Journal of Statistical Physics (2019) 176:773–805
https://doi.org/10.1007/s10955-019-02321-4

Large deviations conditioned on large deviations I: Markov chain and Langevin equation

Bernard Derrida[1] **· Tridib Sadhu**[1,2]

Received: 17 July 2018 / Accepted: 20 May 2019 / Published online: 31 May 2019

Abstract

We present a systematic analysis of stochastic processes conditioned on an empirical observable Q_T defined in a time interval $[0, T]$, for large T. We build our analysis starting with a discrete time Markov chain. Results for a continuous time Markov process and Langevin dynamics are derived as limiting cases. In the large T limit, we show how conditioning on a value of Q_T modifies the dynamics. For a Langevin dynamics with weak noise and conditioned on Q_T, we introduce large deviation functions and calculate them using either a WKB method or a variational formulation. This allows us, in particular, to calculate the typical trajectory and the fluctuations around this trajectory when conditioned on a certain value of Q_T, for large T.

Keywords Conditioned stochastic process · Markov chain · Langevin dynamics · Large deviation function

1 Introduction

Understanding the frequency of rare events and the dynamical trajectories, which generate them has become an important field of research in many physical situations including protein folding [1], chemical reactions [2,3], atmospheric activities [4], glassy systems [5,6], disordered media [7]. From the mathematical point of view, the statistical properties of rare events are characterized by large deviation functions [8–15]. In physics, a particular interest for large deviation functions arose in the context of non-equilibrium statistical physics from the discovery of the fluctuation theorem [16–18], where the rare event consists in observing an atypical value of a current over a long time window. They also had been used for a long time to study stochastic dynamical systems in a weak noise limit [19–21] or extended systems when the system size becomes large [22–24].

✉ Tridib Sadhu
tridib@theory.tifr.res.in

Bernard Derrida
bernard.derrida@college-de-france.fr

1 Collège de France, 11 Place Marcelin Berthelot, 75231 Paris Cedex 05, France

2 Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India

One of the simplest questions one may ask about the large deviation functions is to consider an empirical observable Q_T of the form

$$Q_T = \int_0^T dt\ f(C_t) \tag{1}$$

where $f(C_t)$ is a function of the configuration C_t of a stochastic (or a chaotic) system at time t and to try to determine the probability that this empirical observable takes any atypical value $q\,T$. The corresponding large deviation function $\phi(q)$ is then simply defined by [25–34]

$$\text{Prob}(Q_T = qT) \sim e^{-T\,\phi(q)} \qquad \text{for large } T. \tag{2}$$

(Here, the precise meaning of the symbol $\sim$ is that $\lim_{T\to\infty} \frac{1}{T} \log \text{Prob}(qT) = -\phi(q)$, and this will be used throughout this article.) A rather common situation is when $\phi(q)$ vanishes at a single value q^* of q (the most likely value of q) and where $\phi(q) > 0$ for $q \neq q^*$. The main question raised in the present paper is what are the dominant trajectories of a stochastic process, which contribute to this large deviation function and how to describe their effective dynamics. In particular, we want to determine the probability $P_t(C|Q_T = q\,T)$ of finding the system in a configuration C at an arbitrary time t, conditioned on a certain value of Q_T for large T. Many of these questions have been studied earlier in different contexts spanning over Physics, Mathematics, and Computer Science [33–43].

In recent years, a theory for this conditioning problem in a Markov process has been developed [44–48]. The analysis is based on a canonical approach, which consists in weighting all the events by an exponential of Q_T and then to determine the probability

$$P_t^{(\lambda)}(C) = \frac{\int dQ\, e^{\lambda Q} P_t(C, Q)}{\sum_{C'} \int dQ\, e^{\lambda Q} P_t(C', Q)} \tag{3}$$

where $P_t(C, Q)$ is the joint probability of configuration C at time t and the observable Q_T to take value Q given the system in its steady state. This is in contrast to the previous case where Q_T in (1) was fixed (that we call the microcanonical case). As we will see (in particular, in Sect. 2 and Appendix A), results for the microcanonical case can be obtained using an equivalence of ensembles, which relates these canonical and microcanonical ensembles in the usual way in the large T limit (which plays here the same role as the thermodynamic limit in standard statistical mechanics). This analogy for the two ensembles as canonical and microcanonical has been used earlier [32,35,36,44,46,47,49] as well as their equivalence has been established [47]. (In earlier works the canonical ensemble has been referred as the tilted, biased, or s-ensemble [6,35,44,49].) Using the equivalence of ensembles, it was rigorously shown [47] that, in the large T limit, the conditioned dynamics can be effectively described by a Markov process.

In this paper, we shall follow the canonical approach [44–48]. In the first half of the paper we give an alternative derivation of many results obtained earlier [44–48] mostly in the quasi-stationary regime (region III of Fig. 1). We build our analysis for a discrete time Markov process on a finite configuration space. Then, the continuous time Markov process and Langevin dynamics are obtained as limiting cases. Compared to the rigorous approach in [44–48], our derivation is hopefully easier for a general Physics audience. Moreover, it allows us to easily generalize the results to all other regions of Fig. 1, giving explicit results for the conditioned probability and the effective dynamics.

The second half of this paper is devoted to the weak noise limit of the Langevin dynamics. This limit has been recently studied [50,51] for specific examples with periodic boundary condition, in their quasi-stationary regime. Application of these ideas for interacting many-body systems will be presented in a forthcoming publication [52].

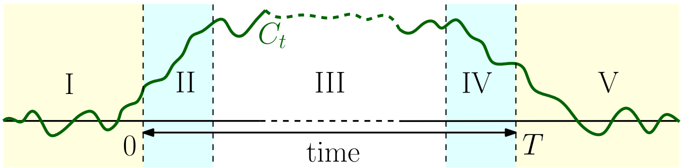

Fig. 1 A schematic of a time evolution of a Markov process C_t when conditioned on an empirical observable Q_T measured in a large time interval $[0, T]$. Different regions denote different parts of the evolution: (I) $t < 0$, (II) $t \geq 0$ but small, (III) t and $T - t$ both large (quasi-stationary regime), (IV) $T - t > 0$ but small, and (V) $t \geq T$

We will start by reviewing and extending some known [44–48] aspects of the conditioning problem for Markov processes and for the Langevin equation (see Sects. 2 and 3). In the large T limit, one has to distinguish five regions (see Fig. 1) for which we calculate how the measure and the dynamics are modified by the conditioning on Q_T. Then, we will consider the Langevin equation in the weak noise limit, first using a Wentzel–Kramers–Brillouin (WKB) approach [53] (Sects. 4 and 5) and then a variational approach (Sect. 6) based on the search of an optimal path which minimizes an action. This will allow us in particular to obtain the equation followed by the optimal trajectory under conditioning for large T. Lastly, we will see in Sect. 7 that an effect of conditioning is to break causality in the sense that a trajectory becomes correlated to the noise in the future.

