Level 2.5 Large Deviations and Uncertainty Relations for Non-Markov Self-Interacting Dynamics

Francesco Coghi,[1,2,*] Amarjit Budhiraja,[3] and Juan P. Garrahan[1,2]
[1]School of Physics and Astronomy, University of Nottingham, Nottingham, NG7 2RD, United Kingdom
[2]Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, Nottingham, NG7 2RD, United Kingdom
[3]Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, North Carolina 27599, USA

(Received 25 March 2026; accepted 20 October 2026; published 6 November 2026)

We address the general problem of formulating the dynamical large deviations of non-Markovian systems in a closed form. Specifically, we consider a broad class of "self-interacting" jump processes whose dynamics depends on the past through a functional of a state-dependent empirical observable. Exploiting a natural separation of timescales, and under the corresponding mixing conditions, we obtain the exact (so-called "level 2.5") large deviation joint statistics of the empirical measure over configurations and of the empirical flux of transitions. As an application of this general framework, we derive explicit general bounds on the fluctuations of trajectory observables, generalizing to the non-Markovian case both thermodynamic and kinetic uncertainty relations. We illustrate our theory with simple examples, and discuss potential applications of these results.

DOI: 10.1103/pt64-m494

Introduction—Communicating information related to past experiences is, arguably, one of the most peculiar features of intelligent beings. Millions of years of evolution have led small organisms, such as *E. coli*, *P. aeruginosa,* and *pharaoh ants*, to develop *external spatial memories*: as these organisms navigate space, they create a memory in the surrounding environment by dropping chemicals [1–6] or by inducing structural changes [7]. This memory then serves other microorganisms of the same *autochemotactic* species to inform future behavior. Such natural behaviors can also inform strategies [8] used to build active self-propelling and interacting artificial agents [9–11].

The typical behavior of autochemotactic organisms can be explained using non-Markov stochastic systems, such as field-based continuum models [12–15], also in combination with stochastic dynamics [16–19] and reinforced random processes [20–26]. Self-interacting processes (SIPs)—including discrete-time chains [27–29], continuous-time jump processes [30] and diffusions [31,32]—describe dynamics mediated by a functional of the "empirical occupation measure," a stochastic field that models changes generated by autochemotactic agents in the environment. Such self-interaction generates long-time memory and strong correlations with the past, leading to nontrivial dynamics, including ergodicity breaking [31,33,34], and modified (potentially accelerated) first-passage dynamics [35,36].

Despite the importance of SIPs, their rare events have attracted systematic attention only recently [29,30,36–39]. While the dynamical large deviations (LDs) [40–48] and their corollaries—such as bounds on fluctuations [49–53]—are now well established for Markovian systems, for non-Markovian dynamics general results are mostly restricted to semi-Markov processes [54–62] (see, however, Refs. [63–65]).

Here we address this gap by studying a broad class of non-Markovian processes, specifically self-interacting jump processes (SIJPs), where interaction with the past is mediated by a functional of a general empirical observable. We analytically characterize dynamical fluctuations of SIJPs by deriving their so-called *level 2.5 of large deviations*, that is, the joint statistics of the long-time empirical occupation of configurations and empirical flux of transitions between them. From this fundamental result, and in analogy to what occurs in the Markovian case [50,51,60], one can derive general fluctuation bounds for dynamical observables, which allows us to extend both thermodynamic [49,50] and kinetic [51,52] uncertainty relations (TURs and KURs) to this wide class of non-Markov processes.

Below we present our general results for the level 2.5 LDs and fluctuation bounds of SIJPs. The key technical advance is the observation at long times of a separation between the relevant timescales for the evolution of

*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.

empirical measures, which determine the instantaneous transition rates of a SIJP, and those necessary for the microscopic jump dynamics to become fully mixing. Formal proofs appear in our accompanying long paper [66]. We illustrate our theory with two simple examples, and discuss broader implications of our results.

Self-interacting jump processes—We define a SIJP as the continuous-time chain $(X_t)_{0\le t\le T}$ of discrete configurations, $x\in\mathcal{S}$ with $|\mathcal{S}|=d$, with $\mathcal{E}=\{(x,y)\in\mathcal{S}\times\mathcal{S}: x\neq y\}$ denoting the set of all possible jumps between them (for notation and conventions see the Appendix). At time t, the SIJP jumps from the current state x to $y\neq x$ according to a rate $Q_{xy}(A_t)\in\mathbb{R}_+$ that depends on a general state-dependent *empirical observable* of the process [43,45]

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

with $f\colon \mathcal{S}\to\mathbb{R}$ a bounded function. In order for the matrix $Q(A_t)$ to be a stochastic generator, we define its diagonal components to be the negative escape rates

$$Q_{xx}(A_t) = -\sum_{y\in\mathcal{S},y\neq x} Q_{xy}(A_t). \tag{2}$$

The dependence of the generator at time t on the trajectory up to that point through Eq. (1) makes the process non-Markovian.

