hence, the cumulative error satisfies

$$\frac{|\mathrm{Err}_T(\Delta t)|}{T} = O\left(\frac{\Delta t \log T}{T} + \frac{\tau_\mu}{\Delta t} \right). \tag{29}$$

The first contribution is the error related to the Riemann sum. Since the frozen cost density changes by order t^{-1} in Eq. (28), the error made on a block starting at t_k is proportional to the block length squared divided by t_k; summing these contributions over the blocks produces the logarithmic factor. The second contribution is due to the replacement of empirical occupations and fluxes by the frozen accompanying distribution. Making the replacement loses only a microscopic transient of order τ_μ; summing over the blocks and dividing by the total time gives the second term. Therefore, choosing for instance $\Delta t = T^\alpha$ with $0 < \alpha < 1$, the accumulated error is subextensive, $\mathrm{Err}_T(\Delta t) = o(T)$, and does not contribute to the large-deviation rate function. For more details on the error we refer the reader to Appendix B.

Finally, by taking the limit $K \to \infty$, the sum can be approximated with the integral over the mesoscopic timescale, yielding

$$\log \frac{d\mathbb{P}_T}{d\tilde{\mathbb{P}}_T}(\omega) = -\sum_{(x,y)\in\mathcal{E}} \int_0^T \left(\tilde{\rho}_x(t) \big(Q_{xy}(A_t) - \tilde{H}_{xy}(t) \big) + \tilde{\rho}_x(t) \tilde{H}_{xy}(t) \log \frac{\tilde{H}_{xy}(t)}{Q_{xy}(A_t)} \right) dt. \tag{30}$$

For compactness, we define

$$\mathcal{I}_T[\tilde{\rho}, \tilde{H}] := \frac{1}{T} \sum_{(x,y)\in\mathcal{E}} \int_0^T \tilde{\rho}_x(t) Q_{xy}(A_t) \Gamma\left(\frac{\tilde{H}_{xy}(t)}{Q_{xy}(A_t)} \right) dt, \tag{31}$$

where $\Gamma(x) = x \log x - x + 1$ is the Cramér (or relative-entropy) rate function for Poisson statistics.

Next, we make two changes of variable inside the time integral of Eq. (31). First, we introduce the adimensional logarithmic time $s = \log(t/t_0)$, where t_0 is interpreted as the system's relevant memory horizon and is then set to 1 without loss of generality. As $t \in [0, T]$, $s \in (-\infty, \log T]$. Within this timescale, the empirical observable A_t makes $O(1)$ changes in $O(1)$ time. We take the liberty to set $t_0 = 1$ because the calculation depends only on time derivatives and long-time asymptotics, while the additive constant generated by the logarithmic rescaling is irrelevant for our purposes. We stress, however, that t_0 may play an important role in concrete experimental settings where one wishes to test the theory developed here. In such cases, identifying the intrinsic timescales of the system under investigation would be necessary.

The first change of variable eventually yields

$$\mathcal{I}_T[\tilde{\rho},\tilde{H}] := \frac{1}{T}\sum_{(x,y)\in\mathcal{E}}\int_{-\infty}^{\log T} e^s \tilde{\rho}_x(e^s) Q_{xy}(A_{e^s})\Gamma \times \left(\frac{\tilde{H}_{xy}(e^s)}{Q_{xy}(A_{e^s})}\right) ds, \tag{32}$$

which effectively sets the system on the (slower) logarithmic timescale.

Then, we make the change $\tau = \log T - s$ to measure backwards from the final logarithmic time $\log T$. Due to this change, the integral boundaries flip to $[0,\infty)$. Eventually, this yields

$$\mathcal{I}[\rho,H] := \sum_{(x,y)\in\mathcal{E}}\int_0^\infty e^{-\tau}\rho_x(\tau)Q_{xy}(M_\tau)\Gamma\left(\frac{H_{xy}(\tau)}{Q_{xy}(M_\tau)}\right)d\tau, \tag{33}$$

where we have (i) dropped T in the notation $\tilde{\rho}(Te^{-\tau}) =: \rho(\tau)$, $A_{Te^{-\tau}} =: M_\tau$, and $\tilde{H}_{xy}(Te^{-\tau}) =: H_{xy}(\tau)$ and (ii) also dropped the explicit dependence on T from the definition of the functional $\mathcal{I}_T =: \mathcal{I}$. Of course, all objects still implicitly depend on T, but in the long-time limit they become T independent. Remarkably, in Eq. (33) the exponential factor $e^{-\tau}$ acts as a *discount*: As τ increases, the contribution of the corresponding time to the rate functional is exponentially suppressed. This aligns with the intuition that, at short physical times—corresponding to large τ under time reversal—the empirical measure fluctuates more rapidly and thus contributes less to the overall fluctuation cost than it does in the long-time limit.

By multiplying and dividing the argument of Γ by $\rho(\tau)$, the rate functional $\mathcal{I}$ takes the equivalent form

$$\tilde{\mathcal{I}}[\rho,\eta] := \sum_{(x,y)\in\mathcal{E}}\int_0^\infty e^{-\tau}\rho_x(\tau)Q_{xy}(M_\tau)\Gamma\left(\frac{\eta_{xy}(\tau)}{\rho_x(\tau)Q_{xy}(M_\tau)}\right)d\tau, \tag{34}$$

where we denote

$$\eta_{xy}(\tau) := \rho_x(\tau)H_{xy}(\tau), \tag{35}$$

the instantaneous flux of the rescaled and reverse-in-time inhomogeneous Markov process. This specific form will be important later when discussing kinetic bounds.

C. Average over paths: Laplace approximation

The rate functional obtained in Eq. (33) [or Eq. (34)] is still related to a specific, albeit long, realization of a trajectory of the nonhomogeneous Markov process $(\tilde{X}_t)_{0\leqslant t\leqslant T}$. The last step is to average over the ensemble of trajectories generated, as in Eq. (20). Given the exponential scaling with time of the Radon-Nikodym derivative, the average is calculated using a Laplace approximation—or contraction principle in the large deviation jargon—which finally yields

$$I_{2.5}(\ell,\phi) = \inf_{(\rho,H)\in\mathcal{S}_{\ell,\phi}} \mathcal{I}[\rho,H], \tag{36}$$

the level-2.5 rate function we are after. The infimum is over trajectories $(\rho(\tau),H(\tau))_{0\leqslant\tau<\infty}$ belonging to the set $\mathcal{S}_{\ell,\phi}$, which contains all normalized densities $\rho(\tau)$ along with the following constraints:

$$\rho(\tau)H(\tau) = 0, \tag{37}$$

$$M_\tau = e^\tau\int_\tau^\infty e^{-\tau'}\left(\sum_{x\in\mathcal{S}} f_x\rho_x(\tau') + \sum_{(x,y)\in\mathcal{E}} g_{xy}\eta_{xy}(\tau')\right)d\tau', \tag{38}$$

$$\ell_x = \int_0^\infty e^{-\tau'}\rho_x(\tau')\,d\tau', \tag{39}$$

$$\phi_{xy} = \int_0^\infty e^{-\tau'}\eta_{xy}(\tau')\,d\tau', \tag{40}$$

set to hold for all $x\in\mathcal{S}$ and all $(x,y)\in\mathcal{E}$. Equation (37) is equivalent to Eq. (18) rescaled and time reversed. The same holds for M_τ in Eq. (38) with respect to Eq. (1) and for the flux ϕ in Eq. (40) with Eq. (5). Finally, a contraction equivalent to (36) can be written for the functional $\tilde{\mathcal{I}}[\rho,\eta]$, over the same set of constraints.

The minimum and zero, which is not necessarily unique, of the rate function (36) characterizes the typical behavior of the SIJP. As $\Gamma(1) = 0$, this is obtained by setting $H_{xy}(\tau) = Q_{xy}(M_\tau)$. Such an equivalence is made possible by the fact that the mapping $Q_{xy}(M_\tau)$ is local in time in the extended state space $\mathcal{S}\times\mathcal{P}_0(\mathcal{S})\times\mathbb{R}_+^{d(d-1)}$. Therefore, it is always possible to find a deterministic operator $H_{xy}(\tau) = (Q_{xy}(M))(\tau)$. In such a typical case, the constraints (37)–(40) simplify to

$$\bar{\rho}(\tau)Q(\bar{M}_\tau) = 0, \tag{41}$$

$$\bar{M}_\tau = e^\tau\int_\tau^\infty e^{-\tau'}\left(\sum_{x\in\mathcal{S}} f_x\bar{\rho}_x(\tau') + \sum_{(x,y)\in\mathcal{E}} g_{xy}\bar{\eta}_{xy}(\tau')\right)d\tau', \tag{42}$$

$$\bar{\pi}_x = \int_0^\infty e^{-\tau'}\bar{\rho}_x(\tau')d\tau', \tag{43}$$

$$\bar{\varphi}_{xy} = \int_0^\infty e^{-\tau'}\bar{\eta}_{xy}(\tau')\,d\tau', \tag{44}$$

where $\bar{\rho}(\tau)$, $\bar{M}(\tau)$, and $\bar{\eta}(\tau)$ are the (rescaled and reverse-in-time) typical trajectories of instantaneous density, empirical observable and instantaneous flux of the original SIJP. Moreover, $(\bar{\pi},\bar{\varphi})$ are interpreted as weighted averages over the set of stationary limiting pairs $\mathcal{A}_{\text{stat}}$. In the case this set is singleton, viz. $\mathcal{A}_{\text{stat}} = \{(\pi,\varphi)\}$, then $\bar{\pi}\equiv\pi$ and $\bar{\varphi}\equiv\varphi$. We denote the set of trajectories $(\bar{\rho}(\tau), Q(\bar{M}_\tau))$ satisfying Eqs. (41)–(44) as $\mathcal{S}_{\mathcal{A}_{\text{stat}}}$.

We conclude this section with a few remarks. First, we can read Eqs. (36) and (33), along with the constraints (37)–(40), as a dynamical reweighting of the relative entropy of a Poisson process—the latter given by $\Gamma(x)$—and therefore as a generalization of the level-2.5 LDs for Markov jump processes [45,46,58,94,95], which follows the seminal work of Donsker and Varadhan [105]. Differently from the latter however, the level-2.5 LDs for SIJPs is not, generally speaking, convex. This is a consequence of the contraction appearing in Eq. (36), which may give rise to interesting dynamical behavior in generating fluctuations in SIJPs. One example in Sec. VII shows interesting features. However, more work is needed to better

understand the dynamical implications of having a nonconvex rate function for SIJPs.

Although we believe we have stated all necessary assumptions to derive the large deviation principle in Eq. (20) along with the rate function (36), one should check the validity of these assumptions on a case-by-case basis—an issue commonly encountered when working formally with tilting methods. For the broad class of affine rate matrices $Q(L(t))$, however, our findings are supported by a rigorous derivation recently presented in Ref. [33], where variational and control-theoretic formulations of Laplace functionals of Poisson random measures are used to establish level-2.5 LDs for self-interacting jump processes; see Eqs. (6.1)–(6.2) in Ref. [33]. While the approach adopted here is less rigorous, it extends beyond affine maps. Generalizing the rigorous derivations in Ref. [33] to the broader class of rate maps considered in this work and in our Ref. [1] also remains an important direction for future research.

Finally, we stress that the general level-2.5 LDs we derived is useful in practice for two reasons. The first is that long-time LDs of any time-extensive trajectory observable, such as activity or a current, as well as the empirical measure itself, can be obtained by contraction from the LDs of empirical occupation measure and flux. The second is that Eq. (36) is variational, so that even if the exact minimization is difficult, it is still possible to get suboptimal solutions as upper bounds to the true rate function, as we will show in the next two sections.

V. KINETIC BOUNDS

We focus on the generalized sample-mean flux observable

$$B_T = \sum_{(x,y)\in\mathcal{E}} \beta_{xy}\Phi_{xy}(T), \tag{45}$$

and $\beta_{xy} \in \mathbb{R}$ and $\Phi_{xy}(T)$ is defined in Eq. (5). In the following, we assume a unique stationary pair such that $\mathcal{A}_{\rm stat} = \{(\pi, \varphi)\}$. Typically, $B_T \to \bar{b}$ in probability, where

$$\bar{b} := \sum_{(x,y)\in\mathcal{E}} \beta_{xy}\varphi_{xy}, \tag{46}$$

where $\varphi \in \mathbb{R}_+^{d\times(d-1)}$ is the typical behavior of the empirical flux and is defined in Eq. (44).

Here we are interested in the fluctuations of B_T in Eq. (45). Its large deviation rate function $I(b)$ can be obtained by solving the variational problem in Eq. (36) along with the usual constraints (37)–(40) and the extra constraint

$$b = \sum_{(x,y)\in\mathcal{E}} \beta_{xy}\phi_{xy}, \tag{47}$$

which we express as follows for ease of readability:

$$I(b) = \inf_{\substack{(\rho,H)\in\mathcal{S}_{\ell,\phi}\\ \ell,b:\,\text{Eq.}(47)}} \mathcal{I}[\rho, H]. \tag{48}$$

The minimum and zero of such a rate function is found at $b = \bar{b}$ by setting $\ell = \pi$, $\phi = \varphi$, and $(\rho(\tau), H(\tau)) = (\bar{\rho}(\tau), \bar{H}(\tau)) \in \mathcal{S}_{\pi,\varphi}$ as solutions of Eqs. (41)–(44).

A. SIJP kinetic uncertainty relation

In this subsection, and similarly when deriving the SIJP-TUR below, we restrict to the case $g = 0$, namely the empirical observable dependence in the SIJP is only given by state contributions, i.e.,

$$A_t = t^{-1}\int_0^t f_{X_{t'}}\, dt'. \tag{49}$$

Such a restriction allows us to derive explicit formulas and link them to well-known results for time-homogeneous Markov processes.

Although an exact minimization of Eq. (48) is out of reach, we can find a suboptimal solution $I^*(b)$ with a simple Ansatz where the density coincides with the stationary state, but the flux is a scalar rescaling of the typical one, that is,

$$\ell^*_x = \pi_x, \tag{50}$$

$$\phi^*_{xy} = (b/\bar{b})\,\varphi_{xy}. \tag{51}$$

This in turn implies

$$(\rho^*(\tau), H^*(\tau)) = (\bar{\rho}(\tau), (b/\bar{b})Q(\bar{M}_\tau)) \in \mathcal{S}_{\pi,(b/\bar{b})\varphi}, \tag{52}$$

such that the constraints (41)–(44) are satisfied. Notice that if we had allowed for jump type contributions in the SIJP dependence on the empirical observable, then we would have not been able to write the equation above due to $M^*_\tau(\neq \bar{M}_\tau)$ including a dependence on the fluctuation b via $\eta^*_{xy}(\tau) = (b/\bar{b})\bar{\eta}_{xy}(\tau)$.