2 Markov process

A schematic time evolution of a Markovian stochastic system conditioned to take a certain value of Q_T is shown in Fig. 1. For large T, one has to consider five regions. The system starts in its steady state far in the past, then evolves to a quasi-stationary regime (region III in Fig. 1), and finally relaxes to the steady state of the unconditioned dynamics. It is known [44–47] how to construct an effective dynamics that describes the conditioned process in the large T limit. For a Markov chain, the effective dynamics is Markovian with transition rate, which can be expressed (for large T) in terms of the largest eigenvalue and eigenvectors of the tilted Markov matrix [44–47] in the biased ensemble. Similar connections between a conditioned ensemble and a biased ensemble appeared earlier in many contexts: rare events problems [26,31,35,37,40,54–57], kinetically constrained models [5,6], optimal control theory [39,48], and also in Quantum systems [58]. The effective dynamics for a Markov process has been studied in depth, mostly for the quasi-stationary regime in [47]. In this section, we give another derivation of the effective dynamics, which extends earlier results in the five regions of Fig. 1.

2.1 The tilted matrix

We focus here our discussion on a discrete time, irreducible, aperiodic, time-homogeneous Markov process [42] on a finite set of configurations. This Markov process is specified by the probability $M_0(C', C)$ that the system jumps from configuration C to C' in one time step. This means that the probability $P_t(C)$ to be in configuration C at time t evolves by

$$P_{t+1}(C') = \sum_C M_0(C', C) P_t(C).$$

We consider that at $t \to -\infty$, the process starts in its steady state. (We will see later that the continuous time Markov process and the Langevin dynamics can be obtained as limiting cases.)

For this discrete time Markov process, we want to condition on a general empirical observable [32,44,47]

$$Q_T = \sum_{t=0}^{T-1} f(\mathcal{C}_t) + \sum_{t=0}^{T-1} g(\mathcal{C}_{t+1}, \mathcal{C}_t) \tag{4}$$

where f and g are arbitrary functions of the configurations. For example, by choosing $f(C) = \delta_{C,C_a}$ and $g(C', C) = 0$, the observable Q_T represents the total time spent in a particular configuration C_a. Another choice $f(C) = 0$ and $g(C', C) = \delta_{C',C_b}\delta_{C,C_a}$ would count the total number of jumps from configuration C_a to configuration C_b. Large deviations of such empirical observables and their conditioning have been studied in the recent past [6,32–34,44–48,59].

Our interest is to describe the dynamics conditioned on a certain value of Q_T in the large T limit. In particular, we want to know what is the conditional probability $\mathcal{P}_t(C|Q_T)$ for the system to be in a configuration C at an arbitrary time t when conditioned on the observable Q_T defined by (4).

Let us first analyze the special case $t = T$. If we define the joint probability $P_T(C, Q|C_0)$ for the system to be in a configuration C at time T and that the observable Q_T defined by (4) takes value Q given its initial configuration C_0 at time 0, it satisfies a recursion relation:

$$P_T(C, Q|C_0) = \sum_{C'} M_0(C, C') P_{T-1}(C', Q - f(C') - g(C, C')|C_0) \tag{5}$$

Then, it is easy to see that the generating function defined by

$$G_T^{(\lambda)}(C|C_0) = \int dQ\, e^{\lambda Q} P_T(C, Q|C_0) \tag{6}$$

satisfies

$$G_T^{(\lambda)}(C|C_0) = \sum_{C'} M_\lambda(C, C') G_{T-1}^{(\lambda)}(C'|C_0) \tag{7}$$

where

$$M_\lambda(C, C') = M_0(C, C') e^{\lambda[f(C')+g(C,C')]} \tag{8}$$

is the tilted matrix. Therefore, $G_T^{(\lambda)}(C|C_0) = M_\lambda^T(C, C_0)$ is the (C, C_0)th element of the matrix $(M_\lambda)^T$. For large T (and for real λ), the matrix elements of $(M_\lambda)^T$ are dominated by the largest eigenvalue $e^{\mu(\lambda)}$ of M_λ, resulting in

$$G_T^{(\lambda)}(C|C_0) \simeq e^{T\mu(\lambda)} R_\lambda(C) L_\lambda(C_0) \tag{9}$$

where $R_\lambda(C)$ and $L_\lambda(C)$ are the associated right and left eigenvectors, respectively. In (9) the symbol $\simeq$ is used, as physicists usually do, to mean that the ratio of the two sides of the equation becomes 1 in the limit $T \to \infty$; in fact, as we are considering an irreducible, aperiodic Markov process on a finite configuration space, the Perron-Frobenius theorem [60] ensures that there is a non-vanishing spectral gap and corrections to (9) are exponentially small. For the prefactor in (9) to be correct the eigenvectors must be normalized with $\sum_C R_\lambda(C) L_\lambda(C) = 1$.

For earlier uses of the tilted matrix see [6,12,26,44,54,55,61] and references therein. For more recent work see [47] where (9) also appears.

Remarks 1. It follows from (6, 9) that the cumulants of Q_T, for large T, can be obtained [6,26,47,54] from the derivatives of $\mu(\lambda)$ at $\lambda = 0$, and that $\lim_{T\to\infty} \frac{1}{T} \log\langle e^{\lambda Q_T}\rangle = \mu(\lambda)$.

2. In the case $\lambda = 0$, the largest eigenvalue is 1, with $L_0(C) = 1$, and $R_0(C)$ is the steady state probability distribution of the Markov process M_0.

2.2 Ensemble equivalence

By an inverse Laplace transformation (6) becomes

$$P_T(C, Q|C_0) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} dz\, e^{-zQ}\, G_T^{(z)}(C|C_0), \tag{10}$$

where the integral is along the imaginary axis on the complex-z plane.

For an irreducible, aperiodic, time-homogeneous Markov chain on a finite configuration space, $\mu(\lambda)$ is differentiable and convex. When $\mu(\lambda)$ is strictly convex, using the asymptotics (9) for large T and the method of steepest descent we get

$$P_T(C, Q = qT|C_0) \simeq e^{-T\phi(q)}\sqrt{\frac{1}{2\pi T\mu''(\lambda)}}R_\lambda(C)L_\lambda(C_0) \tag{11a}$$

where the large deviation function $\phi(q)$ and the eigenvalue $e^{\mu(\lambda)}$ are related by a Legendre transformation (the Gärtner–Ellis theorem [14,62])

$$\phi(q) = \lambda q - \mu(\lambda) \qquad \text{with} \quad \mu'(\lambda) = q. \tag{11b}$$

For convex $\mu(\lambda)$, where the Legendre transformation is self-dual, (11b) gives

$$\mu(\lambda) = \lambda q - \phi(q) \qquad \text{with} \qquad \lambda = \phi'(q). \tag{12}$$

Moreover, as $\mu''(\lambda) = 1/\phi''(q)$, using (11a) we get

$$P_T(C, Q = qT|C_0) \simeq e^{-T\phi(q)}\sqrt{\frac{\phi''(q)}{2\pi T}}R_{\phi'(q)}(C)L_{\phi'(q)}(C_0) \tag{13}$$

One simple way to understand the prefactors in (13) is to use it in (6) to recover (9) by a saddle point calculation. As for (9), the symbol $\simeq$ in (13) means that the ratio of the two sides goes to 1 in the limit $T \to \infty$. However, unlike (9), the higher order correction term to (13) (which could be determined using again (6) and (9)) would be algebraic in T rather than exponential.