Two central objects of interest are the *empirical measure* [43,44]

$$L_x(t) \coloneqq t^{-1}\int_0^t \mathbf{1}_x(X_{t'})dt', \tag{3}$$

viz, the fraction of time that the current realisation of the process has spent in each configuration in $\mathcal{S}$, and the *empirical flux* [43,44]

$$\Phi_{xy}(t) \coloneqq t^{-1}\sum_{t',\Delta X_{t'}\neq 0} \mathbf{1}_x(X_{t'_-})\mathbf{1}_y(X_{t'}). \tag{4}$$

viz, the number of jumps per unit time for every $(x,y)\in\mathcal{E}$.

We assume that $Q(A_t)$ is irreducible, and therefore positively recurrent over $\mathcal{S}$, for all $t\in\mathbb{R}_+$ and $A_t\colon \mathbb{R}_+\to\mathbb{R}$ (generalization to vector valued A_t is straightforward). Additionally, we require the map $a\mapsto Q(a)$ to be at least C^1 for $a\in\mathbb{R}$ and rates $Q_{xy}(A_t)\geq q_->0$ for all $(x,y)\in\mathcal{E}$. That is, for no value of the argument of $Q(\cdot)$ the transition rates for the directed links in $\mathcal{E}$ can be smaller than some constant q_-. We further assume the existence of at least one attracting stationary state, which, unlike in Markov processes for the same assumptions, need not be unique. This implies that the empirical measure and flux converge in the long-time limit to typical values which are not necessarily unique. These assumptions are discussed in detail in Ref. [66].

For a Markov process with fixed generator Q, the stationary state π obeys $\pi Q=0$, and the empiricals converge to π and $\pi\circ Q$ [43,44]. In the SIJP case one needs the long-time limit of the observable,

$$A_t\to\bar{a}=\sum_{x\in\mathcal{S}} f_x\pi_x =: f\pi, \tag{5}$$

and the stationary state is a solution of the (in general nonlinear) equation

$$\pi Q(\bar{a}) = \pi Q(f\pi) = 0. \tag{6}$$

For a SIJP the empiricals converge to

$$L(t)\to\pi, \qquad \Phi(t)\to\varphi \tag{7}$$

for some pair (π,φ), which may depend on the trajectory and is not necessarily unique. We denote the set of limiting pairs (π,φ) as $\mathcal{A}_{\text{stat}}$.

Large deviations at level 2.5 for SIJPs—We now focus on the joint probability of the empirical measure (3) and empirical flux (4), $P_T(\ell,\phi)\coloneqq P(L(T)=\ell,\Phi(T)=\phi)$, at some long time T. We will show below that these dynamical quantities obey a large deviation principle

$$P_T(\ell,\phi)\asymp e^{-TI_{2.5}(\ell,\phi)}, \tag{8}$$

where $I_{2.5}(\ell,\phi)$ is the level 2.5 LD *rate function* for SIJPs. From this general rate function for the fluctuations of L_T and Φ_T, the LDs of any time-additive observable [like that of Eq. (1), but also those dependent on jumps] follow by contraction.

The main ingredients for obtaining Eq. (8) are as follows (see Ref. [66] for full details of derivations). The first is a formulation of the path probability of a SIJP. It has the same product structure of transition rates and waiting time factors as in a Markov process, with two distinctions: the transition rates change in time through their dependence on A_t, and the waiting times factors are no longer the exponential of a fixed escape rate times time, but of the integral of the time-varying escape rate. Noting these departures from Markovianity, the next step is to exponentially "tilt" this path measure, in analogy to what is done in the Markov case.

The third step follows after making the key observation of an asymptotic separation of timescales. At a given time t the SIJP has a characteristic time for relaxation coming from its generator $Q(A_t)$, for example, that set by the smallest escape rate. On the other hand, $Q(A_t)$ itself changes over time due its dependence on the empirical observable. However, A_t accumulates $O(1)$ changes over in $O(t)$ times, cf. Eq. (1). This means that in the long-time (large deviation) limit the dynamics of the "instantaneous"

SIJP has enough time to relax before its rates change appreciably due to the change in A_t. This can be interpreted as an adiabatic relation: fluctuations in the large deviation limit are fundamentally coupled with the dynamical changes of A_t, hence $Q(A_t)$, but homogenized over the instantaneous SIJP.

As a final step, for convenience, we make a sequence of two changes of variable inside time integrals. 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. 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 $[0, \infty)$. Taking the limit $T \to \infty$, we get

$$I_{2.5}(\ell, \phi) = \inf_{(\rho,H)\ \text{s.t.}\ (\ell,\phi)} \mathcal{I}[\rho, H], \tag{9}$$

with

$$\mathcal{I}[\rho, H] \coloneqq \sum_{(x,y)\in\mathcal{E}} \int_0^\infty e^{-\tau} \rho_x(\tau) Q_{xy}(M_\tau) \Gamma\left(\frac{H_{xy}(\tau)}{Q_{xy}(M_\tau)}\right) d\tau. \tag{10}$$