The meaning of the Ansatz above is clear: The time-dependent generator of the time-inhomogeneous auxiliary process is simply the generator of the SIJP rescaled by an overall factor determined by the fluctuation of interest. Eventually, Eq. (48) reduces to

$$I(b) \leqslant I^*(b) := k_{\rm SIJP}\Gamma(b/\bar{b}), \tag{53}$$

where

$$k_{\rm SIJP} := \sum_{(x,y)\in\mathcal{E}} \varphi_{xy} = \sum_{(x,y)\in\mathcal{E}} \int_0^\infty e^{-\tau}\bar{\eta}_{xy}(\tau)\, d\tau, \tag{54}$$

is the hereby-defined average *SIJP dynamical activity*. It corresponds to the number of configuration changes per unit time and it generalizes the standard dynamical activity of a time-homogeneous Markov jump process to the non-Markov case of SIJPs by adding an exponential discount in time.

We now expand $I(b)$ and $\Gamma(b/\bar{b})$ up to second order on, respectively, the left-hand side and right-hand side of Eq. (53). Then, by recalling the definition of asymptotic variance

$$(I''(b))^{-1} = \mathrm{var}(b), \tag{55}$$

we obtain

$$\frac{\mathrm{var}(b)}{\bar{b}^2} \geqslant \frac{1}{k_{\rm SIJP}}, \tag{56}$$

which is what we refer to as SIJP-KUR, generalizing the well-known KUR for time-homogeneous Markov processes [49,50]. Evidently, as for Markov processes, the precision

of flux (or current) observables is bounded by the average dynamical activity of SIJPs.

We remark that the identification of the inverse curvature with the asymptotic variance in Eq. (55) is understood in the fluctuation regime in which the observable satisfies a central limit theorem,

$$\sqrt{T}(B_T - \bar{b}) \to \mathcal{N}(0, \mathrm{var}(b)), \tag{57}$$

where convergence is understood in distribution. For SIJPs, as well as for any other processes, this is not automatic. We believe that the assumptions considered so far in Sec. III should be enough to prove a central limit theorem in the case of a SIJP with a unique hyperbolically stable attractor. However, this goes beyond the scope of this work and leave the endeavor to future research.

In principle, it is possible to improve the bound in Eq. (53) by exploiting the inner time dependency of the rate functional $\mathcal{I}[\rho, H]$ and, therefore, by considering a time-dependent rescaling of the suboptimal rate matrix in Eq. (52) as

$$(\rho^*(\tau), H^*(\tau)) = (\bar{\rho}(\tau), \alpha_\tau Q(\bar{M}_\tau)) \in \mathcal{S}_{\pi,\phi^*}, \tag{58}$$

where $\phi^*_{xy} = \int_0^\infty e^{-\tau}\alpha_\tau \bar{\eta}_{xy}(\tau)$ cannot be cast into an explicit function of the typical flux φ_{xy} as in Eq. (51). The new bound takes the form

$$I(b) \leqslant I^{**}(b) := \sum_{(x,y)\in\mathcal{E}} \int_0^\infty e^{-\tau}\Gamma(\alpha_\tau)\bar{\eta}_{xy}(\tau)\, d\tau. \tag{59}$$

By selecting specific trajectories of α_τ it should be possible to obtain $I^{**}(b) \leqslant I^*(b)$. In particular, one could look after the best solution for α_τ such that $I^{**}(b)$ is minimal. Variationally, one can show that such an optimal α^*_τ satisfies the closed form

$$\alpha^*_\tau = e^{-\lambda \frac{\sum_{(x,y)\in\mathcal{E}} \beta_{xy}\bar{\eta}_{xy}(\tau)}{\sum_{(x,y)\in\mathcal{E}} \bar{\eta}_{xy}(\tau)}}, \tag{60}$$

where the Lagrange multiplier λ plays the role of a *temperature* or *external parameter* and is fixed by the (now dynamical) constraint

$$b = \sum_{(x,y)\in\mathcal{E}} \beta_{xy} \int_0^\infty e^{-\tau}\alpha^*_\tau \bar{\eta}_{xy}(\tau)\, d\tau. \tag{61}$$

The stationary constraint $\bar{\rho}_\tau \alpha_\tau Q(\bar{M}_\tau) = 0$ selects the rescaling α_τ such that it admits $\bar{\rho}_\tau$ as the instantaneous stationary distribution of $\alpha_\tau Q(\bar{M}_\tau)$ as required by (37) for $(\rho^*(\tau), H^*(\tau))$ and is automatically satisfied as long as $\alpha_\tau \neq 0$ for every $\tau \in \mathbb{R}_+$. However, as mentioned earlier, for the timescale separation and the level-2.5 LDs to hold we need $\tilde{H}(\tau)$ to (i) change slowly [cumulate $O(1)$ changes over in $O(t)$ time or longer] and (ii) smoothly. As a consequence, this further restricts the class of functions α_τ we can admit for the bound in Eq. (59) to hold.

B. SIJP ultimate kinetic uncertainty relation

In this subsection, in order to obtain an explicit result, we restrict ourselves to the—opposite with respect to the previous subsection—case $f = 0$, namely the empirical observable dependence in the SIJP is given by

$$A_t = t^{-1} \sum_{\substack{0\leqslant t'\leqslant t \\ (X_{t'_-}, X_{t'})\in\mathcal{E}}} g_{X_{t'_-}, X_{t'}}, \tag{62}$$

which includes only jump type contributions.

By rephrasing the variational problem in Eq. (48) in terms of the functional $\tilde{\mathcal{I}}[\rho, \eta]$ in Eq. (34) we obtain

$$I(b) = \inf_{\substack{(\rho,\eta)\in\mathcal{S}_{\ell,\phi} \\ \ell, b:\,\mathrm{Eq.}(47)}} \tilde{\mathcal{I}}(\rho, \eta). \tag{63}$$

A suboptimal solution to the display above can be found by admitting the Ansatz

$$\rho^*_x(\tau) = \bar{\rho}_x(\tau) + \epsilon\, \delta_x, \tag{64}$$

$$\eta^*_{xy}(\tau) = (1 + \epsilon\, \Delta)\bar{\eta}_{xy}(\tau), \tag{65}$$

which generalizes the approach of Ref. [52] to our non-Markovian setting. Hence, from Eq. (38) we also have

$$M^*_\tau = \bar{M}_\tau + \epsilon\, \Delta\, \delta m_\tau, \tag{66}$$

with

$$\delta m_\tau = e^\tau \int_\tau^\infty e^{-\tau'} \sum_{(x,y)\in\mathcal{E}} g_{xy}\bar{\eta}_{xy}(\tau), \tag{67}$$

where $\epsilon \ll 1$, $\sum_{x\in\mathcal{S}} \delta_x = 0$ as $\rho^*(\tau)$ is normalized, and $\Delta \in \mathbb{R}_+$. As a consequence, Eq. (47) takes the form

$$b = (1 + \epsilon\, \Delta)\bar{b}, \tag{68}$$

with $\bar{b}$ given in Eq. (46). We notice that the assumption on the rate matrix of the nonhomogeneous Markov process in Eq. (52), which led to the SIJP-KUR, is fundamentally different from the assumption in Eq. (65). It is such because the former allows for fluctuations in the density to influence fluctuations in the flux, which is not true for the latter. [Notice that the SIJP-KUR can also be derived following the steps below by assuming Eq. (64) and $H^*_{xy}(\tau) = (1 + \epsilon\, \Delta)Q_{xy}(\bar{M}_\tau)$ in place of Eq. (65), with $b/\bar{b} = 1 + \epsilon\Delta$.]

Using Eqs. (64)–(66) and expanding up to second order in ϵ, Eq. (34) yields

$$\tilde{\mathcal{I}}[\rho^*, \eta^*] = \frac{\epsilon^2}{2} \sum_{(x,y)\in\mathcal{E}} \int_0^\infty \left(e^{-\tau} \frac{Q_{xy}(\bar{M}_\tau)}{\bar{\rho}_x(\tau)} \tilde{\delta}^2_{xy}(\tau) \right) d\tau + O(\epsilon^3), \tag{69}$$

where we define

$$\tilde{\delta}_{xy}(\tau) := \delta_x - \Delta\, \bar{\rho}_x(\tau) \left(1 - \frac{\delta m_\tau\, Q'_{xy}(\bar{M}_\tau)}{Q_{xy}(\bar{M}_\tau)} \right), \tag{70}$$

with $Q'_{xy}(\bar{M}_\tau) := d/dM\, Q_{xy}(M)|_{M=\bar{M}_\tau}$, and such that

$$\sum_{x\in\mathcal{S}} \tilde{\delta}_{xy}(\tau) = -\Delta \left(1 + \delta m_\tau \sum_{x\in\mathcal{S}} \frac{\bar{\rho}_x(\tau) Q'_{xy}(\bar{M}_\tau)}{Q_{xy}(\bar{M}_\tau)} \right), \tag{71}$$

for all $\tau \in \mathbb{R}_+$. By a straightforward application of the Cauchy-Schwarz inequality in the form

$$\left(\sum_{x\in\mathcal{S}} a_{xy}(\tau)^{-1} \right) \left(\sum_{x\in\mathcal{S}} a_{xy}(\tau) b^2_{xy}(\tau) \right) \geqslant \left(\sum_{x\in\mathcal{S}} b_{xy}(\tau) \right)^2, \tag{72}$$

where we identify $a_{xy}(\tau) = Q_{xy}(\bar{M}_\tau)/\bar{\rho}_x(\tau)$ and $b_{xy}(\tau) = \tilde{\delta}_{xy}(\tau)$ and remark that the inequality is saturated for

$$b_{xy}(\tau) = \frac{\sum_{x\in\mathcal{S}} b_{xy}(\tau)}{a_{xy}(\tau)\sum_{x\in\mathcal{S}} a_{xy}^{-1}(\tau)}. \tag{73}$$

Equation (69) simplifies to

$$\tilde{\mathcal{I}}(\rho^*,\eta^*) = \frac{\epsilon^2\Delta^2}{2}\sum_{\substack{y\in\mathcal{S}\\ y\neq x}}\left(\int_0^\infty e^{-\tau}\frac{\left(1-\delta m_\tau\sum_{x\in\mathcal{S}}\frac{\bar{\rho}_x(\tau)Q'_{xy}(\bar{M}_\tau)}{Q_{xy}(\bar{M}_\tau)}\right)^2}{\sum_{x\in\mathcal{S}}\frac{\bar{\rho}_x(\tau)}{Q_{xy}(\bar{M}_\tau)}}d\tau\right) + O(\epsilon^3). \tag{74}$$

Inverting Eq. (68) to extract $\epsilon\,\Delta$ as a function of b and $\bar{b}$ and replacing this into Eq. (74) finally yields the quadratic rate function bound

$$I(b) \leqslant I^{\triangle}(b) := \frac{(b-\bar{b})^2}{2\bar{b}^2} \times \sum_{\substack{y\in\mathcal{S}\\ y\neq x}}\left(\int_0^\infty e^{-\tau}\frac{\left(1-\delta m_\tau\sum_{x\in\mathcal{S}}\frac{\bar{\rho}_x(\tau)Q'_{xy}(\bar{M}_\tau)}{Q_{xy}(\bar{M}_\tau)}\right)^2}{\sum_{x\in\mathcal{S}}\frac{\bar{\rho}_x(\tau)}{Q_{xy}(\bar{M}_\tau)}}d\tau\right). \tag{75}$$

Eventually, by expanding $I(b)$ in Eq. (63) up to second order in ϵ, using the definition of asymptotic variance in Eq. (55), and comparing the form obtained to Eq. (74) we obtain

$$\frac{\mathrm{var}(b)}{\bar{b}^2} \geqslant \left(\sum_{\substack{y\in\mathcal{S}\\ y\neq x}}\int_0^\infty e^{-\tau}\frac{\left(1-\delta m_\tau\sum_{x\in\mathcal{S}}\frac{\bar{\rho}_x(\tau)Q'_{xy}(\bar{M}_\tau)}{Q_{xy}(\bar{M}_\tau)}\right)^2}{\sum_{x\in\mathcal{S}}\frac{\bar{\rho}_x(\tau)}{Q_{xy}(\bar{M}_\tau)}}d\tau\right)^{-1}. \tag{76}$$

This formula states that the asymptotic precision of the generalized flux B_T is controlled by a dynamically discounted average lifetime in the instantaneous—valid for all $\tau \in \mathbb{R}_+$—steady state. Equation (76) extends the UKUR valid for Markov jump processes [52] (see also Ref. [53]) to non-Markov SIJPs. It is immediate to check that for $A_t = 0$—this eliminates the functional dependence of the rate matrix upon the empirical observable—$\delta m_\tau = 0$ and all temporal dependences disappear, retrieving the usual UKUR [52,53].

We conclude by remarking that although in the Markov case UKUR is tighter than KUR—The weighted harmonic mean is less than or equal to the corresponding weighted arithmetic mean—in the SIJP process case it is no longer a general consequence.

VI. THERMODYNAMIC BOUNDS

In this section, we turn the focus on a subset of observables of the generalized flux B_T in Eq. (45), namely generalized current observables of the form

$$J_T = \sum_{x<y}\beta_{xy}\mathfrak{J}_{xy}(T), \tag{77}$$

where we denote the edge currents

$$\mathfrak{J}_{xy}(T) = \Phi_{xy}(T) - \Phi_{yx}(T). \tag{78}$$

The generalized current observable in (77) is obtained from B_T in Eq. (45) by setting $\beta_{xy} = -\beta_{yx}$ for pairs $(x,y) \in \mathcal{E}$. Restricting to these observables allows us to refine the SIJP-KUR in Eq. (56).

In order to derive explicit results as done in the SIJP-KUR case, we assume a unique stationary pair $\mathcal{A}_{\mathrm{stat}} = \{(\pi,\varphi)\}$ and restrict the type of SIJP dependence we consider to Eq. (49), that is, only empirical state contributions are allowed to influence the SIJP dynamics. Furthermore, similarly to the kinetic bounds presented in the previous section, we stress here as well that the thermodynamic bound presented below should be understood as restricted to the limit of validity of the central limit theorem for SIJPs.