We see from (13) that, for large T, the conditional distribution of C at the final time is given by

$$\mathcal{P}_T(C|Q = qT) = \frac{P_T(C, Q = qT|C_0)}{\sum_{C'} P_T(C', Q = qT|C_0)} \simeq \frac{R_{\phi'(q)}(C)}{\sum_{C'} R_{\phi'(q)}(C')} \tag{14}$$

This shows that the initial condition C_0 is forgotten at large T. On the other hand, in the canonical ensemble, using (9) one has for the probability at the final time

$$P_T^{(\lambda)}(C) = \frac{G_T^{(\lambda)}(C|C_0)}{\sum_{C'} G_T^{(\lambda)}(C'|C_0)} \simeq \frac{R_\lambda(C)}{\sum_{C'} R_\lambda(C')} \tag{15}$$

We see that the two expressions (14) and (15) coincide by choosing $\lambda = \phi'(q)$. This shows that, for large T, the two ensembles are equivalent: fixing the value of $Q_T = q\,T$ or weighting the events by a factor $e^{\lambda Q_T}$ with $\lambda = \phi'(q)$ lead to asymptotically the same distribution of the final configuration C. This equivalence of ensembles has been established earlier in [46,47].

Remark The equivalence might not hold for systems with infinitely many configurations, when the spectral gap for the tilted matrix vanish and $\mu(\lambda)$ could become non-differentiable [44,63–65]. See [47,49,66] for conditions for the equivalence of ensembles to hold.

2.3 The measure conditioned on Q_T for large T

As shown in Appendix A, the equivalence of ensembles holds not only at time $t = T$, but at any time t, as long as T is large. The same was established earlier in [45–47]. This states that, by generalizing (15), if we define the canonical measure

$$P_t^{(\lambda)}(C) = \frac{\int dQ\, e^{\lambda Q} P_t(C, Q)}{\sum_{C'} \int dQ\, e^{\lambda Q} P_t(C', Q)} \tag{16}$$

for any time t, then for large T,

$$\mathcal{P}_t(C|Q_T = qT) \simeq P_t^{(\lambda)}(C) \qquad \text{with} \qquad \lambda = \phi'(q) \tag{17}$$

where $P_t(C, Q)$ is the joint probability of configuration C at time t and the observable Q_T to take value Q given the system in its steady state; $\mathcal{P}_t(C|Q)$ is the corresponding conditional probability. As for (13), the symbol $\simeq$ in (17) means that the ratio of the two sides goes to 1 in the limit $T \to \infty$.

This canonical measure (16), for large T, takes different expressions in the five regions indicated in Fig. 1. (A derivation is presented in Appendix A for region II and can be easily extended for other regions. Many of these results can be inferred from the analysis in [46,47].)

- Region I. $t < 0$

$$P_t^{(\lambda)}(C) = \frac{\sum_{C'} L_\lambda(C') M_0^{-t}(C', C) R_0(C)}{\sum_{C'} L_\lambda(C') R_0(C')} \tag{18a}$$

- Region II. $0 \leq t \ll T$. One recovers an earlier [46,47] result

$$P_t^{(\lambda)}(C) = \frac{\sum_{C'} L_\lambda(C) M_\lambda^t(C, C') R_0(C')}{e^{t\mu(\lambda)} \sum_{C'} L_\lambda(C') R_0(C')} \tag{18b}$$

- Region III. $1 \ll t$ and $T - t \gg 1$. One recovers an earlier [46,47] result

$$P_t^{(\lambda)}(C) = R_\lambda(C) L_\lambda(C) \tag{18c}$$

- Region IV. $1 \ll t < T$, i.e. $T - t = \mathcal{O}(1)$

$$P_t^{(\lambda)}(C) = \frac{\sum_{C'} M_\lambda^{T-t}(C', C) R_\lambda(C)}{e^{(T-t)\mu(\lambda)} \sum_{C'} R_\lambda(C')} \tag{18d}$$

- Region V. $T \leq t$

$$P_t^{(\lambda)}(C) = \frac{\sum_{C'} M_0^{t-T}(C, C') R_\lambda(C')}{\sum_{C'} R_\lambda(C')} \tag{18e}$$

To be consistent with the notation of Sect. 2.1 we denote by $R_0(C)$ the steady state measure of the Markov process M_0. Therefore (15) is a special case of (18e). Another special case

$$P_{t=0}^{(\lambda)}(C) = \frac{L_\lambda(C)R_0(C)}{\sum_{C'} L_\lambda(C')R_0(C')} \tag{19}$$

2.4 Time evolution of the tilted process

Again by a straightforward generalization of the reasoning (see Appendix A), one can show that the equivalence of ensembles holds for the dynamics as well [32,44–47]. In fact, the tilted dynamics in the canonical ensemble, where events are weighted by $e^{\lambda Q_T}$, is itself a Markov process [32,44–47] even for small T (see Sect. 2.5). For this process, the probability of jump $W_t^{(\lambda)}(C', C)$ from configuration C at t to C' at $t+1$ depends, in general, on time t. For example, for $t < 0$,

$$W_t^{(\lambda)}(C', C) = \frac{\sum_{C''', C''} M_\lambda^T(C''', C'')\ M_0^{-t-1}(C'', C')\ M_0(C', C)R_0(C)}{\sum_{C''', C''} M_\lambda^T(C''', C'')\ M_0^{-t}(C'', C)R_0(C)}$$

while for $0 \leq t < T$,

$$W_t^{(\lambda)}(C', C) = \frac{\sum_{C'', C_0} M_\lambda^{T-t-1}(C'', C')M_\lambda(C', C)M_\lambda^t(C, C_0)R_0(C_0)}{\sum_{C'', C_0} M_\lambda^{T-t}(C'', C)M_\lambda^t(C, C_0)R_0(C_0)}$$

For $t \geq T$, the transition probability is same as in the unconditioned dynamics, $W_t^{(\lambda)}(C', C) = M_0(C', C)$.

For large T, the dominant contribution comes from the largest eigenvalue of M_λ, and one gets in the five regions of Fig. 1:

– Region I.

$$W_t^{(\lambda)}(C', C) = \frac{\sum_{C''} L_\lambda(C'')\ M_0^{-t-1}(C'', C')\ M_0(C', C)}{\sum_{C''} L_\lambda(C'')\ M_0^{-t}(C'', C)} \tag{20a}$$

– Region II and III.

$$W_t^{(\lambda)}(C', C) = \frac{L_\lambda(C')\ M_\lambda(C', C)}{e^{\mu(\lambda)} L_\lambda(C)} \tag{20b}$$

This result for region III has been obtained earlier in [46,47].

– Region IV.

$$W_t^{(\lambda)}(C', C) = \frac{\sum_{C''}\ M_\lambda^{T-t-1}(C'', C')\ M_\lambda(C', C)}{\sum_{C''}\ M_\lambda^{T-t}(C'', C)} \tag{20c}$$

– Region V.

$$W_t^{(\lambda)}(C', C) = M_0(C', C) \tag{20d}$$

Using these expressions for $W_t^{(\lambda)}$ and the corresponding canonical measure in (18a–18d), one can check that

$$P_{t+1}^{(\lambda)}(C') = \sum_C W_t^{(\lambda)}(C', C)P_t^{(\lambda)}(C) \tag{21}$$

Remarks 1. We have seen that by deforming the matrix M_0 one can condition on two kinds of observables: $f(C_t)$ and $g(C_{t+1}, C_t)$ [see (4)]. It is not possible to condition on other time correlations, like, $Q_T = \sum_{t=1}^{T} g(\mathcal{C}_{t+\tau}, \mathcal{C}_t)$ with $\tau > 1$ by simply deforming the matrix M_0. One could still define a tilted Markov process but this would be on a much larger set of configurations since one would need to keep information about τ consecutive configurations.