Here $\Gamma(x) \coloneqq x \log x - x + 1$ is the Cramér (or relative-entropy) rate function for Poisson statistics. In Eqs. (9) and (10) there are three auxiliary quantities: $H(\tau)$ is the generator of a time-dependent but Markovian auxiliary dynamics, $\rho(\tau)$ is its "accompanying" distribution (instantaneous stationary state)

$$\rho(\tau) H(\tau) = 0, \tag{11}$$

and M_τ the corresponding dynamical observable [Eq. (1)] if the empirical density is $\rho(\tau)$

$$M_\tau = e^\tau \int_\tau^\infty e^{-\tau'} \sum_{x\in\mathcal{S}} f_x \rho_x(\tau') d\tau'. \tag{12}$$

Furthermore, the minimization in Eq. (9) is such that

$$\ell_x = \int_0^\infty e^{-\tau} \rho_x(\tau) d\tau, \tag{13}$$

$$\phi_{xy} = \int_0^\infty e^{-\tau'} \rho_x(\tau) H_{xy}(\tau) d\tau. \tag{14}$$

Equation (9) with definition (10) and constraints Eqs. (11)–(14) provide the explicit form of the level 2.5 LDs for SIJPs. This is the central result of this Letter.

The functional $\mathcal{I}[\rho, H]$ is associated to a trajectory of the SIJP of time extent $T \gg 1$. It is local in $\rho(\tau)$ and $H(\tau)$ and therefore extensive in T. The level 2.5 rate function $I_{2.5}(\ell, \phi)$ is obtained from this functional by averaging over the ensemble of all SIJP's trajectories. In the LD asymptotics, this translates into the infimum over $(\rho(\tau), H(\tau))$, constrained on having the target long-time empirical measure and flux, giving rise to constraints (13) and (14).

Equations (9) and (10) can be interpreted as a dynamical reweighting of the Poissonian relative-entropy density $\Gamma(x)$, thus generalizing the level 2.5 LDs for Markov jump processes [42,44,67]. Unlike the classical Markov case, the contraction over (ρ, H) generically induces nonconvexity of the functional, opening the way to richer dynamical behavior in the generation of fluctuations of SIJPs. Although none of the examples discussed below exhibit this property, in the companion paper [66] we present an example of dynamical bistability which leads to nonconvexity and a change of speed in the large deviations. Additionally, the factor $e^{-\tau}$ acts as an infinite-horizon temporal discount: contributions at large τ (corresponding to short times in the original time frame) are exponentially suppressed, reflecting that early fluctuations of the empirical measure carry a smaller cost than long-time deviations.

The minimiser $(\rho_{\text{opt}}(\tau), H_{\text{opt}}(\tau))$ of Eq. (9), which is not necessarily unique, represents the optimal way to produce the fluctuation (ℓ, ϕ). Note that this corresponds to a time-dependent yet Markovian dynamics generated by $H_{\text{opt}}(\tau)$. For the special case of the SIJP generator $Q(A_\tau)$ being independent of A_τ, which makes the SIJP into a standard Markov jump process, Eqs. (9)–(13) reduce to the standard Markovian level 2.5 LD result [42,44,47]: in this case $(\rho_{\text{opt}}, H_{\text{opt}})$ become time independent, the integral in Eq. (10) is trivial, and (ℓ, ϕ) coincide with $(\rho_{\text{opt}}, \rho_{\text{opt}} \circ H_{\text{opt}})$.

As is customary in tilting arguments, we assume the existence of a time-inhomogeneous process that asymptotically realises the target fluctuation while minimizing the path-space relative entropy with respect to the original SIJP. The resulting calculation leads to a local, time-extensive rate functional, which provides strong evidence for the validity of the former assumption and corresponding large deviation principle. In Ref. [30], this program is made rigorous for affine dependence on A_t. Here, Eqs. (9)–(13) extend the formal derivation to bounded functions on S, leaving a fully rigorous mathematical proof for future work.

Kinetic uncertainty relation for SIJPs—The general level 2.5 LD result above is useful in practice in two ways. The first is that the long-time LDs of any time-extensive trajectory observable can be obtained by *contraction* from the LDs of the empirical measure and flux. The second is that Eq. (9) is variational, so that even if the exact minimization is difficult, an approximate one can provide a bound to the true rate function. This latter is the approach used to prove TURs and KURs using LD methods [50,51].

Consider a generic flux observable, that is, a quantity that depends additively on the jumps in a trajectory. For a given trajectory it can be written as a contraction with the empirical flux

$$B_T = \sum_{(x,y)\in\mathcal{E}} \beta_{xy}\Phi_{xy}(T). \tag{15}$$

Assuming a unique stationary state, at long times B_T will become typical, $B_T \to \bar{b} := \sum_{(x,y)\in\mathcal{E}} \beta_{xy}\varphi_{xy}$ in probability, cf. Eq. (6). The probability of fluctuations will have an LD form, $P(B_T = b) \asymp e^{-TI(b)}$, where the rate function $I(b)$ can be obtained from the level 2.5 one (9) with the extra constraint