First, we derive a large deviation rate function for the edge current observable in Eq. (78), opportunely contracting the level-2.5 large deviation rate function (36)—with $\tilde{\mathcal{I}}[\rho,\eta]$ in Eq. (34). For notation purposes, it will be useful to refer to edge current fluctuations using the rescaled and time-reversed characterization of the flux ϕ in Eq. (40) and write

$$\mathfrak{j}_{xy} = \phi_{xy} - \phi_{yx}, \tag{79}$$

$$= \int_0^\infty e^{-\tau}\upsilon_{xy}(\tau)\,d\tau, \tag{80}$$

where we denote

$$\upsilon_{xy}(\tau) = \eta_{xy}(\tau) - \eta_{yx}(\tau), \tag{81}$$

as the antisymmetric part of the time-dependent flux.

A. SIJP entropy bound on edge currents

We perform a first contraction of the rate functional $\tilde{\mathcal{I}}[\rho,\eta]$ over all possible instantaneous fluxes $\eta(\tau)$ such that the constraint (81) is satisfied along with the conservation of probability in Eq. (37). To do so, we introduce the extended Lagrangian

$$\begin{aligned}\mathcal{L}[\eta,\zeta,\varepsilon] :=& \sum_{(x,y)\in\mathcal{E}}\int_0^\infty e^{-\tau}\rho_x(\tau)Q_{xy}(M_\tau)\Gamma\left(\frac{\eta_{xy}(\tau)}{\rho_x(\tau)Q_{xy}(M_\tau)}\right)d\tau\\ &+\int_0^\infty e^{-\tau}\zeta_{xy}(\tau)(\upsilon_{xy}(\tau)-\eta_{xy}(\tau)+\eta_{yx}(\tau))\,d\tau\\ &+\int_0^\infty e^{-\tau}\varepsilon_y(\tau)\left(\sum_{x\in\mathcal{S}}\eta_{xy}(\tau)\right)d\tau,\end{aligned} \tag{82}$$

which includes the fields $\zeta(\tau)$ and $\varepsilon(\tau)$ fixing the respective constraints. We notice that the symmetry inherent in the definition of flux $\sum_x \eta_{xy}(\tau) = -\sum_x \eta_{yx}(\tau)$ is inherited by the corresponding Lagrangian field, which then satisfies the relation $\varepsilon_y(\tau) = -\varepsilon_x(\tau)$. Additionally, we denote $\triangle_{xy}(\tau) = \exp(\zeta_{xy}(\tau) - \zeta_{yx}(\tau))$. Taking variations of $\eta_{xy}(\tau)$ over the extended Lagrangian in Eq. (82), the optimal flux can be

explicitly written in terms of its asymmetric component, the time-dependent current $v(\tau)$, and its symmetric component, the time-dependent traffic $u(\tau)$, as

$$\eta^*_{xy}(\tau) = \frac{v_{xy}(\tau) + u_{xy}(\tau)}{2}, \tag{83}$$

where the traffic is explicitly given by

$$u_{xy}(\tau) = \sqrt{v^2_{xy}(\tau) + 4\rho_x(\tau)\rho_y(\tau)Q_{xy}(M_\tau)Q_{yx}(M_\tau)}. \tag{84}$$

Eventually, replacing $\eta^*(\tau)$ from Eq. (83) into the rate functional $\tilde{\mathcal{I}}[\rho, \eta]$ and using the relation between "log" and "arcsinh" functions, we can write

$$\tilde{\tilde{\mathcal{I}}}[\rho, v] := \sum_{x<y} \int_0^\infty e^{-\tau} \Psi\big(v_{xy}(\tau), v^\rho_{xy}(\tau), u^\rho_{xy}(\tau)\big)\, d\tau, \tag{85}$$

with the definitions

$$\begin{aligned} v^\rho_{xy}(\tau) &:= \rho_x(\tau)Q_{xy}(M_\tau) - \rho_y(\tau)Q_{yx}(M_\tau) \\ u^\rho_{xy}(\tau) &:= \rho_x(\tau)Q_{xy}(M_\tau) + \rho_y(\tau)Q_{yx}(M_\tau) \\ \Psi(p,q,r) &:= p\,\mathrm{arcsinh}\frac{p}{\sqrt{r^2-q^2}} - p\,\mathrm{arcsinh}\frac{q}{\sqrt{r^2-q^2}} \\ &\quad - (\sqrt{p^2-q^2+r^2} - r), \end{aligned}$$

where the last form also appears in the time-homogeneous Markov case [48] (without time dependence, of course).

Thanks to the explicit characterization of the edge traffic that comes naturally from the minimization above, we can now write the rate function of the edge currents as

$$I(\mathrm{j}) = \inf_{\substack{(\ell,\rho(\tau),v(\tau)) \\ \ell:\text{Eqs.(38),(39)} \\ \mathrm{j}:\text{Eq.(80)}}} \sum_{x<y} \int_0^\infty e^{-\tau} \Psi\big(v_{xy}(\tau), v^\rho_{xy}(\tau), u^\rho_{xy}(\tau)\big)\, d\tau. \tag{86}$$

This object is easily bounded from above using a quadratic function obtained from considering as a suboptimal solution for the asymptotic density the stationary density, i.e.,

$$\ell^*_x = \pi_x, \tag{87}$$

which, in turn, implies $\rho^*(\tau) = \bar{\rho}(\tau)$. Then, Eq. (86) simplifies to

$$I(\mathrm{j}) \leqslant \inf_{\substack{v(\tau) \\ \mathrm{j}:\text{ Eq.(80)}}} \sum_{x<y} \int_0^\infty e^{-\tau} \frac{\big(v_{xy}(\tau) - v^{\bar{\rho}}_{xy}(\tau)\big)^2}{4\big(v^{\bar{\rho}}_{xy}(\tau)\big)^2} \sigma^{\bar{\rho}}_{xy}(\tau)\, d\tau, \tag{88}$$

where we interpret

$$\sigma^{\bar{\rho}}_{xy}(\tau) := 2v^{\bar{\rho}}_{xy}(\tau)\mathrm{arcsinh}\frac{v^{\bar{\rho}}_{xy}(\tau)}{\sqrt{\big(u^{\bar{\rho}}_{xy}(\tau)\big)^2 - \big(v^{\bar{\rho}}_{xy}(\tau)\big)^2}}, \tag{89}$$

as the SIJP instantaneous edge entropy production. This interpretation is valid because the empirical occupation measure evolves slowly: The generator $Q(\bar{M}_\tau)$ changes quasistatically, so that the underlying SIJP can equilibrate at each frozen time and $\bar{\rho}(\tau)$ is stationary with respect to $Q(\bar{M}_\tau)$ [106,107]. Finally, to get rid of the last minimization, one could set the suboptimal $v^*(\tau) = \mathrm{j}$ (constant over time) or replace the optimal minimizing instantaneous current, which takes the explicit form

$$\begin{aligned} v^*_{xy}(\tau) &= \frac{v^{\bar{\rho}}_{xy}(\tau)\big(\mathrm{j} - \int_0^\infty e^{-\tau'} v^{\bar{\rho}}_{xy}(\tau')\, d\tau'\big)}{\int_0^\infty e^{-\tau'} v^{\bar{\rho}}_{xy}(\tau')\, \mathrm{arcsinh}^{-1}\left(\frac{v^{\bar{\rho}}_{xy}(\tau')}{\sqrt{\big(u^{\bar{\rho}}_{xy}(\tau')\big)^2 - \big(v^{\bar{\rho}}_{xy}(\tau')\big)^2}}\right) d\tau'} \\ &\quad \times \mathrm{arcsinh}^{-1}\left(\frac{v^{\bar{\rho}}_{xy}(\tau)}{\sqrt{\big(u^{\bar{\rho}}_{xy}(\tau')\big)^2 - \big(v^{\bar{\rho}}_{xy}(\tau')\big)^2}}\right) + v^{\bar{\rho}}_{xy}(\tau). \end{aligned} \tag{90}$$

Regardless of the specific choice for the current $v^*(\tau)$, the bound (88) has a specific meaning: the instantaneous entropy production rate of the SIJP, which we can interpret as a measure of dynamical dissipation, bounds all asymptotic current fluctuations.

B. SIJP thermodynamic uncertainty relation

We now restrict our analysis to the scalar observable J_T in Eq. (77), whose most general realization can be written as

$$j = \sum_{x<y} \beta_{xy} \mathrm{j}_{xy}, \tag{91}$$

where j_{xy} is defined in Eq. (79). We aim at calculating an upper bound of the rate function $I(j)$ as follows:

$$I(j) \leqslant \inf_{\substack{v(\tau) \\ j:\text{Eq.(91)}}} \sum_{x<y} \int_0^\infty e^{-\tau} \frac{\big(v_{xy}(\tau) - v^{\bar{\rho}}_{xy}(\tau)\big)^2}{4\big(v^{\bar{\rho}}_{xy}(\tau)\big)^2} \sigma^{\bar{\rho}}_{xy}(\tau)\, d\tau. \tag{92}$$

We do so by introducing a specific extended Lagrangian incorporating the constraint (91) via a Lagrangian field. The minimization is explicit and yields the optimal instantaneous edge current

$$v^*_{xy}(\tau) = v^{\bar{\rho}}_{xy}(\tau) + \frac{(j - \bar{j})\beta_{xy}\big(v^{\bar{\rho}}_{xy}(\tau)\big)^2}{\sigma^{\bar{\rho}}_{xy}(\tau) \sum_{x<y} \int_0^\infty e^{-\tau'} \frac{(\beta_{xy} v^{\bar{\rho}}_{xy}(\tau'))^2}{\sigma^{\bar{\rho}}_{xy}(\tau')}\, d\tau'}, \tag{93}$$

where

$$\bar{j} := \int_0^\infty e^{-\tau} j^{\bar{\rho}}_\tau\, d\tau, \tag{94}$$

$$j^{\bar{\rho}}_\tau := \sum_{x<y} \beta_{xy} v^{\bar{\rho}}_{xy}(\tau). \tag{95}$$

Inserting $v^*(\tau)$ from Eq. (93) in Eq. (92) we get

$$I(j) \leqslant \frac{(j - \bar{j})^2}{4 \sum_{x<y} \int_0^\infty e^{-\tau'} \frac{(\beta_{xy} v^{\bar{\rho}}_{xy}(\tau'))^2}{\sigma^{\bar{\rho}}_{xy}(\tau')}\, d\tau'}. \tag{96}$$

Eventually, by applying the Cauchy-Schwarz relation in Eq. (72), with $a^{-1}_{xy}(\tau) = \sigma^{\bar{\rho}}_{xy}(\tau)$ and $b_{xy}(\tau) = \beta_{xy} v^{\bar{\rho}}_{xy}(\tau)$ in its saturated form, we obtain

$$I(j) \leqslant I^{\square}(j) := \frac{(j - \bar{j})^2}{4} \left(\int_0^\infty e^{-\tau} \frac{\big(j^{\bar{\rho}}_\tau\big)^2}{\sigma_\tau}\, d\tau \right)^{-1}, \tag{97}$$

where σ_τ is the total instantaneous entropy production rate of the SIJP, given by

$$\sigma_\tau = \sum_{x<y} \sigma^{\bar{\rho}}_{xy}(\tau), \tag{98}$$

$$= \sum_{x<y} \log\left(\frac{\bar{\rho}_x(\tau)Q_{xy}(\bar{M}_\tau)}{\bar{\rho}_y(\tau)Q_{yx}(\bar{M}_\tau)}\right) v^{\bar{\rho}}_{xy}(\tau), \tag{99}$$

where the second line can be derived from Eq. (89) using the "arcsinh- log" relation.

Eventully, the SIJP-TUR is obtained by expanding the left-hand side of Eq. (97) up to second order locally around $\bar{j}$. This yields

$$\frac{(\mathrm{var}(j))}{\bar{j}^2} \geqslant 2\int_0^\infty e^{-\tau}\frac{\left(j^{\bar{\rho}}_\tau\right)^2}{\sigma_\tau}\,d\tau, \tag{100}$$

as a generalization to SIJPs of the well-known TUR for continuous-time Markov processes [47,48]. In analogy to the Markovian case, Eq. (100) shows that reducing current fluctuations in a SIJP incurs a minimal cost in terms of the instantaneous entropy production σ_τ. Notice also that it is not possible to disentangle the instantaneous entropy cost from the instantaneous generalized current. As a consequence, we can interpret the instantaneous entropy production as an additional discount for the SIJP, on top of the exponential one, that is necessary to reach a certain level of precision for the current. We further note that in the time-homogeneous case, the Eq. (100) clearly reduces to the standard TUR of Markov jump processes.

The present TUR and entropy production construction is similar in spirit to Refs. [91,92] but different in implementation. The similarity is that these works also introduce specific auxiliary processes and quantify irreversibility through a comparison of their path weights, or dynamical costs, with respect to the original process. Here, however, the non-Markovianity does not relate to Langevin processes and does not come from colored noise or an explicit memory kernel but from the dependence of the jump rates on empirical observables of the past trajectory. The instantaneous cost entering our entropy-production rate is obtained from the level-2.5 description of empirical occupations and fluxes after introducing an auxiliary time-inhomogeneous Markov generator that makes the atypical SIJP trajectory typical.

VII. ILLUSTRATIVE EXAMPLES

In this section, we illustrate the general results above with several simple concrete examples of two- and three-state systems.

We initially focus on a two-state model where some analytical treatment of the fluctuations can be done, although for the exact rate functions shown the solution is approachable only by numerical methods. We will derive so-called level-2 LD functions (LD functions for the state occupation) and compare them to simulations and to the upper bounds given by Markov approximations. Along with this, we will study the LD functions of the unidirectional flux in the system and compare them to simulations and to the SIJP-KUR upper bounds given in Eq. (59).

Then, we discuss three-state examples which have a nonequilibrium stationary state. We will see that even for such simple systems an efficient method to compute numerically the exact rate function for the total current through the system is out of reach at the moment and we leave this as an open problem to study in future work. We therefore compare simulations of the process with the SIJP-TUR of Eq. (97), testing the validity of the latter as an upper bound to the true fluctuations.

Finally, in the last example, we go back to a simpler, effectively one-state, stochastic birth process where the specific self-interaction considered leads to two (not one) basins of attractions for the long-time empirical birth rate. The large deviations of such an observable reveals striking dynamical features, such as a change of convergence speed.