2. In a similar analysis one can describe the time reversed process [42] conditioned on Q_T. We define $\mathbb{W}_t^{(\lambda)}(C, C')$ as the transition probability to jump from C' at $t+1$ to C at t in the time reversed process. In all five regions of time, they could be expressed in terms of the corresponding $W_t^{(\lambda)}$ and $P_t^{(\lambda)}$ of the forward process.

$$\mathbb{W}_t^{(\lambda)}(C, C') = W_t^{(\lambda)}(C', C)\frac{P_t^{(\lambda)}(C)}{P_{t+1}^{(\lambda)}(C')} \tag{22}$$

For example, in the quasi-stationary regime ($1 \ll t$ and $T - t \gg 1$),

$$\mathbb{W}_t^{(\lambda)}(C, C') = \frac{M_\lambda(C', C)R_\lambda(C)}{e^{\mu(\lambda)}R_\lambda(C')}. \tag{23}$$

The time reversed process is useful in describing how a fluctuation is created. For example, the fluctuation leading to an atypical configuration can be described by relaxation from the same configuration in the time reversed process [52].

2.5 A generalization

The above expressions (18a–18e) and (20a–20d) can be extended for a more general observable of the form

$$Q = \sum_t f_t(C_t) + \sum_t g_t(C_{t+1}, C_t) \tag{24}$$

where $f_t(C)$ and $g_t(C', C)$ are arbitrary functions of configurations in a discrete time irreducible Markov process $M_0(C', C)$ on a finite configuration space. The observable (4) is just a particular case of (24) with $f_t(C) = f(C)$ and $g_t(C', C) = g(C', C)$ for $t \in [0, T]$ with large T, and both being zero outside this time window.

We consider that the system started at $t \to -\infty$ in its steady state and evolves till $t \to \infty$, but this can be changed without affecting much of our analysis. One can even generalize to the case when the Markov process $M_0(C', C)$ depends on time.

Using a reasoning similar to that in Appendix A, we can show that in the canonical ensemble, where the events are weighted by $e^{\lambda Q}$, the probability $P_t^{(\lambda)}(C)$ is given by

$$P_t^{(\lambda)}(C) = \frac{Z_t^{(\lambda)}(C)\,\mathbb{Z}_t^{(\lambda)}(C)}{\sum_{C'} Z_t^{(\lambda)}(C')\,\mathbb{Z}_t^{(\lambda)}(C')} \tag{25a}$$

where $Z_t^{(\lambda)}(C)$ and $\mathbb{Z}_t^{(\lambda)}(C)$ follow the recursion relations

$$Z_t^{(\lambda)}(C) = \sum_{C'} e^{\lambda f_{t-1}(C') + \lambda g_{t-1}(C, C')} M_0(C, C') Z_{t-1}^{(\lambda)}(C') \tag{25b}$$

$$\mathbb{Z}_t^{(\lambda)}(C) = \sum_{C'} e^{\lambda f_t(C) + \lambda g_t(C', C)} M_0(C', C) \mathbb{Z}_{t+1}^{(\lambda)}(C') \tag{25c}$$

We can also show that the tilted dynamics remains Markovian, and $P_t^{(\lambda)}(C)$ follows (21) with the transition probability

$$W_t^{(\lambda)}(C', C) = \frac{\mathbb{Z}_{t+1}^{(\lambda)}(C') M_0(C', C) e^{\lambda f_t(C) + \lambda g_t(C', C)} Z_t^{(\lambda)}(C)}{\sum_{C''} \mathbb{Z}_{t+1}^{(\lambda)}(C'') M_0(C'', C) e^{\lambda f_t(C) + \lambda g_t(C'', C)} Z_t^{(\lambda)}(C)}$$
$$= \frac{\mathbb{Z}_{t+1}^{(\lambda)}(C')}{\mathbb{Z}_t^{(\lambda)}(C)} e^{\lambda f_t(C) + \lambda g_t(C', C)} M_0(C', C) \tag{26}$$

One can verify using (25c) that $\sum_{C'} W_t^{(\lambda)}(C', C) = 1$.

The expressions (18a–18e) and (20a–20d) for $Q = Q_T$ in (4) can be easily recovered from (25a) and (26) by using the corresponding $f_t(C)$ and $g_t(C', C)$ and taking large T limit.

2.6 Continuous time Markov process

The case of a continuous time Markov process can be obtained [60] by choosing a Markov matrix M_0 in the discrete time case of the form

$$M_0(C', C) = \left(1 - \sum_{C''} \mathcal{M}_0(C'', C) dt\right) \delta_{C', C} + \mathcal{M}_0(C', C)\, dt + \cdots \tag{27}$$

and subsequently taking the limit $dt \to 0$ in the corresponding Master equation. The $\mathcal{M}_0(C', C)$ is the jump rate from configuration C to C'. Following this construction it is straightforward to extend the results of conditioned process in the discrete time case to the continuous time case. The details are given in Appendix B.

3 The Langevin dynamics

We now extend the above discussion to a Langevin process on an infinite line defined by the stochastic differential equation

$$\dot{X}_t = F(X_t) + \eta_t \tag{28}$$

where $F(x)$ is an external force and η_t is a Gaussian white noise of mean zero and covariance $\langle \eta_t \eta_{t'} \rangle = \epsilon\, \delta(t - t')$ with ϵ being the noise strength. It is well known [60] that the probability $P_t(x)$ of the process X_t to be in x at time t follows a Fokker–Planck equation

$$\frac{d}{dt} P_t(x) = \mathcal{L}_0 \cdot P_t(x) := -\frac{d}{dx} [F(x) P_t(x)] + \frac{\epsilon}{2} \frac{d^2}{dx^2} P_t(x) \tag{29}$$

3.1 The tilted Fokker–Planck operator

Our interest is the dynamics conditioned on an empirical observable, considered already in [32,44–47],

$$Q_T = \int_0^T dt\, f(X_t) + \int_0^T dX_t\, h(X_t) \tag{30}$$

where f and h are functions of X_t. In writing the second integral we mean a special class of observables whose discrete analogue

$$\int_0^T dX_t \; h(X_t) \equiv \sum_t (X_{t+dt} - X_t) \left[\alpha \, h(X_{t+dt}) + (1-\alpha) \, h(X_t)\right] \tag{31}$$

with $\alpha \in [0, 1]$. The choice $\alpha = 0$ corresponds to the Îto integral and $\alpha = \frac{1}{2}$ corresponds to the Stratonovich integral in stochastic calculus [67]. One may view (30) as a limiting case of (4).

A large number of relevant empirical observables in statistical physics are of the form (30). For example, integrated current, work, entropy production, empirical density, etc [18,32–34,68,69].