$$b = \sum_{(x,y)\in\mathcal{E}} \beta_{xy}\phi_{xy}. \tag{16}$$

While exact minimisation of Eq. (9) under Eq. (16) is in general not doable, we can get an approximation with a simple ansatz where the empirical density coincides with the stationary state, $\ell^* = \pi$, but the flux is a rescaling of the stationary one, $\phi^* = (b/\bar{b})\varphi$; cf. Refs. [50,51]. In order to satisfy the Eqs. (12)–(14) we have to set

$$\rho^*(\tau) = \bar{\rho}(\tau), \qquad H^*(\tau) = (b/\bar{b})Q(\bar{M}_\tau), \tag{17}$$

that is, the time-dependent generator of the auxiliary process is simply the typical generator of the SIJP rescaled by an overall factor, and where the typical $\bar{\rho}$ is defined via Eqs. (11) and (12) for the typical $H(\tau) = Q(\bar{M}_\tau)$.

Inserting into Eq. (9) we obtain the general bound

$$I(b) \le k_{\mathrm{SIJP}}\Gamma(b/\bar{b}), \tag{18}$$

where

$$k_{\mathrm{SIJP}} := \sum_{(x,y)\in\mathcal{E}} \int_0^\infty e^{-\tau}\bar{\rho}_x(\tau)Q_{xy}(\bar{M}_\tau)d\tau \tag{19}$$

is the average *SIJP dynamical activity*. It generalises the standard dynamical activity [40,41,68] of Markov jump processes to the non-Markovian case of SIJPs which corresponds to the number of configuration changes per unit time.

Assuming that the observable (15) satisfies a central limit theorem in the long-time limit, expanding Eq. (18) to second order around $\bar{b}$ we obtain a bound on the asymptotic variance of B_T,

$$\frac{\mathrm{var}(b)}{\bar{b}^2} \ge \frac{1}{k_{\mathrm{SIJP}}}, \tag{20}$$

which we refer to as SIJP-KUR, generalizing the KUR of time-homogeneous Markov processes [51,52]. As for Markov processes, the average dynamical activity of a SIJP provides an elementary bound to the quadratic precision of observables.

Thermodynamic uncertainty relation for SIJPs—While the SIJP-KUR is valid for all fluxlike observables (15), it can be refined, as in the Markovian case, for currents, for which $\beta_{xy} = -\beta_{yx}$. This is done by writing the time-dependent flux in terms of its symmetric and antisymmetric parts, $\rho(\tau)\circ H(\tau) =: (v(\tau) + u(\tau))/2$, and solving exactly in Eqs. (9) and (10) for the symmetric part $u(\tau)$, cf. Refs. [42,44,67]. Doing this, the functional $\mathcal{I}$ in Eq. (10) is replaced by

$$\tilde{\mathcal{I}}[\rho, v] := \sum_{x<y} \int_0^\infty e^{-\tau}\Psi(v_{xy}(\tau), v^\rho_{xy}(\tau), u^\rho_{xy}(\tau))d\tau, \tag{21}$$

where

$$v^\rho_{xy}(\tau) := \rho_x(\tau)Q_{xy}(M_\tau) - \rho_y(\tau)Q_{yx}(M_\tau), \tag{22}$$

$$u^\rho_{xy}(\tau) := \rho_x(\tau)Q_{xy}(M_\tau) + \rho_y(\tau)Q_{yx}(M_\tau), \tag{23}$$

$$\Psi(p,q,r) := p\,\mathrm{arcsinh}\frac{p}{\sqrt{r^2-q^2}} - p\,\mathrm{arcsinh}\frac{q}{\sqrt{r^2-q^2}} - \left(\sqrt{p^2-q^2+r^2} - r\right), \tag{24}$$

and the minimization in Eq. (9) is done over $\rho(\tau)$ and $v(\tau)$, with constraint (14) replaced by

$$\phi_{xy} - \phi_{yx} = \int_0^\infty e^{-\tau}v_{xy}(\tau)d\tau. \tag{25}$$

The LD rate function $I(j)$ of any current observable,

$$J_T = \sum_{x<y}\beta_{xy}[\Phi_{xy}(T) - \Phi_{yx}(T)], \tag{26}$$

can be obtained from the level 2.5 (9) by contraction using Eq. (21) with constraints (11)–(13), and (25), (26). Even if the exact calculation is not practical, one can obtain a general upper bound to $I(j)$. The steps to do this are analogous to those taken to derive the TUR using an LD approach in the Markovian case [50], although more general derivations are available nowadays in that context [63,69].