A. Feedback two-state SIJP

We consider SIJPs for a two-state system, $\mathcal{S} = \{0, 1\}$, with rate matrix

$$Q(A_t) := \begin{pmatrix} -1 - h(L_1(t)) & 1 + h(L_1(t)) \\ 1 & -1 \end{pmatrix}. \tag{101}$$

The specific dynamics depends on the dependence on the empirical distribution through the function h. We study two cases,

$$h^{(L)}(L_1(t)) = L_1(t), \tag{102}$$

$$h^{(E)}(L_1(t)) = \exp\left[\alpha\, L_1(t)\right], \tag{103}$$

for $\alpha > 0$. This describes a spin whose flip-up rate grows either linearly (L) or exponentially (E) with its past occupation. Either of these forms define a SIJP as in Eq. (6), where the empirical observable $A_t = L_1(t)$ is defined by setting $(f_0, f_1) = (0, 1)$ and $g = 0$ in Eq. (6).

In both cases, the stationary state is found by explicitly solving Eq. (10). In the linear case (L), the explicit solution to the stationary distribution is

$$\left(\pi_0^{(L)}, \pi_1^{(L)}\right) = \left(\frac{3-\sqrt{5}}{2}, \frac{\sqrt{5}-1}{2}\right), \tag{104}$$

which shows that the occupation of state 0 is approximately 0.382, compared with 0.618 for state 1. In the second case, the stationary state $(\pi_0^{(E)}, \pi_1^{(E)})$ is solution of the following transcendental equation:

$$2 + e^{\alpha(1-\pi_0^{(E)})} = \frac{1}{\pi_0^{(E)}}, \tag{105}$$

which, depending on α, returns a value in the interval $0 < \pi_0^{(E)} < \frac{1}{3}$, where the right boundary is obtained by setting $\alpha = 0$ and the left for $\alpha \to \infty$. The larger α, the less time the system typically spends in state 0 (for $\alpha = 2$ we have $\pi_0^{(E)} \approx 0.13$).

1. Level-2 large deviations

In the following, we study the level-2 rate function $I_2(\ell)$ for the occupation distribution at 0, that is, $L_0(t)$, along with the optimal time-dependent fluxes that lead to its form.

For such a simple two-state system we can simplify notation. In the following, we drop the superscript denoting the specific form of h, as the treatment is general, and restore it when deriving specific results. For compactness we also define $\pi := \pi_0$, $\rho_\tau := \rho_0(\tau)$, and recall that M_τ is the the empirical observable A_t over the backward logarithmic timescale, and due to normalization we have $\pi_1 = 1-\pi$, $\rho_1(\tau) = 1-\rho_\tau$. Additionally, by conservation of probability, the overall net current is zero and therefore $\eta_{01}(\tau) = \eta_{10}(\tau) =: \eta_\tau$. All this helps simplify Eq. (34), which takes the form

$$\tilde{\mathcal{I}}[\rho,\eta] = \int_0^\infty e^{-\tau}\left(\rho_\tau(1+h(M_\tau))\Gamma\left(\frac{\eta_\tau}{\rho_\tau(1+h(M_\tau))}\right) + (1-\rho_\tau)\Gamma\left(\frac{\eta_\tau}{1-\rho_\tau}\right)\right)d\tau. \tag{106}$$

The level-2 rate function is obtained by minimizing $\tilde{\mathcal{I}}[\rho,\eta]$ over (ρ_τ,η_τ) under the constraints given by Eqs. (39) and (38). A free minimization over η_τ yields the following optimal flux:

$$\eta_\tau^* = \sqrt{\rho_\tau(1-\rho_\tau)(1+h(M_\tau))}, \tag{107}$$

and the simplified functional

$$\tilde{\mathcal{I}}[\rho,\eta^*] = \int_0^\infty e^{-\tau}\left(\sqrt{\rho_\tau(1+h(M_\tau))}-\sqrt{1-\rho_\tau}\right)^2 d\tau. \tag{108}$$

Noticeably, its quadratic form guarantees that there is a unique minimizing trajectory ρ_τ^* leading to the realization of the fluctuation $L_0(t) = \ell$ in the long-time limit. In the typical state, obtained by equating the last display to 0, we readily get

$$\rho_\tau^* = 1 - M_\tau^* = \pi, \tag{109}$$

along with the corresponding optimal constant flux in Eq. (107).

Minimizing Eq. (108) over ρ_τ with the constraints (38) and (39) does not lead to an explicit formula for ρ_τ^* and neither for the level-2 rate function. Yet, a closed form solution for ρ_τ^* can be found by introducing the extended Lagrangian

$$\mathcal{L}[\rho_\tau,\zeta_\tau,\varepsilon] := \int_0^\infty e^{-\tau}\big[(\sqrt{\rho_\tau(1+h(M_\tau))}-\sqrt{1-\rho_\tau})^2 + \zeta_\tau(\dot{M}_\tau - M_\tau + 1 - \rho_\tau) + \varepsilon e^{-\tau}(\rho_\tau - \pi)\big]d\tau, \tag{110}$$

where $\zeta_\tau \in \mathbb{R}$ for all $\tau \in \mathbb{R}_+$ and $\varepsilon \in \mathbb{R}$ are the Lagrangian fields enforcing the constraints. In particular, notice that Eq. (38) has been enforced as a differential equation rather than an integral equation, which is more handy.

Minimizing variationally Eq. (110) and replacing the Euler-Lagrange equation for ρ_τ in the costate equation for M_τ we obtain, after some algebra, the following system of equations:

$$\dot{M}_\tau = M_\tau - 1 - \rho_\tau, \tag{111}$$

$$\dot{\rho}_\tau = r(M_\tau,\rho_\tau), \tag{112}$$

where the function r takes the form

$$r(M_t,\rho_t) = -\frac{dh(M_\tau)}{dM_\tau}\left(\frac{\rho_\tau^3(1-2M_\tau)+\rho_\tau^2(3M_\tau-1-2\sqrt{1+h(M_\tau)}1-M_\tau\sqrt{\rho_\tau(1-\rho_\tau)})}{1+h(M_\tau)} + \frac{M_\tau\rho_\tau(2\sqrt{1+h(M_\tau)}\sqrt{\rho_\tau(1-\rho_\tau)}-1)}{1+h(M_\tau)}\right). \tag{113}$$

The system of equations is to be solved with the initial condition $M_0 = 1-\ell$ and an initial condition for ρ_0, which can only be found *a posteriori* by setting the integral constraint over the solution ρ_τ to match ℓ, that is, $\int_0^\infty e^{-\tau}\rho_\tau\, d\tau = \ell$.

For both cases, the system of ordinary differential equations above is highly unstable because all its fixed points, $(\rho, M) \in \{(0,0),(\pi,\pi),(1,1)\}$, are unstable. The phase diagrams in Fig. 1 show that (0,0) and (1,1) act as saddle points whose one-dimensional stable manifolds intersect at (π,π). Consequently, all initial conditions (ρ_0, M_0) must lie on this stable manifold (black solid lines). It is therefore sufficient to compute two sets of trajectories, (ρ_τ, M_τ), with initial conditions chosen near (π,π) on the stable manifold—one on each side of the maximum corresponding to the two saddle points. Since these trajectories evolve along the stable manifold and asymptotically converge to their respective saddle points, together they span all possible initial conditions. Hence, to study trajectories starting from different initial states, it suffices to vary the backward-time parameter τ until (ρ_τ, M_τ) reaches the desired condition. We plot the two sets of trajectories for both systems under analysis in Fig. 2. [Numerically, to deal with instability, we study the system of differential equations starting in the vicinity of the saddle points and follow trajectories as they converge to (π,π).]

In Fig. 3, we compare the level-2 rate function $I_2(\ell)$ for the SIJP, obtained by plugging (ρ_τ^*, M_τ^*) for each value $L_0(T) = \ell$ into Eq. (108) with Monte Carlo simulations—whose detailed explanations is deferred to Appendix C—and with the Markov upper bound $I_{\mathrm{DV}}(\ell)$. The latter is provided by the naïve (constant) suboptimal solution

$$\rho_\tau^* = 1 - M_\tau^* = \ell, \tag{114}$$

which would be the correct Ansatz, leading to the Donsker-Varadhan rate function, if the process were Markov and time-homogeneous.

As expected, in a neighborhood of π, where the rate functions attain their minimum and vanish, the Markov upper bound closely matches the exact rate function. However, for fluctuations further away from π, $I_2(\ell)$ departs

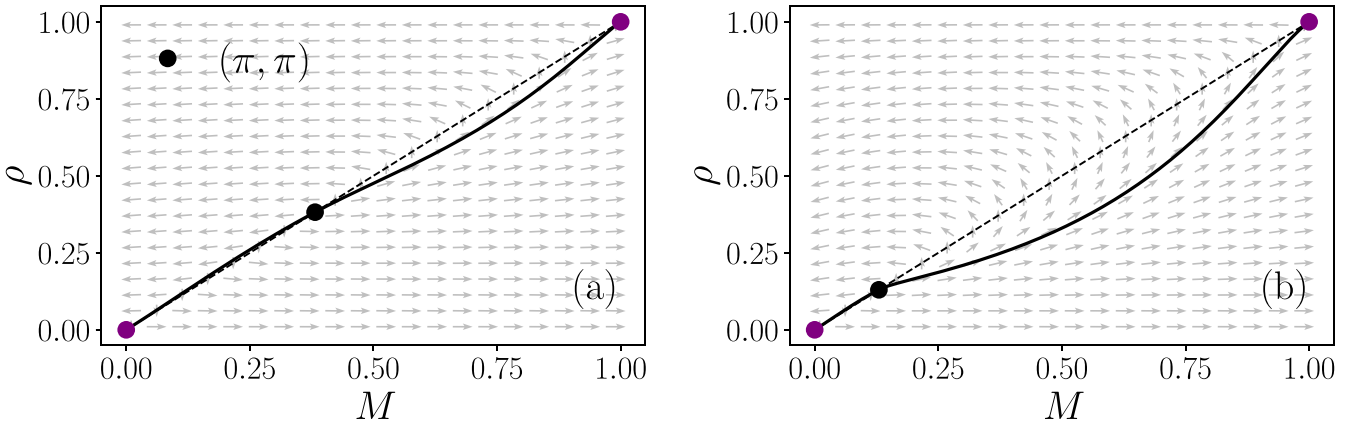


FIG. 1. Phase portraits of the optimal trajectories (ρ_τ, M_τ) solving Eqs. (111) and (112) for the two-state SIJP with (a) linear feedback $h^{(L)}(L_1) = L_1$ and (b) exponential feedback $h^{(E)}(L_1) = \exp(\alpha L_1)$ with $\alpha = 2$. Fixed points at (0,0), (π, π), and (1,1) are shown, with (0,0) and (1,1) acting as saddle points (purple circles). The black solid line denotes the one-dimensional stable manifold connecting the saddle points through (π, π), along which all admissible optimal trajectories lie. Arrows indicate the direction of increasing reversed time τ.

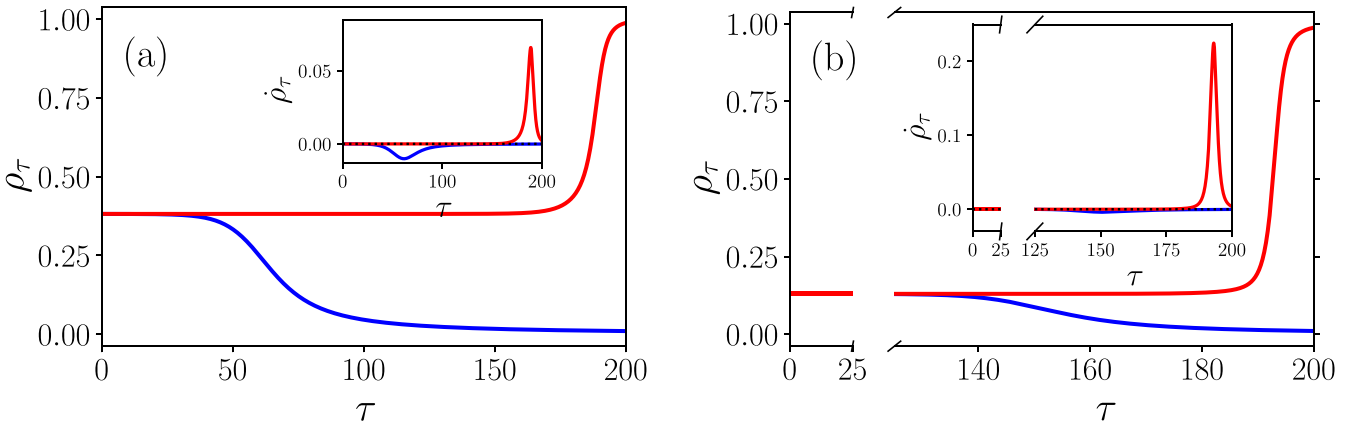


FIG. 2. Optimal trajectories (ρ_τ, M_τ) along the stable manifold for the two-state SIJP with (a) linear and (b) exponential feedback. Trajectories are obtained by integrating, backward with respect to time τ, Eqs. (111) and (112) from initial conditions close to the saddle points: the red trajectory starting close to (1,1) and the blue trajectory starting close to (0,0). Each point on the curves corresponds to a distinct initial value $M_0 = 1-\ell$, generating a fluctuation of the empirical occupation measure $L_0(T) = \ell$ in the (physical) long-time limit. Trajectories converge to (π, π) as $\tau \to 0$.