The Langevin dynamics in (28) can be viewed as a continuous space and time limit of a jump process on a one-dimensional chain (see Appendix C). This way, the effective dynamics conditioned on Q_T in (30) can be obtained from our results in Sect. 2 by suitably taking the continuous limit. For example, a continuous limit of (7) gives (see Appendix C)

$$\frac{d}{dT} G_T^{(\lambda)}(x|y) = \mathcal{L}_\lambda \cdot G_T^{(\lambda)}(x|y) \tag{32}$$

where the tilted Fokker–Planck operator

$$\mathcal{L}_\lambda := \lambda f(x) - \left(\frac{d}{dx} - \lambda h(x)\right) F(x) + \frac{\epsilon}{2}\left(\frac{d}{dx} - \lambda h(x)\right)^2 + \epsilon \left(\alpha - \frac{1}{2}\right) \lambda h'(x) \tag{33}$$

For an earlier derivation of (33) when $f(x) = 0$ see [68,69], and for the general case see [46,47].

If there is a non-vanishing spectral gap for the largest eigenvalue $\mu(\lambda)$ of $\mathcal{L}_\lambda$, then for large T, one gets, analogous to (9),

$$G_T^{(\lambda)}(x|y) \simeq e^{T\mu(\lambda)} r_\lambda(x) \ell_\lambda(y) \tag{34}$$

where $r_\lambda(x)$ and $\ell_\lambda(x)$ are the corresponding right and left eigenvectors defined by

$$\mathcal{L}_\lambda \cdot r_\lambda(x) = \mu(\lambda) r_\lambda(x) \quad \text{and} \quad \mathcal{L}_\lambda^\dagger \cdot \ell_\lambda(x) = \mu(\lambda) \ell_\lambda(x) \tag{35}$$

where $\mathcal{L}_\lambda^\dagger$ is the operator conjugate to $\mathcal{L}_\lambda$.

$$\mathcal{L}_\lambda^\dagger := \lambda f(x) + F(x)\left(\frac{d}{dx} + \lambda h(x)\right) + \frac{\epsilon}{2}\left(\frac{d}{dx} + \lambda h(x)\right)^2 + \epsilon \left(\alpha - \frac{1}{2}\right) \lambda h'(x) \tag{36}$$

In (34) the symbol $\simeq$ means that the sub-leading terms are exponentially small in T. Analogous to (9), for the expression (34) the eigenfunctions should satisfy $\int dx \, \ell(x) r(x) = 1$, as discussed in [32,47].

Remark Unlike (9) in the discrete Markov process, the existence of a spectral gap for (34) is not assured (see the discussion in [47,59]). On a one-dimensional line, where $F(x)$ and $h(x)$ are gradients, (35) can be mapped [32,59,69,70] to a Schrödinger equation with potential

$$V(x) = \frac{F(x)^2}{2\epsilon} + \frac{F'(x)}{2} - \lambda f(x) - \epsilon\left(\alpha - \frac{1}{2}\right) \lambda h'(x). \tag{37}$$

In this case, the question of a spectral gap for $\mathcal{L}_\lambda$ maps to the existence of a bound state of the Schrödinger equation with a potential $V(x)$, which is a well studied problem in Quantum mechanics (see [53,71–74]). On an infinite line, if $V(x)$ grows when $|x| \to \infty$, then there is a bound state [70].

3.2 Canonical measure for the Langevin dynamics

One could similarly derive the canonical measure and the corresponding rate equation. This way (18a–18e) become, for the continuous analogue $P_t^{(\lambda)}(x)$ of the canonical probability (16) in the five regions of Fig. 1 (see the derivation in Appendix C)

- Region I

$$P_t^{(\lambda)}(x) = \frac{\left[e^{-t\mathcal{L}_0^\dagger} \cdot \ell_\lambda\right](x)\, r_0(x)}{\int dy\, \ell_\lambda(y)\, r_0(y)} \tag{38a}$$

- Region II

$$P_t^{(\lambda)}(x) = \frac{\ell_\lambda(x)\left[e^{t\mathcal{L}_\lambda} \cdot r_0\right](x)}{e^{t\mu(\lambda)} \int dy\, \ell_\lambda(y)\, r_0(y)} \tag{38b}$$

- Region III

$$P_t^{(\lambda)}(x) = \ell_\lambda(x) r_\lambda(x) \tag{38c}$$

- Region IV

$$P_t^{(\lambda)}(x) = \frac{\left[e^{(T-t)\mathcal{L}_\lambda^\dagger} \cdot \ell_0\right](x)\, r_\lambda(x)}{e^{(T-t)\mu(\lambda)} \int dy\, r_\lambda(y)} \quad \text{with} \quad \ell_0(x) = 1 \tag{38d}$$

- Region V

$$P_t^{(\lambda)}(x) = \frac{\left[e^{(t-T)\mathcal{L}_0} \cdot r_\lambda\right](x)}{\int dy\, r_\lambda(y)} \tag{38e}$$

These expressions of $P_t^{(\lambda)}(x)$, particularly (38b–38d), were already written in [46,47].

The time evolution of the tilted dynamics is described by a Langevin equation (28) with a modified force $F_t^{(\lambda)}(x)$, which, in general, depends on time. The force takes different expressions in the five regions indicated in Fig. 1.

- Region I

$$F_t^{(\lambda)}(x) = F(x) + \epsilon \frac{d}{dx} \log\left[e^{-t\mathcal{L}_0^\dagger} \cdot \ell_\lambda(x)\right] \tag{39a}$$

- Region II and III, we recover an earlier result [46,47],

$$F_t^{(\lambda)}(x) = F(x) + \epsilon\left(\lambda h(x) + \frac{d}{dx} \log \ell_\lambda(x)\right) \tag{39b}$$

- Region IV

$$F_t^{(\lambda)}(x) = F(x) + \epsilon\left(\lambda h(x) + \frac{d}{dx} \log\left[e^{(T-t)\mathcal{L}_\lambda^\dagger} \cdot \ell_0(x)\right]\right) \tag{39c}$$

- Region V

$$F_t^{(\lambda)}(x) = F(x) \tag{39d}$$

A derivation is given in Appendix C. One can easily verify that the probability (38a–38e) follows a Fokker–Planck equation with the corresponding force (39a–39d). To see this, for example in region I, one can simply use that $\left[e^{-t\mathcal{L}_0^\dagger}\cdot \ell_\lambda\right](x) \equiv V_t(x)$ in (38a) is a solution of $\frac{d}{dt}V_t(x) = -\mathcal{L}_0^\dagger \cdot V_t(x)$ and that $\mathcal{L}_0 \cdot r_0(x) = 0$.

Remark We have considered the noise amplitude ϵ in (28) to be a constant. A generalization where the amplitude is a function of X_t involves a choice of the Îto-Stratonovich discretization [67]. The analysis could be easily extended to such cases, as well as in higher dimensions.

3.3 The Ornstein–Uhlenbeck process

As an illustrative easy example [32] one can consider the Langevin equation in a harmonic potential, $F(x) = -\gamma x$. This is known as the Ornstein–Uhlenbeck process [60,70]. To make our discussion simple, we choose the observable $Q_T = \int_0^T ds\ X_s$ which corresponds to $f(x) = x$ and $h(x) = 0$ in (30). In this case, the tilted Fokker–Planck operator (33) gives

$$\mathcal{L}_\lambda := \lambda x + \gamma \frac{d}{dx} x + \frac{\epsilon}{2}\frac{d^2}{dx^2}$$

Its largest eigenvalue and the corresponding eigenvectors are [32,70]

$$\mu(\lambda) = \frac{\epsilon\lambda^2}{2\gamma^2}; \qquad r_\lambda(x) = \mathcal{N} e^{-\frac{\gamma}{\epsilon}\left(x-\frac{\mu}{\lambda}\right)^2}; \qquad \ell_\lambda(x) = e^{\frac{\lambda}{\gamma}x} \tag{40}$$

with $\mathcal{N}$ determined from normalization $\int dx \ell_\lambda(x) r_\lambda(x) = 1$. The Legendre transformation (12) gives the large deviation function $\phi(q) = \frac{\gamma^2}{2\epsilon}q^2$.