The first step, after assuming a unique stationary state, is to choose the same ansatz as in the SIJP-KUR for the empirical density, $\ell^* = \pi$, which in turn implies that $\rho^*(\tau) = \bar{\rho}(\tau)$. The second step is to bound from above the function $\Psi(v_{xy}, v^{\bar{\rho}}_{xy}, u^{\bar{\rho}}_{xy})$ with a form quadratic in v_{xy}, as is done in the proof of the TUR [50]. The third step is to perform the minimization over $v(\tau)$ in the resulting bound of the functional $\tilde{\mathcal{I}}[\bar{\rho}, v]$. An application of the

Cauchy-Schwarz relation eventually yields (see Ref. [66] for detailed derivations)

$$I(j) \leq \frac{(j-\bar{j})^2}{4\int_0^\infty e^{-\tau}\frac{(j_\tau^{\bar{\rho}})^2}{\sigma_\tau}d\tau}. \tag{27}$$

where

$$j_\tau^{\bar{\rho}} := \sum_{x<y}\beta_{xy}v_{xy}^{\bar{\rho}}(\tau), \tag{28}$$

$$\sigma_\tau := \sum_{x<y}\log\left(\frac{\bar{\rho}_x(\tau)Q_{xy}(\bar{M}_\tau)}{\bar{\rho}_y(\tau)Q_{yx}(\bar{M}_\tau)}\right)v_{xy}^{\bar{\rho}}(\tau). \tag{29}$$

Over the backward logarithmic timescale, $j_\tau^{\bar{\rho}}$ is the typical instantaneous current, such that $\int_0^\infty e^{-\tau}j_\tau^{\bar{\rho}}d\tau = \bar{j}$, and σ_τ is the instantaneous *entropy production rate of the SIJP*. The interpretation follows by viewing $\bar{M}_\tau$ as an "internal" adiabatic protocol: since the empirical occupation measure evolves slowly, $Q(\bar{M}_\tau)$ changes quasistatically, allowing the underlying SIJP to equilibrate at each frozen value of time and $\bar{\rho}_\tau$ to be stationary for $Q(\bar{M}_\tau)$ [70,71]. Instantaneous entropy-production rates have also appeared in non-Markovian Langevin settings, including active Ornstein-Uhlenbeck particles [72] and generalized Langevin dynamics with memory [64], where irreversibility is likewise quantified through path-weight or dynamical-cost comparisons with auxiliary processes.

Assuming that the observable (26) satisfies a central limit theorem in the long-time limit, expanding Eq. (27) to second order around $\bar{j}$ gives a bound to the quadratic fluctuations of the current

$$\frac{\mathrm{var}(j)}{\bar{j}^2} \geq 2\int_0^\infty e^{-\tau}\frac{(j_\tau^{\bar{\rho}}/\bar{j})^2}{\sigma_\tau}d\tau, \tag{30}$$

which we refer to as SIJP-TUR, generalizing the TUR of continuous-time Markov processes [49,50]. As in the case of the TUR, Eq. (30) shows that increasing precision for the current incurs a cost in terms of instantaneous entropy production. Since it is not possible to disentangle in the integral the instantaneous entropy cost from the instantaneous current, we can interpret the instantaneous entropy production as an additional discount on the SIJP necessary to reach a certain level of precision.

Note that both the SIJP-KUR (20) and SIJP-TUR (30) reduce to the standard Markovian KUR and TUR in the limit when the SIJP becomes a Markov process. Note also that the SIJP-TUR (30) would hold in the case of *external* driving, as long as the protocol evolves as slowly as A_t in a SIJP.

Simple examples—We now illustrate the general results above with two simple examples. The first example is a two-state SIJP, $\mathcal{S} = \{0, 1\}$, with transition rates $Q_{01} = 1 + e^{\alpha A_t}$, with $A_t = L_1(t)$, and $Q_{10} = 1$, with $\alpha > 0$. This describes a spin whose flip-up rate grows exponentially with its past occupation. There is a unique equilibrium state π, which for example for $\alpha = 2$ is $\pi_0 = 0.12\ldots$ (with $\pi_1 = 1 - \pi_0$), thus favouring occupation of the up state.

Figure 1(a) illustrates the level 2.5 result by showing the rate function at "Level 2" $I_2(\ell)$, obtained minimizing Eq. (9) over the flux ϕ. This exact result coincides with Monte Carlo simulations. We also show an upper bound obtained by assuming solutions of the variational problem to be time independent [73]: for larger deviations $I_2(\ell)$ departs the Markov approximation because larger fluctuations entail sharper variations in $\rho(\tau)$ and M_τ; see Fig. 1(b).

Figure 1(c) shows the exact rate function $I(b)$ for the flow observable $B_T = \Phi_{01}(T)$ obtained by contraction from Eq. (9), and which coincides with the Monte Carlo simulations (whose range is limited to the values around the