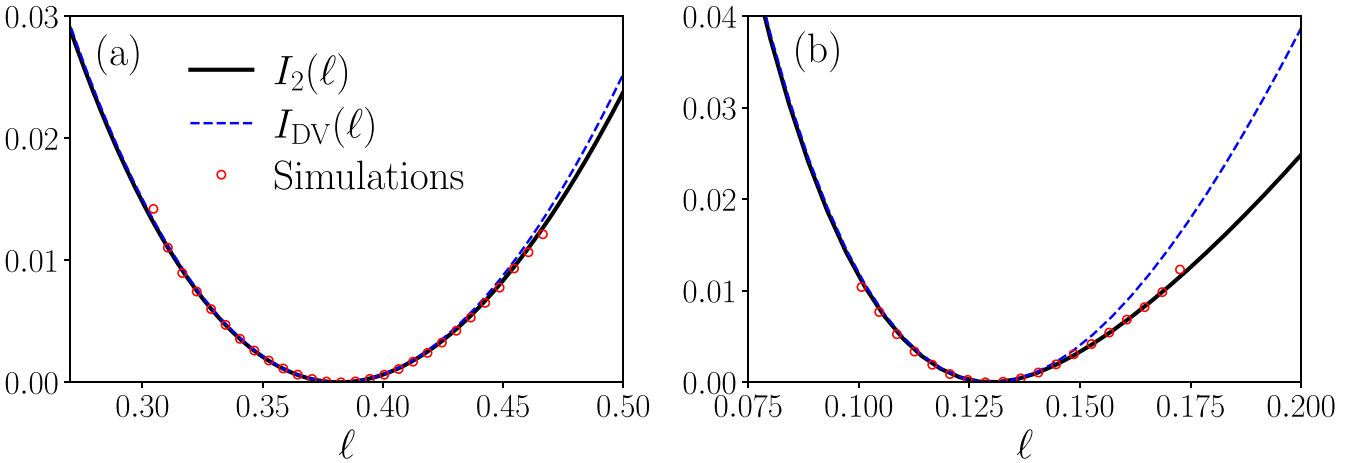


FIG. 3. Level-2 large deviation rate function $I_2(\ell)$ of the empirical occupation measure $L_0(T)$ for the two-state SIJP with (a) linear and (b) exponential feedback. Solid black lines show the exact rate function obtained by evaluating Eq. (108) along the optimal trajectories (ρ^*_τ, M^*_τ). Symbols correspond to Monte Carlo simulations. The dashed curve shows the Markov (Donsker-Varadhan) upper bound $I_{\rm DV}(\ell)$ obtained from the constant-density Ansatz $\rho_\tau = \ell$. Deviations from the Markov bound become pronounced for large fluctuations away from π.

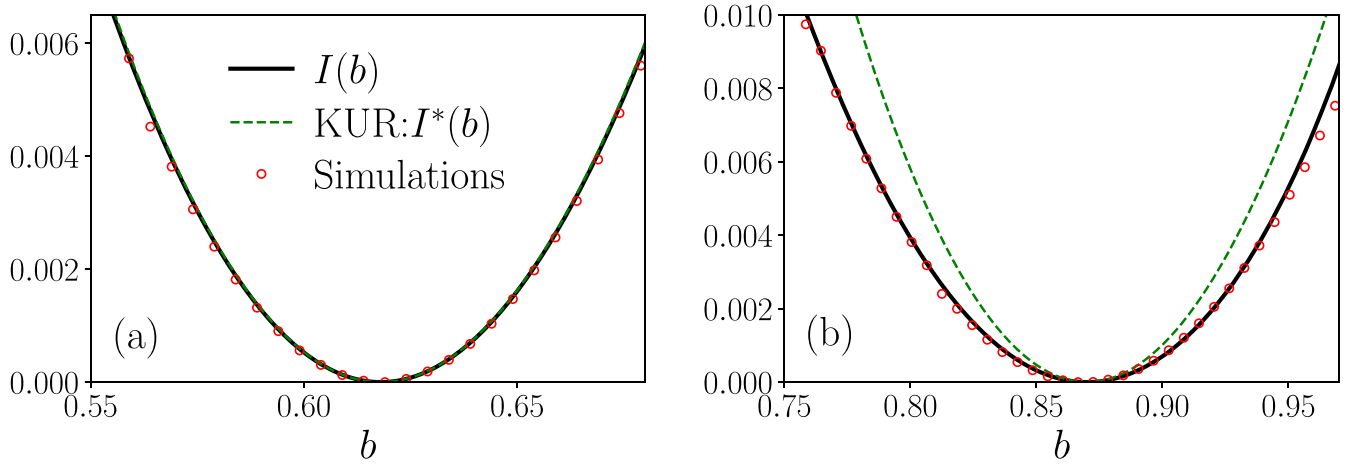


FIG. 4. Large deviation rate function $I(b)$ of the unidirectional flux $B_T = \Phi_{01}(T)$ for the two-state SIJP with (a) linear and (b) exponential feedback. Solid black lines denote the exact rate functions obtained by minimizing $\tilde{\mathcal{I}}[\rho, \eta]$ under the flux constraint. Symbols correspond to Monte Carlo simulations. The dashed curve shows the SIJP-KUR bound given by Eq. (53), which correctly captures typical fluctuations and provides an upper bound on rare events.

from $I_{\rm DV}(\ell)$. This discrepancy arises because LDs from the mean require increasingly abrupt changes in the optimal trajectories (ρ_τ, M_τ), as illustrated in Fig. 2. In particular, trajectories producing fluctuations with $\ell > \pi$ evolve more rapidly—exhibiting larger slopes in absolute value—so that significant deviations from π are reached at earlier times, when the exponential discount $e^{-\tau}$ is still appreciable. As a result, the right tails of $I_2(\ell)$ and $I_{\rm DV}(\ell)$ differ markedly, whereas for $\ell < \pi$ the left tails do not show relevant discrepancies.

2. *Flux large deviations and the SIJP-KUR*

In the following, we carry out a similar analysis to derive the unidirectional flux rate function $I(b)$, where b is an instance of the generalized observable $B_T = \Phi_{01}(T)$ in (45). The minimization over $\tilde{\mathcal{I}}[\rho, \eta]$ in Eq. (106) requires different constraints for (ρ_τ, η_τ). First, we must impose Eq. (47) where $\phi_{01} = \phi_{10}$ is fixed by Eq. (40). Minimizing over η yields

$$\eta^*_\tau = \sqrt{\rho_t(1-\rho_t)(1+h(M_\tau))}\kappa, \tag{115}$$

as the optimal flux, where κ is the Lagrange parameter fixing the constraint (47) playing a useful reparametrization of the flux $B_T = b$.

The next step is to minimize the resulting functional $\tilde{\mathcal{I}}[\rho, \eta^*]$ with respect to ρ_τ with the constraint given by Eq. (38) (with the specific f and g introduced at the beginning of this section). Yet, there is no fixed initial condition as we do not *a priori* know the optimal density that leads to a specific fluctuation of the flux. To circumvent this obstacle, we consider all possible densities—as done above to derive the level-2 rate function—and by sweeping over them keep only the trajectories (ρ^*_τ, M^*_τ) that minimize the large deviation functional for a specific flux b. This procedure returns the rate functions represented as solid black lines in Fig. 4.

These exact rate functions are matched at the level of typical fluctuations by Monte Carlo simulations and are upper bounded by the SIJP-KUR function in Eq. (53).

B. Occuaption-self-induced rotatory three-state SIJP

Current large deviations and the SIJP-TUR

We consider the three-state SIJP, $\mathcal{S} = \{0, 1, 2\}$, with rate matrix

$$Q(A_t) := \begin{pmatrix} -1 & h(L_1(t)) & 1-h(L_1(t)) \\ 1-h(L_1(t)) & -1 & h(L_1(t)) \\ h(L_1(t)) & 1-h(L_1(t)) & -1 \end{pmatrix}, \tag{116}$$

where

$$h(L_1(t)) = \frac{1}{2}\left[1 + \frac{\tanh[\alpha(L_1(t) - 1/2)]}{\tanh[\alpha/2]}\right], \tag{117}$$

with $\alpha > 0$. This describes a system with a self-induced nonlinear push and cyclic transitions. The dynamics tends to accelerate when state 1 is visited frequently. Yet, the more the particle lingers in a region, the harder it is to sustain a current. Similarly to the previous subsection, the empirical observable $A_t = L_1(t)$ is obtained by setting $(f_0, f_1, f_2) = (0, 1, 0)$ and $g = 0$ in Eq. (6). Given the inherent rotational symmetry of the system, the stationary distribution solution of Eq. (10) is uniform over $\mathcal{S}$.

We now focus on the total current observable J_T in Eq. (77) and aim at studying its large deviation function. A priori, given the self-interacting nature of the process, it is not even obvious what the typical behavior of J_T is in the long-time limit as time-dependent dynamics might lead to deviations from the stationary current of the Markov generator $Q(\bar{a})$. *A posteriori*, however, after analyzing the time-dependent system current averaged over a large ensemble of trajectories, we conclude that the averaged SIJP current is constant over time and, therefore

$$\bar{j} = 3\mathrm{j} = 3((\pi \circ Q(\bar{a}))_{01} - (\pi \circ Q(\bar{a}))_{10}) \approx -0.42, \tag{118}$$

where thanks to the rotational symmetry typical edge currents are all equivalent.

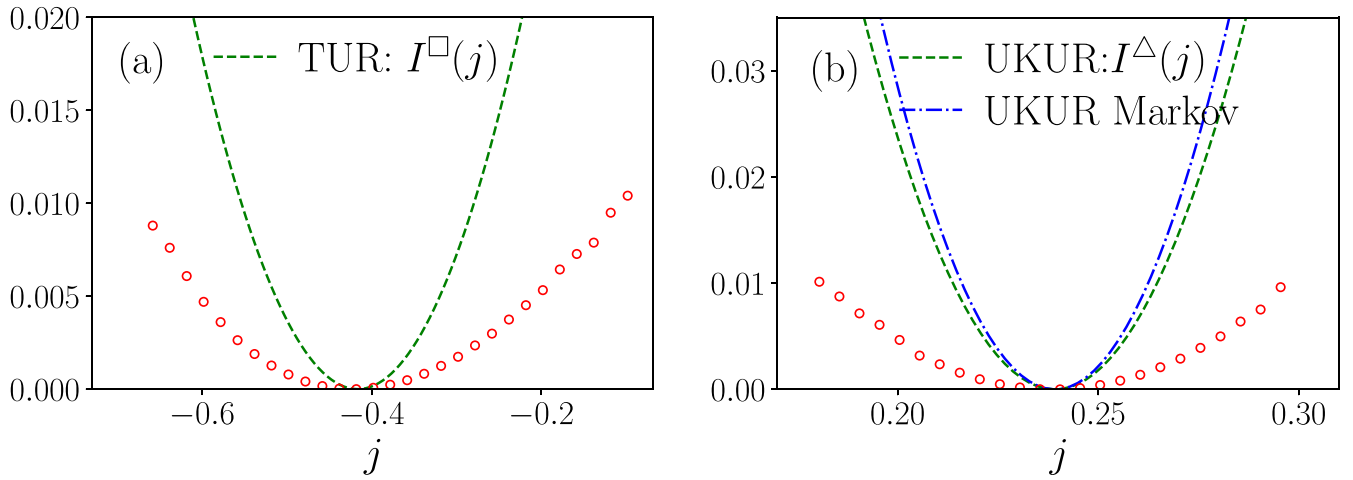


FIG. 5. Large deviation behavior of the total current J_T in three-state SIJPs. (a) Rate function of the total current for the SIJP with state-dependent feedback $h(L_1)$ defined in Eq. (117), obtained from Monte Carlo simulations (symbols) and compared with the SIJP-TUR upper bound $I^{\square}(j)$ given by Eq. (97) (solid line). (b) Current fluctuations for the SIJP with jump-dependent feedback $A_t = J_t$, compared with the SIJP-UKUR bound $I^{\triangle}(j)$ in Eq. (122) (solid line). The dashed line shows the corresponding Markov UKUR for comparison.

This scenario is too complicated to attempt analytically the variational minimization in Eq. (86) and therefore we can only accumulate statistics from simulating the process for a long-enough time and compare these simulations to the SIJP-TUR rate function bound $I^{\square}(j)$ in Eq. (97) [see Fig. 5(a)]. We remark that, for this specific system, the SIJP-TUR bound reduces to the well-known Markov TUR bound as $j^{\bar{p}}_{\tau} = \bar{j}$ and $\sigma_\tau = \sigma$, where

$$\sigma = \sum_{(x,y)\in\mathcal{E}} (\pi \circ Q(\bar{a}))_{xy} \log\left(\frac{(\pi \circ Q(\bar{a}))_{xy}}{(\pi \circ Q(\bar{a}))_{yx}}\right) \approx 0.37, \quad (119)$$

is the well-known average rate of entropy production for the Markov jump process with generator $Q(\bar{a})$ in the stationary state.

C. Current-self-induced rotatory three-state SIJP

Current large deviations and the SIJP-UKUR

We consider another three-state SIJP, $\mathcal{S} = \{0, 1, 2\}$, with rate matrix

$$Q(A_t) := \begin{pmatrix} -1 & e^{-\alpha J_t} & 1 - e^{-\alpha J_t} \\ 1 - e^{-\alpha J_t} & -1 & e^{-\alpha J_t} \\ e^{-\alpha J_t} & 1 - e^{-\alpha J_t} & -1 \end{pmatrix}, \quad (120)$$

with $\alpha > 0$. This again describes a cyclic system. The self-induced exponential drag is enhanced by having a higher current through the system. The empirical observable $A_t = J_t$ is obtained by setting $f = 0$ and $g_{01} = g_{12} = g_{20} = 1 = -g_{10} = -g_{21} = -g_{02}$ in Eq. (6). Even in this case, the inherent rotational symmetry of the system guarantees a uniform stationary distribution over $\mathcal{S}$.

We now study the LDs of the total current observable J_T in (77). Similarly to the previous example, the averaged SIJP current as well as edge currents are constant over time and, therefore,

$$\bar{j} = 3\mathrm{j} = 3((\pi \circ Q(\bar{a}))_{01} - (\pi \circ Q(\bar{a}))_{10}) \approx 0.24. \quad (121)$$

In Fig. 5(b) we compare the accumulated statistics from simulating the process for a long enough time with the SIJP-UKUR rate function bound $I^{\triangle}(j)$ in Eq. (75). Due to the time-independent behavior of the average current through the system, the SIJP-UKUR bound simplifies to

$$I^{\triangle}(j) = \frac{(j-\bar{j})^2}{2\bar{j}^2} \frac{e^{-\alpha\bar{j}}}{1-e^{-\alpha\bar{j}}}((3+\bar{j})(1-e^{-\alpha\bar{j}}) - \bar{j}e^{-\alpha\bar{j}})^2, \quad (122)$$

which is, however, different from and improves the time-homogeneous Markov UKUR [blue dashed line in Fig. 5(b)]. This is because in Eq. (75) a derivative with respect to the self-interaction term appears.

D. Self-induced bistable birth process

We now consider a self-interacting birth process. A birth process is a counting process N_t that increases by one at each birth event, with its dynamics specified by the birth rate r_t. Here this rate depends on the empirical birth activity itself and is given by

$$r_t = K_t - \frac{(K_t - a)(K_t - b)(K_t - c)}{(K_t^2 + p^2)}, \quad (123)$$

for $a < b < c \in \mathbb{R}^+$, $p \in \mathbb{R}$, and

$$K_t = \frac{N_t}{t}, \quad (124)$$

the time-normalized number of births N_t up to time t.