The canonical measure (38a–38e) and the effective force (39a–39d) can be explicitly evaluated in this example. One would essentially need to evaluate terms like $\left[e^{-t\mathcal{L}_0^\dagger}\cdot \ell_\lambda\right](x) \equiv V_t(x)$ which is a solution of $\frac{d}{dt}V_t(x) = -\mathcal{L}_0^\dagger \cdot V_t(x)$ with an initial condition $V_0(x) = \ell_\lambda(x)$. It is simple to verify that the solution is

$$\left[e^{-t\mathcal{L}_0^\dagger}\cdot \ell_\lambda\right](x) = \exp\left[\frac{\lambda x}{\gamma}e^{\gamma t} + \frac{\lambda^2\epsilon}{4\gamma^3}\left(1-e^{2\gamma t}\right)\right] \qquad \text{for } t \le 0$$

Similarly, one can verify

$$\left[e^{t\mathcal{L}_\lambda}\cdot r_0\right](x) = \mathcal{N}\exp\left[\left(1-e^{-\gamma t}\right)\left\{\frac{\lambda x}{\gamma} - \frac{\epsilon\lambda^2}{4\gamma^3}\left(3-e^{-\gamma t}\right)\right\} + \frac{\epsilon\lambda^2 t}{2\gamma^2} - \frac{\gamma x^2}{\epsilon}\right]$$
$$\text{for } t \ge 0,$$
$$\left[e^{(T-t)\mathcal{L}_\lambda^\dagger}\cdot \ell_0\right](x) = \exp\left[\left(1-e^{-\gamma(T-t)}\right)\left\{\frac{\lambda x}{\gamma} - \frac{\epsilon\lambda^2}{4\gamma^3}\left(3-e^{-\gamma(T-t)}\right)\right\} + \frac{\epsilon\lambda^2(T-t)}{2\gamma^2}\right] \qquad \text{for } t \le T,$$
$$\left[e^{(t-T)\mathcal{L}_0}\cdot r_\lambda\right](x) = \mathcal{N}\exp\left[-\frac{\gamma}{\epsilon}\left(x - \frac{\epsilon\lambda}{2\gamma^2}e^{-\gamma(t-T)}\right)^2\right] \qquad \text{for } t \ge T.$$

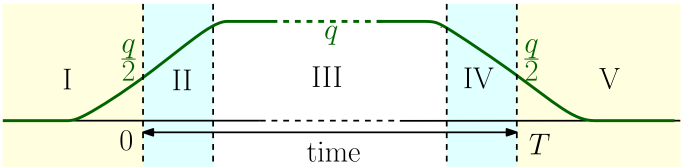

Fig. 2 A schematic of the most probable trajectory, for large T, of the conditioned Ornestein-Uhlenbeck process defined in Sect. 3.3. The most probable position changes with time, only reaching a time independent value $q = \frac{\epsilon\lambda}{\gamma^2}$ at the intermediate quasi-stationary region III. The evolution is symmetric under time reversal, with most probable position $\frac{q}{2}$ at $t = 0$ and $t = T$

Using these in the general expression (38a–38e) and (39a–39d) we find that, in all regions, the canonical measure and the effective force are of the form

$$P_t^{(\lambda)}(x) = \sqrt{\frac{\gamma}{\pi\epsilon}} \exp\left[-\frac{\gamma}{\epsilon}(x - a_t)^2\right] \qquad \text{and} \qquad F_t^{(\lambda)}(x) = -\gamma\,(x - \epsilon\, b_t) \tag{41}$$

This means that the tilted dynamics is another Langevin equation in a harmonic potential whose minimum is at $\epsilon\, b_t$. We get, in region I, $a_t = \frac{\epsilon\lambda}{2\gamma^2} e^{\gamma t}$ and $b_t = \frac{\lambda}{\gamma^2} e^{\gamma t}$; in region II, $a_t = \frac{\epsilon\lambda}{\gamma^2}\left(1 - \frac{1}{2}e^{-\gamma\, t}\right)$ and $b_t = \frac{\lambda}{\gamma^2}$; in region III, $a_t = \frac{\epsilon\lambda}{\gamma^2}$ and $b_t = \frac{\lambda}{\gamma^2}$; in region IV, $a_t = \frac{\epsilon\lambda}{\gamma^2}\left(1 - \frac{1}{2}e^{-\gamma\,(T-t)}\right)$ and $b_t = \frac{\lambda}{\gamma^2}\left(1 - e^{-\gamma(T-t)}\right)$; in region V, $a_t = \frac{\epsilon\lambda}{2\gamma^2} e^{-\gamma\,(t-T)}$ and $b_t = 0$.

For large T, one can get the microcanonical probability $\mathcal{P}_t(x|q)$ using $\frac{\epsilon\lambda}{\gamma^2} = q$ (from $\lambda = \phi'(q)$) in the above expression for $P_t^{(\lambda)}(x)$. From this solution, one can also see that the most likely trajectory followed by the system is $x(t) = a_t$. A schematic of the trajectory is given in Fig. 2.

Remarks 1. In this example, both X_t and Q_T are Gaussian variables. The direct calculation of the covariance is an alternative way of re-deriving (41).

2. Here, the canonical measure $P_t^{(\lambda)}(x)$ is symmetric under $t \to T-t$, thus symmetric under time reversal. This is because on a one-dimensional line the force $F(x)$ can be written as the gradient of a potential and the Langevin dynamics satisfies detailed balance. This would not necessarily be the case on a ring or in higher dimensions.

4 Large deviations in the conditioned Langevin dynamics

We shall now discuss the Langevin dynamics on the infinite line when the noise strength ϵ is small. This weak noise limit has been of interest in the past [50,51,68,69,75–77] particularly in the Freidlin-Wentzel theory of large deviations for stochastic differential equations [19]. One may also view the fluctuating hydrodynamics description of interacting many-body systems as a generalization of the Langevin equation, where the weak noise limit comes from the large system size [12,24,78,79]. A generalization of our discussion here to a many-body system will be presented in a forthcoming publication [52].

In this weak noise limit, one can describe rare fluctuations in terms of a large deviation function [19–21]. For example, the steady state probability of a Langevin equation describing a particle in a potential $U(x)$ has a large deviation form

$$P(x) \sim e^{-\frac{2}{\epsilon}U(x)} \qquad \text{for small } \epsilon.$$

In this Section, we shall show that a similar large deviation description holds, in general, for the canonical measure in the Langevin equation and also for the conditioned probability using the equivalence of ensembles.