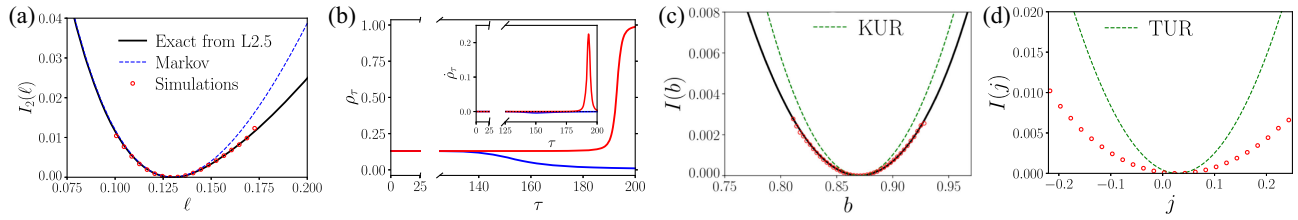


FIG. 1. (a) Exact rate function for the empirical measure, $I_2(\ell)$ (full black curve), obtained by solving Eqs. (9) and (10) with constraints (11)–(13), for the two-level SIJP with $\alpha = 2$, Monte Carlo simulations (open red circles), and Markovian approximation (dashed blue curve). (b) Optimal trajectories $\rho(\tau)$: in backward time τ, (π_0, π_1) is an unstable fixed point and the dynamics is integrated from $\tau = 200$ (which for this example is large enough to be in the asymptotic limit in τ) back to $\tau = 0$ for numerical stability. By progressively truncating the red and blue trajectories at different $\tau = \hat{\tau}$, one uncovers the most likely dynamics responsible for the fluctuation $\ell \equiv \rho_{\hat{\tau}}$. Since trajectories yielding $\ell < \pi_0$ evolve slowly (see inset), the left tail of $I_2(\ell)$ is close to that of the Markovian approximation (as the exponential discount suppresses distant-in-time contributions faster than they can accumulate). This is not the case for $\ell > \pi_0$, and the non-Markovian character becomes apparent in the higher likelihood of those fluctuations. (c) Exact rate function $I(b)$ for the flux observable $B_T = \Phi_{01}(T)$ (full black curve) for the two-level SIJP, Monte Carlo simulations (open red circles), and the SIJP-KUR bound (dashed green curve). (d) Fluctuations of the particle current in the three-state model: Monte Carlo simulations (open red circles) and the SIJP-TUR bound. All simulation results are for 10^6 trajectories of time extent $T = 1000$.

typical). We also show the SIJP-KUR upper bound from the rhs of Eq. (18).

We illustrate the SIJP-TUR in a system with a nonzero stationary current, that of a particle on a three-state ring, $\mathcal{S} = \{0, 1, 2\}$, with a self-induced drag and cyclic transitions: $Q_{01} = Q_{12} = Q_{20} = e^{-\alpha L_1(t)}$ and $Q_{10} = Q_{21} = Q_{02} = 1 - e^{-\alpha L_1(t)}$. Dynamics slows down when state 1 is visited frequently, and the more the particle lingers in one region, the harder it becomes to sustain a current. In Fig. 1(d) we compare the statistics for the particle current, $J_T = \sum_{x<y}(\Phi_{xy}(T) - \Phi_{yx}(T))$ from Monte Carlo simulations, with the SIJP-TUR bound for the rate function given in Eq. (27).

Conclusions—Self-interacting processes offer a rich setting for uncovering universal principles of non-Markovian dynamics. In the regime where the relevant timescales are well separated and the dynamics is sufficiently mixing, we have derived the exact dynamical large deviations at level 2.5, that is the full joint statistics of the empirical measure and fluxes, of self-interacting continuous-time processes.

For concreteness, we focused on SIJPs with generators dependent on empirical observables of configurations. Our results, however, are easily generalisable to self-interacting systems where the dynamics depends also on the accumulated fluxes: if A_t in Eq. (1) also adds $t^{-1} g_{xy}$ for every jump between x and y in a trajectory, the corresponding level 2.5 LDs are also given by Eqs. (9)–(14), with the change $\sum_x f_x \rho_x \to \sum_x f_x \rho_x + \sum_{xy} g_{xy} \rho_x H_{xy}$ in the integrand on the r.h.s. of Eq. (12). See Ref. [66] for a detailed derivation.

As is the case with Markovian systems, an important application of the level 2.5 LDs framework is in the derivation of bounds on the fluctuations (to all orders) of dynamical observables, which for quadratic fluctuations are the so-called thermodynamic and kinetic uncertainty relations. Here, we obtained the corresponding versions for SIJPs of the thermodynamic [49,50] and kinetic [51,52] uncertainty relations, the SIJP-TUR and SIJP-KUR. More refined and tighter bounds, such as a generalization of the "ultimate" KUR [60] (also known as the "clock" uncertainty relation [74]) can be obtained by generalising the approach of [60] to SIJPs; see Ref. [66].

Looking ahead, the present framework could, in principle, also be adapted to derive bounds for first-passage times [51,75] beyond the Markovian regime. More complex illustrative examples and applications—such as extensions to open quantum dynamics with memory—are important for clarifying the scope and implications of the theory. Another natural direction is to reweight past memory, for example, by damping contributions from the distant past or restricting the memory to a finite time window, as would arise from rapidly evaporating pheromone trails in autochemotactic microorganisms [1,2].

Acknowledgments—F. C. is grateful to Tobias Grafke for help on clarifying numerical aspects behind the two-state examples. F. C. is supported by a Leverhulme Early Career Fellowship No. ECF-2025-482. We acknowledge support from the EPSRC Grants No. EP/V031201/1 and No. EP/T022140/1. A. B. supported in part by the NSF (No. DMS-2134107, No. DMS-2152577, and No. DMS-2506010) and DARPA ARC Compass Project No. HR0011-25-3-0240.