Equivalently, the same counting process can be represented as a two-state SIJP with two equivalent states, 0 and 1, where each jump $0 \leftrightarrow 1$ counts as one birth event and both transition rates are given by Eq. (123). Since the two states are equivalent, the relevant observable is the total number of jumps rather than the occupation of either state. In this representation, Eq. (124) is therefore the empirical total activity of the

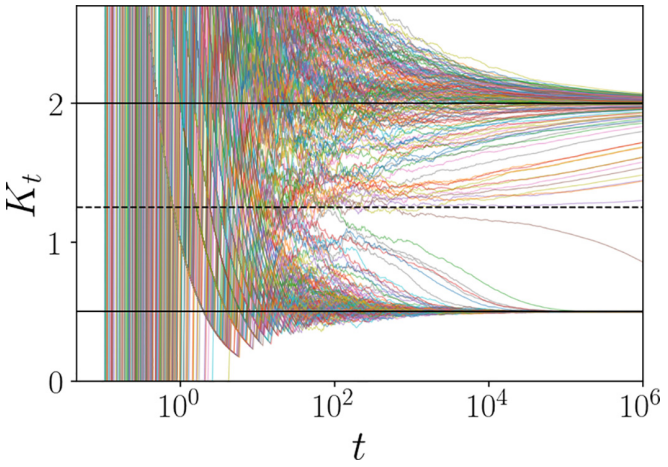


FIG. 6. Empirical activity trajectory simulations for the birth process with self-interacting birth rate in Eq. (123). Trajectories settle at the two stable fixed points $a = 0.5$ and $c = 2$. The unstable fixed point $b = 1.25$ lies between them and dynamically splits trajectories towards the two basins of attraction. Noticeably, trajectories crossing the unstable fixed point settle at a lower "speed" than those starting outside the interval. Simulations for $p = 0.7$.

process, i.e.,

$$K_t = \Phi_{01}(t) + \Phi_{10}(t), \tag{125}$$

where $\Phi_{xy}(t)$ is the empirical flux defined in Eq. (5).

The specific form of the rate r_t guarantees bistable behavior for the observable K_t itself, while the quadratic denominator ensures that the rate remains bounded. This is shown via stochastic approximation by considering the limiting ordinary differential equation (ODE) for the empirical activity. The ODE is

$$\frac{dK_t}{dt} = -\frac{(K_t - a)(K_t - b)(K_t - c)}{t\left(K_t^2 + p^2\right)}. \tag{126}$$

The equation above can also be obtained from the Doob-Meyer decomposition by dropping the martingale contribution, which can be shown to be asymptotically negligible.

Using the logarithmic time variable introduced earlier, $s = \log(t/t_0)$ with $t_0 = 1$, Eq. (126) becomes

$$\frac{dK_s}{ds} = -\frac{(K_s - a)(K_s - b)(K_s - c)}{\left(K_s^2 + p^2\right)}, \tag{127}$$

where, with a slight abuse of notation, we define $K_s := K_{t_0 e^s} = K_{e^s}$. There are three fixed points for this ODE: $k_\infty = a$, b, and c. A simple linear stability analysis reveals that a and c are stable, whereas b is unstable. Numerical simulations of the process are reported in Fig. 6 and confirm this analysis.

To study the large deviations of the empirical activity in Eq. (125), we start from the functional (34). We denote $\rho_\tau := \rho_0(\tau)$. Minimization with respect to ρ_τ, with the normalization constraint, yields the optimal occupation

$$\rho_\tau = \frac{\eta_{01}(\tau)}{\eta_{01}(\tau) + \eta_{10}(\tau)}, \tag{128}$$

where $\eta_{xy}(\tau)$ is the instantaneous flux from x to y in the backward rescaled time. Plugging Eq. (128) back into the functional yields

$$\begin{aligned}\tilde{\mathcal{I}}[\eta] = \int_0^\infty e^{-\tau}\Bigg(&\eta_{01}(\tau)\log\frac{\eta_{01}(\tau)}{\rho_\tau Q(M_\tau)} \\ &+ \eta_{10}(\tau)\log\frac{\eta_{10}(\tau)}{(1-\rho_\tau)Q(M_\tau)} - \eta_{01}(\tau) \\ &- \eta_{10}(\tau) + Q(M_\tau)\Bigg)\,d\tau.\end{aligned} \tag{129}$$

This can be further minimized with respect to $\eta_{01}(\tau)$ and $\eta_{10}(\tau)$, together with the two constraints $\dot{M}_\tau = M_\tau - \eta_{01}(\tau) - \eta_{10}(\tau)$ and

$$k = \int_0^\infty e^{-\tau}(\eta_{01}(\tau) + \eta_{10}(\tau))d\tau, \tag{130}$$

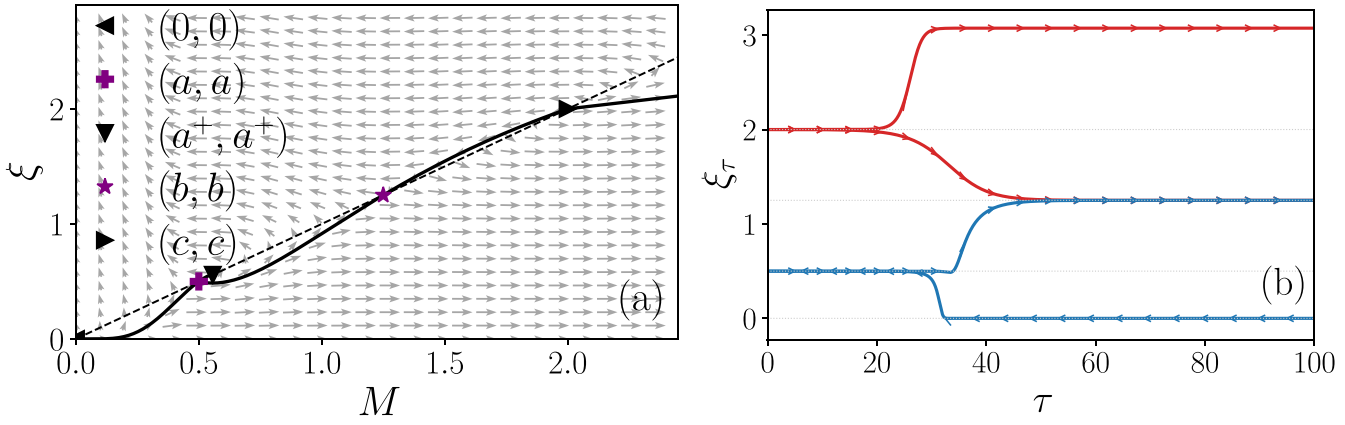


FIG. 7. (a) Phase portrait of the optimal trajectories (ξ_τ, M_τ) solving Eqs. (131) and (132) for the self-interacting birth process with $a = 0.5$, $b = 1.25$, $c = 2$, and $p = 0.7$. Fixed points of the 2D dynamics are marked with black and purple symbols. The black solid line denotes the one-dimensional stable manifold connecting the saddle points (a, a) and (b, b) with the other unstable points (0,0), (a^+, a^+), and (c, c), along which all admissible optimal trajectories lie. Arrows indicate the direction of increasing reversed time τ. (b) Optimal trajectories (ξ_τ, M_τ) along the stable manifold. Trajectories are obtained by integrating Eqs. (131) and(132): The red trajectories connect to the fixed point (c, c), while the blue trajectory connects to (a, a). Each point on the curves corresponds to a distinct initial value $M_0 = k$, generating a fluctuation of the empirical total activity $K_T = k$ in the physical long-time limit.

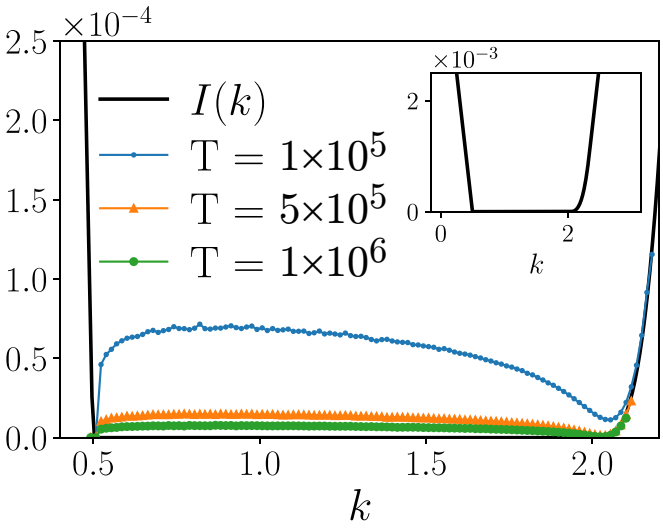


FIG. 8. Large deviation rate function $I(k)$ of the total activity K_T in Eq. (125) for the self-interacting birth process. The solid black line denotes the exact rate function obtained by minimizing $\tilde{\mathcal{I}}[\rho, \eta]$ under the activity constraint. Symbols correspond to Monte Carlo simulations.

which represents the long-time atypical value of the empirical activity in the backward rescaled time frame.

This yields the following two-dimensional system of ODEs:

$$\dot{M}_\tau = M_\tau - \xi_\tau \tag{131}$$

$$\dot{\xi}_\tau = \xi_\tau Q'(M_\tau)\left[\frac{M_\tau - Q(M_\tau)}{Q(M_\tau)}\right], \tag{132}$$

where $\xi_\tau = \eta_{01}(\tau) + \eta_{10}(\tau)$ is the total instantaneous flux. The system must be solved with the initial condition $M_0 = k$ and the terminal condition $\lim_{\tau\to\infty} e^{-\tau} M_\tau = 0$. Both conditions follow from Eq. (130) using Eq. (131). We plot the phase portrait and the trajectory solutions of the ODE system in Figs. 7(a) and 7(b).

Finally, in Fig. 8 we plot the rate function of the observable total activity K_t in Eq. (125). The set of zeros of the rate function is not finite and includes all points between $k_\infty = a$ and $k_\infty = c$. Simulations confirm this picture, suggesting that the speed of large deviations may differ for trajectories settling into one of the two stable fixed points from the inner side of the interval compared with trajectories approaching from outside, as Fig. 6 initially suggested.

VIII. CONCLUSION

In this work we have developed a general framework to characterize fluctuations in SIJPs, a broad class of non-Markovian stochastic dynamics in which transition rates depend on empirical observables of the process itself. Using an exponential tilting construction, we derived the level-2.5 large deviation principle governing the joint fluctuations of the empirical occupation measure and empirical flux for SIJPs with general functional dependence on these quantities. The resulting rate functional reveals a clear separation between fast microscopic dynamics and slow, memory-driven evolution, encoded through an exponential temporal discount. Building on this structure, we obtained variational bounds on flux and current fluctuations, leading to kinetic and thermodynamic uncertainty relations that extend classical Markovian results to non-Markovian settings. Through low-dimensional examples, we demonstrated how self-interaction and time-dependent feedback modify fluctuation behavior and assessed the tightness of the derived bounds against exact or numerical large deviation results.

Several open questions emerge from this analysis. From a theoretical perspective, the level-2.5 rate function derived here relies on assumptions underpinning the tilting procedure and timescale separation, which should be examined on a case-by-case basis, particularly for strongly nonsmooth feedbacks. While rigorous large deviation results are available for specific classes of SIJPs, notably those with affine dependence on the empirical measure, extending such results to more general rate maps remains an open challenge. In addition, the possible nonconvexity of the rate function after contraction deserves further investigation, as it may signal richer fluctuation mechanisms such as multiple optimal trajectories or dynamical phase transitions. On the level of uncertainty relations, it remains to be understood under which conditions kinetic and thermodynamic bounds are tight, and whether systematically sharper bounds can be obtained by exploiting time-dependent or state-dependent variational Ansätze.

Looking ahead, the framework developed here opens several directions for future research. Extending rigorous large deviation techniques to broader classes of self-interacting dynamics would strengthen the mathematical foundations of fluctuation theory for non-Markovian systems. Another important direction is the development of a macroscopic fluctuation theory analog of the present construction, where the empirical observables entering the rates become spatially extended fields rather than finite-dimensional quantities. The variational structure of the level-2.5 rate function naturally connects with stochastic control formulations, suggesting alternative approaches to the characterization of rare events. From a physical standpoint, the results contribute to ongoing efforts to generalize stochastic thermodynamics beyond Markovian dynamics, complementing existing work on semi-Markov processes, delayed systems, and coarse-grained descriptions. It would also be interesting to extend the definition of the memory observable by introducing kernels that damp the contribution of the distant past, reset memory, or restrict it to a finite time window. Such extensions would be relevant, for example, for biophysical systems in which the memorized observable is an environmental trace with finite lifetime, such as evaporating chemical trails in autochemotactic motion. More broadly, SIJPs provide a controlled setting to investigate how memory, feedback, and information encoded in empirical observables shape fluctuations, irreversibility, and precision-dissipation trade-offs, with potential relevance for adaptive biological systems, active matter, and learning-driven stochastic processes.