4.1 WKB solution of the eigenfunctions

For small ϵ, one can try the WKB method [53] to determine the largest eigenvalue and associated eigenvectors of the tilted operator $\mathcal{L}_\lambda$ in (33). This means that we look for a solution of the type

$$r_{\frac{\kappa}{\epsilon}}(x) \sim e^{-\frac{1}{\epsilon}\psi^{(\kappa)}_{\mathrm{right}}(x)}, \qquad \ell_{\frac{\kappa}{\epsilon}}(x) \sim e^{-\frac{1}{\epsilon}\psi^{(\kappa)}_{\mathrm{left}}(x)} \tag{42a}$$

by setting

$$\lambda = \frac{\kappa}{\epsilon} \qquad \text{and} \qquad \mu\left(\frac{\kappa}{\epsilon}\right) \simeq \frac{1}{\epsilon}\chi(\kappa) \tag{42b}$$

in the eigenvalue equations (35). The scaling (42b) is known [50,51,77] for specific examples of (30), e.g. work and entropy production. We find that, for small ϵ, this is indeed a consistent solution to the leading order when $\psi^{(\kappa)}_{\mathrm{left}}$ and $\psi^{(\kappa)}_{\mathrm{right}}$ satisfy

$$F(x)^2 - \left(\frac{d}{dx}\psi^{(\kappa)}_{\mathrm{left}}(x) - \kappa h(x) - F(x)\right)^2 = 2\kappa f(x) - 2\chi(\kappa) \tag{43a}$$

$$F(x)^2 - \left(\frac{d}{dx}\psi^{(\kappa)}_{\mathrm{right}}(x) + \kappa h(x) + F(x)\right)^2 = 2\kappa f(x) - 2\chi(\kappa) \tag{43b}$$

When we use such a solution in (34) we get

$$G_T^{(\frac{\kappa}{\epsilon})}(x|y) \sim e^{\frac{T}{\epsilon}\chi(\kappa) - \frac{1}{\epsilon}\psi^{(\kappa)}_{\mathrm{right}}(x) - \frac{1}{\epsilon}\psi^{(\kappa)}_{\mathrm{left}}(y)} \tag{44}$$

for small ϵ. This also gives a large deviation form for the canonical measure. In particular, the canonical measure (15) and (19), for small ϵ, gives

$$P_T^{(\frac{\kappa}{\epsilon})}(x) \sim e^{-\frac{1}{\epsilon}\psi^{(\kappa)}_T(x)} \qquad \text{and} \qquad P_0^{(\frac{\kappa}{\epsilon})}(x) \sim e^{-\frac{1}{\epsilon}\psi^{(\kappa)}_0(x)} \tag{45}$$

where $\psi^{(\kappa)}_T(x) = \psi^{(\kappa)}_{\mathrm{right}}(x)$ and $\psi^{(\kappa)}_0(x) = \psi^{(\kappa)}_{\mathrm{left}}(x) + \mathcal{F}(x)$, up to an additive constant [we denote by $\mathcal{F}(x)$ the large deviation function associated to the steady state probability of the original Langevin equation (28)].

Remarks 1. The solution (44) implies that, for large T and small ϵ, the joint probability (13) also has a large deviation form given by

$$P_T(x, Q_T = qT|y) \sim e^{-\frac{T}{\epsilon}\Phi(q) - \frac{1}{\epsilon}\psi_{\mathrm{right}}(x,q) - \frac{1}{\epsilon}\psi_{\mathrm{left}}(y,q)}$$

where $\Phi(q)$, $\psi_{\mathrm{right}}(x,q)$, and $\psi_{\mathrm{left}}(x,q)$ are related to their counterparts $\chi(\kappa)$, $\psi^{(\kappa)}_{\mathrm{right}}(x)$, and $\psi^{(\kappa)}_{\mathrm{left}}(x)$, respectively, by the Legendre transformation (11b), which gives

$$\Phi(q) = \kappa\, q - \chi(\kappa) \qquad \text{with} \quad \chi'(\kappa) = q. \tag{46}$$

This is due to the ensemble equivalence discussed in Sect. 2.2. See [47,49,66] for general conditions for the equivalence of ensembles to hold.

2. Later, in Sect. 6.3, we will see that (43a–43b) are the Hamilton–Jacobi equations in a variational formulation of the problem.

3. The ansatz (42a–42b) is not always applicable, for example, in a double well potential. See [51,80,81] for other recent applications of the WKB analysis for conditioned stochastic processes.

4.2 Large deviation of the canonical measure

The WKB solution (42a) gives that the conditioned probability at any time t, for large T and small ϵ, in the two ensembles, has a large deviation form

$$P_t^{(\frac{\kappa}{\epsilon})}(x) \sim e^{-\frac{1}{\epsilon}\psi_t^{(\kappa)}(x)} \quad \text{and} \quad \mathcal{P}_t(x|Q = qT) \sim e^{-\frac{1}{\epsilon}\psi_t(x,q)} \tag{47}$$

with the two large deviation functions related by the Legendre transformation (46). This is already seen in (45). For other times, this comes from using the WKB solution (42a–42b) in the expressions (38a–38e) for small ϵ.

Among these, the simplest case is the quasi-stationary regime, i.e. $1 \ll t$ and $T - t \gg 1$, where $P_t^{(\lambda)}(x) = r_\lambda(x)\ell_\lambda(x)$ given in (38c). Using (42a) we get

$$\psi_t^{(\kappa)}(x) \equiv \psi_{\text{qs}}^{(\kappa)}(x) = \psi_{\text{right}}^{(\kappa)}(x) + \psi_{\text{left}}^{(\kappa)}(x) \tag{48}$$

(The subscript qs refers to the quasi-stationary state.)

5 Langevin equation on the infinite line

On the infinite line, $F(x)$ is a gradient of a potential $U(x)$, i.e. $F(x) = -\partial_x U(x)$. (For a finite line with a periodic boundary see [46,47,50,51,82].) For simplicity, we consider $Q_T = \int_0^T dt\, f(X_t)$, i.e. $h(x) = 0$ in (30). Moreover, we consider that $F(x)^2 - 2\kappa f(x)$ has a global single minimum [see the remark after (51)] and it grows at $|x| \to \infty$. For small ϵ with (37), this ensures that the spectral gap is non-vanishing.

In this case, the two solutions of the Hamilton–Jacobi equations (43a–43b) are related,

$$\psi_{\text{left}}^{(\kappa)}(x) = \psi_{\text{right}}^{(\kappa)}(x) - 2U(x) + \text{constant} \tag{49}$$

(This would not be true, in general, when $F(x)$ is not a gradient of a potential. For example, on a ring with a circular driving force.)

Moreover, using (49), the effective force (39b) in the quasi-stationary regime, for small ϵ, can be written as

$$F_t^{(\frac{\kappa}{\epsilon})}(x) \simeq F(x) - \partial_x \psi_{\text{left}}^{(\kappa)}(x) = -\frac{1}{2}\partial_x \psi_{\text{qs}}^{(\kappa)}(x) \tag{50}$$

(This is only the leading order term for small ϵ.) This shows that the tilted process can be viewed as a Langevin dynamics in the potential landscape of the large deviation function $\psi_{\text{qs}}^{(\kappa)}(x)$.