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.

[1] E. O. Budrene and H. C. Berg, Nature (London) **376**, 49 (1995).
[2] M. P. Brenner, L. S. Levitov, and E. O. Budrene, Biophys. J. **74**, 1677 (1998).
[3] D. J. Sumpter and M. Beekman, Animal Behaviour **66**, 273 (2003).
[4] D. E. Jackson, S. J. Martin, M. Holcombe, and F. L. W. Ratnieks, Animal Behaviour **71**, 351 (2006).
[5] K. Zhao, B. S. Tseng, B. Beckerman, F. Jin, M. L. Gibiansky, J. J. Harrison, E. Luijten, M. R. Parsek, and G. C. Wong, Nature (London) **497**, 388 (2013).
[6] A. Gelimson, K. Zhao, C. K. Lee, W. T. Kranz, G. C. L. Wong, and R. Golestanian, Phys. Rev. Lett. **117**, 178102 (2016).
[7] C. R. Reid, T. Latty, A. Dussutour, and M. Beekman, Proc. Natl. Acad. Sci. U.S.A. **109**, 17490 (2012).
[8] J. R. Howse, R. A. L. Jones, A. J. Ryan, T. Gough, R. Vafabakhsh, and R. Golestanian, Phys. Rev. Lett. **99**, 048102 (2007).
[9] S. Thutupalli, R. Seemann, and S. Herminghaus, New J. Phys. **13**, 073021 (2011).
[10] B. V. Hokmabad, J. Agudo-Canalejo, S. Saha, R. Golestanian, and C. C. Maass, Proc. Natl. Acad. Sci. U.S.A. **119**, e2122269119 (2022).
[11] B. Nakayama, H. Nagase, H. Takahashi, Y. Saito, S. Hatayama, K. Makino, E. Yamamoto, and T. Saiki, Proc. Natl. Acad. Sci. U.S.A. **120**, e2213713120 (2023).
[12] E. F. Keller and L. A. Segel, J. Theor. Biol. **26**, 399 (1970).
[13] Y. Tsori and P.-G. de Gennes, Eur. Phys. Lett. **66**, 599 (2004).
[14] A. Gelimson, K. Kawasaki, and R. Golestanian, Phys. Rev. Lett. **114**, 028101 (2015).
[15] T. Grafke, M. E. Cates, and E. Vanden-Eijnden, Phys. Rev. Lett. **119**, 188003 (2017).
[16] R. Grima, Phys. Rev. Lett. **95**, 128103 (2005).
[17] A. Sengupta, S. van Teeffelen, and H. Löwen, Phys. Rev. E **80**, 031122 (2009).
[18] O. Pohl and H. Stark, Phys. Rev. Lett. **112**, 238303 (2014).
[19] W. T. Kranz, A. Gelimson, K. Zhao, G. C. L. Wong, and R. Golestanian, Phys. Rev. Lett. **117**, 038101 (2016).
[20] B. Tóth, *Self-Interacting Random Motions* (Birkhäuser, Basel, 2001), 10.1007/978-3-0348-8268-2_34.
[21] H. G. Othmer and A. Stevens, SIAM J. Appl. Math. **57**, 1044 (2006).