ACKNOWLEDGMENTS

The authors warmly thank Amarjit Budhiraja for many insightful and stimulating discussions. F.C. is grateful to Tobias

Grafke for help on clarifying numerical aspects behind the example in Sec. VII A. F.C. is supported by a Leverhulme Early Career Fellowship (ECF-2025-482). J.P.G. is supported by EPSRC Grants No. EP/V031201/1 and No. EP/T022140/1.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: DETAILS ON SIJP ASSUMPTIONS AND COUNTEREXAMPLES

In this appendix, we clarify the role of the fixed-point argument. In the simpler case $g_{xy} = 0$, the observable is bounded by construction and any limiting value of A_t belongs to the compact interval

$$\mathcal{I}_f = \Big[\min_{x\in\mathcal{S}} f_x, \max_{x\in\mathcal{S}} f_x\Big]. \tag{A1}$$

The frozen average

$$\Psi_f(a) = \sum_{x\in\mathcal{S}} f_x \gamma_x(a) \tag{A2}$$

is continuous and maps $\mathcal{I}_f$ into itself. Hence Brouwer's fixed-point theorem guarantees the existence of at least one $a^\star \in \mathcal{I}_f$ such that

$$a^\star = \Psi_f(a^\star), \quad \gamma(a^\star)Q(a^\star) = 0. \tag{A3}$$

This proves the existence of a self-consistent frozen stationary point but not convergence of the SIJP towards it. Convergence requires additional stability assumptions on the limiting dynamics. In the scalar case, this corresponds to stability of the fixed points of the limiting ODE,

$$\dot{a} = \Psi_f(a) - a. \tag{A4}$$

When $g \neq 0$, boundedness of f and g alone is not sufficient to produce a compact set for A_t. We further require that the empirical jump fluxes $\Phi_{xy}(t)$ are bounded and this can be obtained by imposing a uniform bound on the jump rates on the relevant range of a. Under such an assumption, one can formulate the self-consistency problem for stationary pairs as

$$\bar{a} = \sum_{x\in\mathcal{S}} f_x \pi_x + \sum_{(x,y)\in\mathcal{E}} g_{xy}\varphi_{xy}, \tag{A5}$$

$$\pi Q(\bar{a}) = 0, \tag{A6}$$

$$\sum_{y\neq x} \varphi_{xy} = \sum_{y\neq x} \varphi_{yx}. \tag{A7}$$

The existence of such a pair is then part of the standing assumptions defining $\mathcal{A}_{\text{stat}}$. There are simple ways in which these assumptions can fail. First, imagine a feedback that dynamically removes part of the state space. Take $\mathcal{S} = \{0, 1\}$, $g = 0$, $f_0 = 0$, $f_1 = 1$, so that

$$A_t = L_1(t) \in [0, 1]. \tag{A8}$$

Then, let

$$Q_{01}(a) = \begin{cases} 1, & a \geqslant 1/2, \\ 0, & a < 1/2, \end{cases} \quad Q_{10}(a) = 1. \tag{A9}$$

This is a discontinuous map that induces a breaking of irreducibility: When $A_t \geqslant 1/2$, both states communicate and the process is irreducible, however, as soon as $A_t < 1/2$, the transition $0 \to 1$ is switched off. State 0 then becomes absorbing, and the effective state space collapses from $\{0, 1\}$ to the single stae $\{0\}$. Consequently, once the trajectory enters this regime and reaches state 0, it remains there, and

$$L_0(t) \to 1 \quad L_1(t) \to 0, \tag{A10}$$

$$A_t \to 0 \quad \Phi_{01}(t), \quad \Phi_{10}(t) \to 0. \tag{A11}$$

Physically, this corresponds to a memory-induced collapse into an inactive state. The long-time behavior is therefore selected by a reducible limiting dynamics, not by an irreducible frozen generator. Another interesting example not covered by our theory is obtained from a specific self-exciting activity feedback. Take $\mathcal{S} = \{0, 1\}$, $f_0 = f_1 = 0$, $g_{01} = g_{10} = 1$, so that

$$A_t = \Phi_{01}(t) + \Phi_{10}(t) = \frac{N_t}{t}, \tag{A12}$$

where N_t is the total number of jumps up to time t. Let

$$Q_{01}(a) = Q_{10}(a) = 1 + a, \quad a \geqslant 0. \tag{A13}$$

For every finite a, the frozen chain is irreducible. However, independently of the current state of the process, the instantaneous escape rate is $1 + A_t$. Since $A_t = N_t/t$, its deterministic drift is

$$\frac{dA_t}{dt} = \frac{1}{t}. \tag{A14}$$

Thus A_t grows as $\log t$ rather than converging to a finite stationary value. Therefore no stationary limiting pair (π, φ) of the type assumed above exists.

APPENDIX B: CONTROL OF THE ERROR DUE TO TIMESCALE SEPARATION AND BLOCK-AVERAGING CALCULATION

We detail here why the error term in Eq. (27) is subextensive in the long-time limit. We denote

$$F(t) = \sum_{(x,y)\in\mathcal{E}} \Big[\tilde{\rho}_x(t)(Q_{xy}(A_t) - \tilde{H}_{xy}(t)) + \tilde{\rho}_x(t)\tilde{H}_{xy}(t)\log\frac{\tilde{H}_{xy}(t)}{Q_{xy}(A_t)}\Big]. \tag{B1}$$

The error $\text{Err}_T(\Delta t)$ has two contributions. The first comes from freezing the time-dependent cost density $F(t)$ over each block, viz. it comes from replacing an integral with a sum. The second comes from replacing empirical occupations and fluxes within a block by the frozen accompanying distribution. We first consider the freezing error. On the block $[t_k, t_k + \Delta t]$, this error is

$$E_k^{\text{freeze}} = \int_{t_k}^{t_k+\Delta t} F(s)\,ds - F(t_k)\Delta t = \int_{t_k}^{t_k+\Delta t} [F(s) - F(t_k)]\,ds. \tag{B2}$$

The timescale separation assumption is equivalently expressed by requiring the frozen cost density to vary on the slow SIJP timescale,

$$\left|\frac{dF}{dt}(t)\right| = O(t^{-1}). \tag{B3}$$

Therefore, for $s \in [t_k, t_k + \Delta t]$,

$$|F(s) - F(t_k)| \leqslant C \int_{t_k}^{s} \frac{du}{u} = C \log \frac{s}{t_k} = C \log\left(1 + \frac{s - t_k}{t_k}\right). \tag{B4}$$

Since $\Delta t \ll t_k$ on the slow scale, this implies

$$\log\left(1 + \frac{s - t_k}{t_k}\right) \leqslant \frac{s - t_k}{t_k}. \tag{B5}$$

Hence

$$\left|E_k^{\text{freeze}}\right| \leqslant \frac{C}{t_k} \int_{t_k}^{t_k + \Delta t} (s - t_k)\, ds = O\left(\frac{(\Delta t)^2}{t_k}\right). \tag{B6}$$

Summing over blocks gives

$$\sum_k \left|E_k^{\text{freeze}}\right| = O\left(\sum_k \frac{(\Delta t)^2}{t_k}\right). \tag{B7}$$

Using $t_k = k\Delta t$, up to an initial finite-time block which gives only a subextensive contribution,

$$\sum_k \frac{(\Delta t)^2}{t_k} = \Delta t \sum_k \frac{1}{k} = O(\Delta t \log T). \tag{B8}$$

Thus

$$E_T^{\text{freeze}} = O(\Delta t \log T). \tag{B9}$$

We now consider the second source of error. Within each block, the empirical occupations and fluxes are replaced by the frozen accompanying distribution, as in Eqs. (24) and (25). This replacement is valid after the frozen dynamics has relaxed. If τ_μ denotes a uniform upper bound on the mixing times of the frozen generators, then the nonstationary transient contributes an error of order τ_μ per block. Since the number of blocks is of order $T/\Delta t$, the total mixing error is

$$E_T^{\text{mix}} = O\left(\frac{T}{\Delta t}\tau_\mu\right). \tag{B10}$$

Combining the two estimates, we obtain

$$\frac{|\text{Err}_T(\Delta t)|}{T} = O\left(\frac{\Delta t \log T}{T} + \frac{\tau_\mu}{\Delta t}\right). \tag{B11}$$

Therefore, choosing for instance $\Delta t = T^\alpha$ with $0 < \alpha < 1$, while keeping $\tau_\mu = O(1)$, gives

$$\frac{|\text{Err}_T(\Delta t)|}{T} \to 0, \tag{B12}$$

as $T \to \infty$. Thus the accumulated block error is subextensive and does not contribute to the large-deviation rate function.

APPENDIX C: SIMULATING A SIJP

We now detail how to run simulations of SIJPs. Naively applying a Gillespie algorithm does not work as the rate matrix $Q(A_t)$ of a SIJP changes with time due to the fact that the empirical observable (1) is explicitly time dependent.

The survival probability, viz. the probability to be at, say, state $x \in \mathcal{S}$ for a time longer than u, having landed onto x at time t, is therefore a time-dependent quantity and takes the form

$$\exp\left(\int_t^{t+u} Q_{xx}(A_{t'})\, dt'\right), \tag{C1}$$

where the diagonal term $Q_{xx}(A_{t'})$ is defined in Eq. (3). Equivalently, introducing the instantaneous escape rate

$$\lambda_x(t') := -Q_{xx}(A_{t'}) = \sum_{y \neq x} Q_{xy}(A_{t'}), \tag{C2}$$

we define the corresponding integrated escape rate as

$$\Lambda_x(t, t+u) := \int_t^{t+u} \lambda_x(t')\, dt' = -\int_t^{t+u} Q_{xx}(A_{t'})\, dt', \tag{C3}$$

as in Eq. (13).

A trajectory of length T for such a SIJP can then be realized in two equivalent ways. First, we can implement an accept-reject algorithm, which is based on the (Ogata) thinning construction, see also Ex. 2.1.6 in Ref. [97]. Imagine the process sits at state x at time t. The instantaneous escape rate can be bounded on the time horizon of interest by a constant $\bar{\lambda}_x$ such that

$$\bar{\lambda}_x \geqslant \lambda_x(t') = -Q_{xx}(A_{t'}) \tag{C4}$$

for all relevant $t' \in [t, T]$. Such a bound can be chosen, for instance, from an *a priori* bound on the admissible values of $A_{t'}$, or updated adaptively after each rejected proposal. Then, it is enough to (i) propose exponential waiting times u with rate $\bar{\lambda}_x$ and (ii) accept them with probability $\lambda_x(t+u)/\bar{\lambda}_x$. If accepted, (iii) then we update time, empirical observable, and choose next state $y \in \mathcal{S}$ with probability

$$\frac{Q_{xy}(A_{t+u})}{\lambda_x(t+u)}. \tag{C5}$$

Otherwise, (iii) we shift time by u, i.e., $t \leftarrow t + u$, keep the process at x, update the empirical observable, compute a new bound $\bar{\lambda}_x$ and loop back to (i), repeating until $t + u$ exceeds the final observation time T.

A second method that can be used to simulate SIJPs is based on direct numerical integration. At time t, while the process sits at $x \in \mathcal{S}$, (i) we extract $E \sim \text{Exp}(1)$, namely a unit-rate exponential random variable, and (ii) numerically solve $\Lambda_x(t, t+u) = E$ for $u \geqslant 0$ using, e.g., bisection methods. Then, (iii) we advance to time $t \leftarrow t + u$, update the empirical observable, and choose next state $y \in \mathcal{S}$ with probability (C5). We loop back to (i) and iterate until $t + u$ exceeds T.

The first method is cheaper in terms of computational resources, but it might be affected by a low acceptance rate, which would increase the overall simulation time. The second method is heavier than the first from the computational

point of view as quadrature techniques need to be included. However, it is not affected by an accept-reject step. In all the examples discussed in this work we made use of the second algorithm.

We conclude by remarking that the construction above and the algorithms mentioned are not restricted to SIJPs. They can be applied to—in fact, they were developed to work with—time-inhomogeneous Markov processes. In our SIJP, the time nonhomogeneity is implicit and is evaluated through the empirical observable. We refer the reader to Chap. 2 of Ref. [97] for further details on how to treat time-inhomogeneous Poisson processes.

[1] F. Coghi, A. Budhiraja, and J. P. Garrahan, companion paper, Level 2.5 large deviations and uncertainty relations for non-Markov self-interacting dynamics, Phys. Rev. Lett. **137**, 197101 (2026).

[2] E. O. Budrene and H. C. Berg, Dynamics of formation of symmetrical patterns by chemotactic bacteria, Nature (London) **376**, 49 (1995).

[3] M. P. Brenner, L. S. Levitov, and E. O. Budrene, Physical mechanisms for chemotactic pattern formation by bacteria, Biophys. J. **74**, 1677 (1998).

[4] D. J. Sumpter and M. Beekman, From nonlinearity to optimality: Pheromone trail foraging by ants, Anim. Behav. **66**, 273 (2003).

[5] D. E. Jackson, S. J. Martin, M. Holcombe, and F. L. W. Ratnieks, Longevity and detection of persistent foraging trails in Pharaoh's ants, Monomorium pharaonis (L.), Anim. Behav. **71**, 351 (2006).

[6] K. Zhao, B. S. Tseng, B. Beckerman, F. Jin, M. L. Gibiansky, J. J. Harrison, E. Luijten, M. R. Parsek, and G. C. Wong, Psl trails guide exploration and microcolony formation in Pseudomonas aeruginosa biofilms, Nature (London) **497**, 388 (2013).

[7] A. Gelimson, K. Zhao, C. K. Lee, W. T. Kranz, G. C. L. Wong, and R. Golestanian, Multicellular self-organization of *P. aeruginosa* due to interactions with secreted trails, Phys. Rev. Lett. **117**, 178102 (2016).

[8] C. R. Reid, T. Latty, A. Dussutour, and M. Beekman, Slime mold uses an externalized spatial "memory" to navigate in complex environments, Proc. Natl. Acad. Sci. USA **109**, 17490 (2012).

[9] J. R. Howse, R. A. L. Jones, A. J. Ryan, T. Gough, R. Vafabakhsh, and R. Golestanian, Self-motile colloidal particles: From directed propulsion to random walk, Phys. Rev. Lett. **99**, 048102 (2007).

[10] S. Thutupalli, R. Seemann, and S. Herminghaus, Swarming behavior of simple model squirmers, New J. Phys. **13**, 073021 (2011).

[11] B. V. Hokmabad, J. Agudo-Canalejo, S. Saha, R. Golestanian, and C. C. Maass, Chemotactic self-caging in active emulsions, Proc. Natl. Acad. Sci. USA **119**, e2122269119 (2022).

[12] B. Nakayama, H. Nagase, H. Takahashi, Y. Saito, S. Hatayama, K. Makino, E. Yamamoto, and T. Saiki, Tunable pheromone interactions among microswimmers, Proc. Natl. Acad. Sci. USA **120**, e2213713120 (2023).

[13] M. Kumar, A. Murali, A. G. Subramaniam, R. Singh, and S. Thutupalli, Emergent dynamics due to chemo-hydrodynamic self-interactions in active polymers, Nat. Commun. **15**, 4903 (2024).

[14] D. Coppersmith and P. Diaconis, Random walks with reinforcement (unpublished).

[15] B. Tóth, Self-interacting random motions, in *European Congress of Mathematics*, Vol. 1 (Birkhäuser, Basel, 2001), pp. 555–564.

[16] A. Stevens and H. G. Othmer, Aggregation, blowup, and collapse: The ABC's of taxis in reinforced random walks, SIAM J. Appl. Math. **57**, 1044 (2006).

[17] R. Pemantle, A survey of random processes with reinforcement, Probab. Surv. **4**, 1 (2007).

[18] A. Erschler, B. Tóth, and W. Werner, Stuck walks, Probab. Theory Relat. Fields **154**, 149 (2012).

[19] D. Kious, Stuck walks: A conjecture of Erschler, Tóth and Werner, Ann. Probab. **44**, 883 (2016).

[20] A. Barbier-Chebbah, O. Bénichou, and R. Voituriez, Self-interacting random walks: Aging, exploration, and first-passage times, Phys. Rev. X **12**, 011052 (2022).

[21] J. Brémont, O. Bénichou, and R. Voituriez, Exact propagators of one-dimensional self-interacting random walks, Phys. Rev. Lett. **133**, 157101 (2024).

[22] E. F. Keller and L. A. Segel, Initiation of slime mold aggregation viewed as an instability, J. Theor. Biol. **26**, 399 (1970).

[23] Y. Tsori and P.-G. de Gennes, Self-trapping of a single bacterium in its own chemoattractant, Europhys. Lett. **66**, 599 (2004).

[24] A. Gelimson and R. Golestanian, Collective dynamics of dividing chemotactic cells, Phys. Rev. Lett. **114**, 028101 (2015).

[25] T. Grafke, M. E. Cates, and E. Vanden-Eijnden, Spatiotemporal self-organization of fluctuating bacterial colonies, Phys. Rev. Lett. **119**, 188003 (2017).

[26] R. Grima, Strong-coupling dynamics of a multicellular chemotactic system, Phys. Rev. Lett. **95**, 128103 (2005).

[27] A. Sengupta, S. van Teeffelen, and H. Löwen, Dynamics of a microorganism moving by chemotaxis in its own secretion, Phys. Rev. E **80**, 031122 (2009).

[28] O. Pohl and H. Stark, Dynamic clustering and chemotactic collapse of self-phoretic active particles, Phys. Rev. Lett. **112**, 238303 (2014).

[29] W. T. Kranz, A. Gelimson, K. Zhao, G. C. L. Wong, and R. Golestanian, Effective dynamics of microorganisms that interact with their own trail, Phys. Rev. Lett. **117**, 038101 (2016).

[30] P. Del Moral and L. Miclo, Self-interacting Markov chains, Stochast. Anal. Appl. **24**, 615 (2006).

[31] D. Kannan and J. Zhang, Self-interacting Markov chains: Some asymptotics, Stochast. Anal. Appl. **27**, 196 (2009).

[32] F. Coghi and J. P. Garrahan, Self-interacting processes via Doob conditioning, J. Stat. Mech. (2025) 103202.

[33] A. Budhiraja and F. Coghi, Jump processes with self-interactions: Large deviation asymptotics, arXiv:2510.14659.

[34] M. Benaïm, M. Ledoux, and O. Raimond, Self-interacting diffusions, Probab. Theory Relat. Fields **122**, 1 (2002).

[35] M. Benaïm, Dynamics of Stochastic Approximation Algorithms, Technical Report Seminar on Probability, Strasbourg, 1999.
[36] A. Kurtzmann, The ODE method for some self-interacting diffusions on $\{R\}^d$, Ann. Inst. H. Poincaré Probab. Statist. **46**, 618 (2010).
[37] M. Benaïm and O. Raimond, Self-interacting diffusions IV: Rate of convergence, Electron. J. Probab. **16**, 1815 (2011).
[38] F. Coghi, Current fluctuations of a self-interacting diffusion on a ring, J. Phys. A: Math. Theor. **58**, 015002 (2025).
[39] S. Franchini, Large deviations for generalized Polya urns with arbitrary urn function, Stochast. Process. Appl. **127**, 3372 (2017).
[40] A. Budhiraja and A. Waterbury, Empirical measure large deviations for reinforced chains on finite spaces, Syst. Control Lett. **169**, 105379 (2022).
[41] A. Budhiraja, A. Waterbury, and P. Zouboulouglou, Large deviations for empirical measures of self-interacting Markov chains, Stochast. Process. Appl. **186**, 104640 (2025).
[42] F. Coghi, R. Duvezin, and J. S. Wettlaufer, Accelerated first-passage dynamics in a non-Markovian feedback Ornstein–Uhlenbeck process, J. Stat. Phys. **192**, 128 (2025).
[43] M. Benaïm and O. Raimond, Self-interacting diffusions. III. Symmetric interactions, Ann. Probab. **33**, 1716 (2005).
[44] A. Aleksian, P. Del Moral, A. Kurtzmann, and J. Tugaut, Self-interacting diffusions: Long-time behaviour and exit-problem in the uniformly convex case, ESAIM Probab. Stat. **28**, 46 (2024).
[45] C. Maes and K. Netočný, Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states, Europhys. Lett. **82**, 30003 (2008).
[46] L. Bertini, A. Faggionato, and D. Gabrielli, Large deviations of the empirical flow for continuous time Markov chains, Ann. Inst. H. Poincaré Probab. Statist. **51**, 867 (2012).
[47] A. C. Barato and U. Seifert, Thermodynamic uncertainty relation for biomolecular processes, Phys. Rev. Lett. **114**, 158101 (2015).
[48] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Dissipation bounds all steady-state current fluctuations, Phys. Rev. Lett. **116**, 120601 (2016).
[49] J. P. Garrahan, Simple bounds on fluctuations and uncertainty relations for first-passage times of counting observables, Phys. Rev. E **95**, 032134 (2017).
[50] I. Di Terlizzi and M. Baiesi, Kinetic uncertainty relation, J. Phys. A: Math. Theor. **52**, 02LT03 (2019).
[51] J. M. Horowitz and T. R. Gingrich, Thermodynamic uncertainty relations constrain non-equilibrium fluctuations, Nat. Phys. **16**, 15 (2020).
[52] K. Macieszczak, Ultimate kinetic uncertainty relation and optimal performance of stochastic clocks, arXiv:2407.09839.
[53] K. Prech, G. T. Landi, F. Meier, N. Nurgalieva, P. P. Potts, R. Silva, and M. T. Mitchison, Optimal time estimation and the clock uncertainty relation for stochastic processes, Phys. Rev. X **15**, 031068 (2025).
[54] V. Lecomte, C. Appert-Rolland, and F. van Wijland, Thermodynamic formalism for systems with Markov dynamics, J. Stat. Phys. **127**, 51 (2007).
[55] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Dynamical first-order phase transition in kinetically constrained models of glasses, Phys. Rev. Lett. **98**, 195702 (2007).
[56] H. Touchette, The large deviation approach to statistical mechanics, Phys. Rep. **478**, 1 (2009).
[57] R. Chetrite and H. Touchette, Nonequilibrium Markov processes conditioned on large deviations, Ann. Henri Poincaré **16**, 2005 (2015).
[58] R. Chetrite and H. Touchette, Variational and optimal control representations of conditioned and driven processes, J. Stat. Mech. (2015) P12001.
[59] J. P. Garrahan, Aspects of non-equilibrium in classical and quantum systems: Slow relaxation and glasses, dynamical large deviations, quantum non-ergodicity, and open quantum dynamics, Physica A **504**, 130 (2018).
[60] A. Barato, R. Chetrite, A. Faggionato, and D. Gabrielli, A unifying picture of generalized thermodynamic uncertainty relations, J. Stat. Mech. (2019) 084017.
[61] R. L. Jack, Ergodicity and large deviations in physical systems with stochastic dynamics, Eur. Phys. J. B **93**, 74 (2020).
[62] G. Carugno, P. Vivo, and F. Coghi, Graph-combinatorial approach for large deviations of Markov chains, J. Phys. A **55**, 295001 (2022).
[63] G. M. Schütz and S. Trimper, Elephants can always remember: Exact long-range memory effects in a non-Markovian random walk, Phys. Rev. E **70**, 045101(R) (2004).
[64] A. Rebenshtok and E. Barkai, Distribution of time-averaged observables for weak ergodicity breaking, Phys. Rev. Lett. **99**, 210601 (2007).
[65] R. J. Harris, Random walkers with extreme value memory: Modelling the peak-end rule, New J. Phys. **17**, 053049 (2015).
[66] R. L. Jack and R. J. Harris, Giant leaps and long excursions: Fluctuation mechanisms in systems with long-range memory, Phys. Rev. E **102**, 012154 (2020).
[67] R. J. Harris and H. Touchette, Current fluctuations in stochastic systems with long-range memory, J. Phys. A **42**, 342001 (2009).
[68] S. Franchini, Corrigendum to "Large deviations for generalized Polya urns with arbitrary urn function" [Stochastic processes and their applications 127 (2017) 3372–3411], Stochast. Process. Appl. **189**, 104745 (2025).
[69] A. C. Barato and R. Chetrite, Current fluctuations in periodically driven systems, J. Stat. Mech. (2018) 053207.
[70] C. Maes, K. Netočný, and B. Wynants, Dynamical fluctuations for semi-Markov processes, J. Phys. A **42**, 365002 (2009).
[71] J. P. Garrahan and I. Lesanovsky, Thermodynamics of quantum jump trajectories, Phys. Rev. Lett. **104**, 160601 (2010).
[72] M. Shreshtha and R. J. Harris, Thermodynamic uncertainty for run-and-tumble–type processes, Europhys. Lett. **126**, 40007 (2019).
[73] F. Carollo, R. L. Jack, and J. P. Garrahan, Unraveling the large deviation statistics of Markovian open quantum systems, Phys. Rev. Lett. **122**, 130605 (2019).
[74] B. Ertel, J. van der Meer, and U. Seifert, Operationally accessible uncertainty relations for thermodynamically consistent semi-Markov processes, Phys. Rev. E **105**, 044113 (2022).
[75] C. Jia, D.-Q. Jiang, and B. Wu, Large deviations for the empirical measure and empirical flow of Markov renewal processes with a countable state space, Electron. J. Probab. **29**, 1 (2024).

[76] F. Liu, S. Xia, and S. Su, Semi-Markov processes in open quantum systems. III. Large deviations of first-passage-time statistics, Phys. Rev. E **110**, 064145 (2024).

[77] A. M. Maier, J. H. Fritz, and U. Seifert, Pedestrian's approach to large deviations in semi-Markov processes with an application to entropy production, Phys. Rev. E **113**, 014119 (2026).

[78] C. Maes, S. Safaverdi, P. Visco, and F. van Wijland, Fluctuation-response relations for nonequilibrium diffusions with memory, Phys. Rev. E **87**, 022125 (2013).

[79] H. Mori, Transport, collective motion, and Brownian motion, Prog. Theor. Phys. **33**, 423 (1965).

[80] R. Zwanzig, Nonlinear generalized Langevin equations, J. Stat. Phys. **9**, 215 (1973).

[81] C. Hijón, P. Español, E. Vanden-Eijnden, and R. Delgado-Buscalioni, Mori–Swanzig formalism as a practical computational tool, Faraday Discuss. **144**, 301 (2009).

[82] T. Schilling, Coarse-grained modelling out of equilibrium, Phys. Rep. **972**, 1 (2022).

[83] G. Jung, Non-Markovian systems out of equilibrium: Exact results for two routes of coarse graining, J. Phys.: Condens. Matter **34**, 204004 (2022).

[84] Y. Zhu, H. Lei, and C. Kim, General validity of the second fluctuation-dissipation theorem in the nonequilibrium steady state: Theory and applications, Phys. Scr. **98**, 115402 (2023).

[85] C. Maes, On the second fluctuation–dissipation theorem for nonequilibrium baths, J. Stat. Phys. **154**, 705 (2014).

[86] A. Chen, H. Qiu, T. Tian, and T. Zhou, Generalized fluctuation-dissipation theorem for non-Markovian reaction networks, Phys. Rev. E **105**, 064409 (2022).

[87] T.-L. Vu and J. Barés, Soft-grain compression: Beyond the jamming point, Phys. Rev. E **100**, 042907 (2019).

[88] Y. Hasegawa and T. Van Vu, Fluctuation theorem uncertainty relation, Phys. Rev. Lett. **123**, 110602 (2019).

[89] M. L. Rosinberg and G. Tarjus, Comment on thermodynamic uncertainty relation for time-delayed Langevin systems, arXiv:1810.12467.

[90] M. K. Chari, On reversible semi-Markov processes, Oper. Res. Lett. **15**, 157 (1994).

[91] L. Caprini, U. M. B. Marconi, A. Puglisi, and A. Vulpiani, The entropy production of Ornstein–Uhlenbeck active particles: a path integral method for correlations, J. Stat. Mech. (2019) 053203.

[92] I. Di Terlizzi and M. Baiesi, A thermodynamic uncertainty relation for a system with memory, J. Phys. A: Math. Theor. **53**, 474002 (2020).

[93] B. K. Agarwalla and D. Segal, Assessing the validity of the thermodynamic uncertainty relation in quantum systems, Phys. Rev. B **98**, 155438 (2018).

[94] A. de La Fortelle, Large deviation principle for Markov chains in continuous time, Probl. Inf. Trans. **37**, 120 (2001).

[95] A. C. Barato and R. Chetrite, A formal view on level 2.5 large deviations and fluctuation relations, J. Stat. Phys. **160**, 1154 (2015).

[96] S. Kusuoka, K. Kuwada, and Y. Tamura, Large deviation for stochastic line integrals as L^p -currents, Probab. Theory Relat. Fields **147**, 649 (2010).

[97] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes, Vol. I: Elementary Theory and Methods*, 2nd ed. (Springer, New York, 2003).

[98] G. Torrie and J. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, J. Comput. Phys. **23**, 187 (1977).

[99] J. A. Bucklew, *Introduction to Rare Event Simulation* (Springer, New York, 2004).

[100] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).

[101] U. Ray, G. K.-L. Chan, and D. T. Limmer, Importance sampling large deviations in nonequilibrium steady states. I, J. Chem. Phys. **148**, 124120 (2018).

[102] K. Klymko, P. L. Geissler, J. P. Garrahan, and S. Whitelam, Rare behavior of growth processes via umbrella sampling of trajectories, Phys. Rev. E **97**, 032123 (2018).

[103] A. Guyader and H. Touchette, Efficient large deviation estimation based on importance sampling, J. Stat. Phys. **181**, 551 (2020).

[104] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 3 ed. (Academic Press, 2023).

[105] M. D. Donsker and S. R. S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, II, Comm. Pure. Appl. Math. **28**, 279 (1975).

[106] C. Jarzynski, Nonequilibrium equality for free energy differences, Phys. Rev. Lett. **78**, 2690 (1997).

[107] U. Seifert, Stochastic thermodynamics: From principles to the cost of precision, Physica A **504**, 176 (2018).