[22] R. Pemantle, Probab. Surv. **4**, 1 (2007).
[23] A. Erschler, B. Tóth, and W. Werner, Probab. Theory Relat. Fields **154**, 149 (2011).
[24] D. Kious, Ann. Probab. **44**, 883 (2016).
[25] A. Barbier-Chebbah, O. Bénichou, and R. Voituriez, Phys. Rev. X **12** (2022).
[26] J. Brémont, O. Bénichou, and R. Voituriez, Phys. Rev. Lett. **133**, 157101 (2024).
[27] P. Del Moral and L. Miclo, Stoch. Anal. Appl. **24**, 615 (2007).
[28] S. J. Schreiber, SIAM J. Appl. Math. **61**, 2148 (2001).
[29] F. Coghi and J. P. Garrahan, J. Stat. Mech. (2025) P103202.
[30] A. Budhiraja and F. Coghi, arXiv:2510.14659.
[31] M. Benaïm, M. Ledoux, and O. Raimond, Probab. Theory Relat. Fields **122**, 1 (2002).
[32] A. Kurtzmann, Ann. Inst. Henri Poincare Probab. Stat. **46**, 618 (2010).
[33] M. Benaïm and O. Raimond, Ann. Probab. **33**, 1716 (2005).
[34] F. Coghi, J. Phys. A **58**, 015002 (2024).
[35] A. Aleksian, P. D. Moral, A. Kurtzmann, and J. Tugaut, ESAIM Probab. Stat. **28**, 46 (2024).
[36] F. Coghi, R. Duvezin, and J. S. Wettlaufer, J. Stat. Phys. **192**, 128 (2025).
[37] S. Franchini, Stochastic Processes Appl. **127**, 3372 (2017).
[38] A. Budhiraja and A. Waterbury, Syst. Control Lett. **169**, 105379 (2022).
[39] A. Budhiraja, A. Waterbury, and P. Zoubouloglou, Stochastic Processes Appl. **186**, 104640 (2025).
[40] V. Lecomte, C. Appert-Rolland, and F. van Wijland, J. Stat. Phys. **127**, 51 (2007).
[41] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).
[42] C. Maes and K. Netočný, Eur. Phys. Lett. **82**, 30003 (2008).
[43] H. Touchette, Phys. Rep. **478**, 1 (2009).
[44] L. Bertini, A. Faggionato, and D. Gabrielli, Ann. Inst. Henri Poincare Probab. Stat. **51**, 867 (2012).
[45] R. Chetrite and H. Touchette, Ann. Henri Poincare **16**, 2005 (2015).
[46] J. P. Garrahan, Physica A (Amsterdam) **504**, 130 (2018).
[47] A. Barato, R. Chetrite, A. Faggionato, and D. Gabrielli, J. Stat. Mech. (2019) P084017.
[48] R. L. Jack, Eur. Phys. J. B **93**, 74 (2020).
[49] A. C. Barato and U. Seifert, Phys. Rev. Lett. **114**, 158101 (2015).
[50] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Phys. Rev. Lett. **116**, 120601 (2016).
[51] J. P. Garrahan, Phys. Rev. E **95**, 032134 (2017).
[52] I. D. Terlizzi and M. Baiesi, J. Phys. A **52**, 02LT03 (2018).
[53] J. M. Horowitz and T. R. Gingrich, Nat. Phys. **16**, 15 (2020).
[54] C. Maes, K. Netočný, and B. Wynants, J. Phys. A **42**, 365002 (2009).
[55] J. P. Garrahan and I. Lesanovsky, Phys. Rev. Lett. **104**, 160601 (2010).
[56] M. Shreshtha and R. J. Harris, Eur. Phys. Lett. **126**, 40007 (2019).
[57] F. Carollo, R. L. Jack, and J. P. Garrahan, Phys. Rev. Lett. **122**, 130605 (2019).
[58] B. Ertel, J. van der Meer, and U. Seifert, Phys. Rev. E **105**, 044113 (2022).
[59] C. Jia, D. Q. Jiang, and B. Wu, Electron. J. Probab **29**, 1 (2022).
[60] K. Macieszczak, arXiv:2407.09839.
[61] F. Liu, S. Xia, and S. Su, Phys. Rev. E **110**, 064145 (2024).
[62] A. M. Maier, J. H. Fritz, and U. Seifert, Phys. Rev. E **113**, 014119 (2026).
[63] G. Falasco, M. Esposito, and J.-C. Delvenne, New J. Phys. **22**, 053046 (2020).
[64] I. Di Terlizzi and M. Baiesi, J. Phys. A **53**, 474002 (2020).
[65] T. Koyuk and U. Seifert, Phys. Rev. Lett. **125**, 260604 (2020).
[66] F. Coghi and J. P. Garrahan, companion paper, Phys. Rev. E **114**, 054112 (2026).
[67] A. C. Barato and R. Chetrite, J. Stat. Phys. **160**, 1154 (2015).
[68] C. Maes, Phys. Rep. **850**, 1 (2020).
[69] A. Dechant and S.-i. Sasa, Phys. Rev. E **97**, 062101 (2018).
[70] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[71] U. Seifert, Physica A (Amsterdam) **504**, 176 (2018).
[72] L. Caprini, U. M. B. Marconi, A. Puglisi, and A. Vulpiani, J. Stat. Mech. (2019) P053203.
[73] M. D. Donsker and S. R. S. Varadhan, Commun. Pure Appl. Math. **28**, 279 (1975).
[74] K. Prech, G. T. Landi, F. Meier, N. Nurgalieva, P. P. Potts, R. Silva, and M. T. Mitchison, Phys. Rev. X **15**, 031068 (2025).
[75] T. R. Gingrich and J. M. Horowitz, Phys. Rev. Lett. **119**, 170601 (2017).

End Matter

Appendix: Notation and conventions—We use a compact notation for multidimensional objects: given vector $\pi \coloneqq (\pi_x)_{x=1:d}$ and matrix $Q \coloneqq (Q_{xy})_{x,y=1:d}$, πQ denotes their internal product, $(\pi Q)_y = \sum_{x=1}^{d} \pi_x Q_{xy}$, and use $\circ$ for 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. $\mathbf{1}_x(\cdot)$ 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 symbol $\asymp$ indicates equality up to subexponential factors in a large parameter, so that $P(\cdot) \asymp e^{-TI(\cdot)}$ is equivalent to $I(\cdot) = -\lim_{T\to\infty} T^{-1} \log P(\cdot)$. We define the asymptotic variance of an (intensive) trajectory observable such as B_T as $\mathrm{var}(b) \coloneqq \lim_{T\to\infty} T\mathrm{var}(B_T)$. We use the "maths" convention of probabilities being row vectors, stochastic generators having rows adding up to zero, and time propagation being left to right, e.g., $\partial_t \mu = \mu Q$.