An explicit solution

The Hamilton–Jacobi equations (43a–43b) are simple to solve. For example, let's take (43b), which is quadratic and has two solutions $\psi_\pm^{(\kappa)}(x)$ which follows

$$\partial_x \psi_\pm^{(\kappa)}(x) = -F(x) \pm \sqrt{F(x)^2 - 2\kappa f(x) + 2\chi(\kappa)}$$

When $F(x)^2 - 2\kappa f(x)$ has a single global minimum at a value $x = u$ and it grows at $x \to \pm\infty$, the only possible choice is that

$$\partial_x \psi^{(\kappa)}_{\text{right}}(x) = \begin{cases} \partial_x \psi^{(\kappa)}_{+}(x), & \text{for } x \geq u, \\ \partial_x \psi^{(\kappa)}_{-}(x), & \text{for } x \leq u. \end{cases}$$

At the meeting point, the eigenfunction $r_{\frac{\kappa}{\epsilon}}(x)$ and its derivative are continuous which leads to continuity of $\partial_x \psi^{(\kappa)}_{\text{right}}(x)$. The latter condition gives

$$\chi(\kappa) = \kappa f(u) - \frac{1}{2} F(u)^2 \qquad \text{with} \qquad \kappa = \frac{F(u)F'(u)}{f'(u)} \tag{51}$$

Remark The reason for imposing the condition that $F(x)^2 - 2\kappa f(x)$ has a single global minimum is that otherwise, one can not straightforwardly extend the asymptotic solutions $\psi^{(\kappa)}_{\pm}(x)$ to all values of x, similar to the WKB analysis of double well potential in Quantum Mechanics [53]. This is because between the minima the eigenfunction is a superposition of the $\psi^{(\kappa)}_{+}(x)$ and $\psi^{(\kappa)}_{-}(x)$ solutions and one has to carefully match the solutions at each minimum.

The second Hamilton–Jacobi equation (43a) is similarly solved. Integrating these solutions we write

$$\psi^{(\kappa)}_{\text{right}}(x) = \int_{x^\star}^{x} dz \left\{ -F(z) + \text{sgn}(x-u)\sqrt{F(z)^2 - F(u)^2 - 2\kappa[f(z) - f(u)]} \right\} \tag{52a}$$

$$\psi^{(\kappa)}_{\text{left}}(x) = K + \int_{x^\star}^{x} dz \left\{ F(z) + \text{sgn}(x-u)\sqrt{F(z)^2 - F(u)^2 - 2\kappa[f(z) - f(u)]} \right\} \tag{52b}$$

where K and $x^\star$ are a priori arbitrary constants. To satisfy the normalization $\int dx\, r_\lambda(x)\ell_\lambda(x) = 1$, one can choose $K = 0$ for $x^\star = u$ (using $F(x)^2 - 2\kappa f(x)$ has minimum at $x = u$).

Using (52a–52b) in (45) one can see that $\psi^{(\kappa)}_{T}(x)$ and $\psi^{(\kappa)}_{0}(x)$ both have minimum at x_0 given by $f(x_0) = f(u) - \frac{1}{2\kappa} F(u)^2$. This makes x_0 the most likely position at time $t = 0$ and $t = T$, which is different from the quasi-stationary position u.

As a consequence of (52a–52b) we get the large deviation function (48) in the quasi-stationary regime

$$\psi^{(\kappa)}_{\text{qs}}(x) = 2\,\text{sgn}(x-u) \int_{u}^{x} dz\, \sqrt{F(z)^2 - F(u)^2 - 2\kappa[f(z) - f(u)]} \tag{53}$$

This shows that $x = u$ is the most likely position in the quasi-stationary regime.

Remarks 1. In this example, one could systematically calculate sub-leading corrections in the eigenvalue and eigenvector. Writing

$$r_{\frac{\kappa}{\epsilon}}(x) = e^{-\frac{1}{\epsilon}\psi^{(\kappa)}_{\text{right}}(x) - \widetilde{\psi}^{(\kappa)}_{\text{right}}(x) + \cdots}, \qquad \mu\left(\frac{\kappa}{\epsilon}\right) = \frac{1}{\epsilon}\chi(\kappa) + \widetilde{\chi}(\kappa) + \cdots$$

in (35) (we are using $h(x) = 0$) and expanding in powers of ϵ one would get in the sub-leading order

$$-F'(x) + \left[F(x) + \partial_x \psi^{(\kappa)}_{\text{right}}(x)\right] \partial_x \widetilde{\psi}^{(\kappa)}_{\text{right}}(x) - \frac{1}{2}\partial_x^2 \psi^{(\kappa)}_{\text{right}}(x) = \widetilde{\chi}(\kappa) \tag{54}$$

Using (52a) we see that the term $F(x)+\partial_x\psi^{(\kappa)}_{\text{right}}(x)$ in (54) vanishes at $x=u$. Moreover, from (52a) we get

$$\lim_{x\to u}\partial_x^2\psi^{(\kappa)}_{\text{right}}(x)=-F'(u)+\sqrt{F'(u)^2+F(u)F''(u)-\kappa f''(u)}$$

This and the fact that $\partial_x\psi^{(\kappa)}_{\text{right}}(x)=-F(x)$ for $x=u$ gives for the sub-leading order correction to the eigenvalue

$$\widetilde{\chi}(\kappa)=-\frac{1}{2}\left[F'(u)+\sqrt{F'(u)^2+F(u)F''(u)-\kappa f''(u)}\right] \tag{55}$$

An explicit expression for $\widetilde{\psi}^{(\kappa)}_{\text{right}}(x)$ could also be deduced from (52a) and (54).

2. One can also check that the results for the Ornstein–Uhlenbeck process in Sect. 3.3 can be recovered by choosing $f(x)=x$ and $F(x)=-\gamma x$.

6 A variational formulation

In this section, we use the path integral formulation of the Langevin equation [19–21,83]. A similar formulation has been used recently for large deviation of empirical observables [51,69,75,76,82]. This gives an alternative approach for the conditioned dynamics. This conditioned process has been used in [50,51] for specific examples of diffusion on a ring, giving explicit results for the effective dynamics in the quasi-stationary state and revealing dynamical phase transitions.

As in Sect. 5, we consider a Langevin equation on an infinite line and $Q_T=\int_0^T dt\, f(X_t)$, such that $F(x)^2-2\kappa f(x)$ has a single global minimum and it grows at $|x|\to\infty$.

We introduce the formulation for the generating function $G_T^{(\lambda)}(x|y)$ for the Langevin dynamics. Using a path integral solution of (32) (see Appendix D for details) one can write, for small ϵ,

$$G_T^{(\frac{\kappa}{\epsilon})}(x|y)\sim\int_{z(0)=y}^{z(T)=x}\mathcal{D}[z]e^{\frac{1}{\epsilon}S_T^{(\kappa)}[z(t)]} \tag{56}$$

where the Action

$$S_T^{(\kappa)}[z]=\int_0^T dt\left\{\kappa\, f(z)-\frac{\dot{z}^2}{2}+\dot{z}\,F(z)-\frac{F(z)^2}{2}\right\} \tag{57}$$

One may view (56) as a sum over all paths (connecting y to x during time T) weighted by $\exp(\frac{1}{\epsilon}S_T^{(\kappa)}[z])$.

In the small ϵ limit, if we assume that (56) is dominated by a single path, we get (44) with

$$T\chi(\kappa)-\psi^{(\kappa)}_{\text{right}}(x)-\psi^{(\kappa)}_{\text{left}}(y)=\max_{z(t)}S_T^{(\kappa)}[z(t)] \tag{58}$$

where the maximum is over all possible trajectories $z(t)$ with $z(0)=y$ and $z(T)=x$.

6.1 An explicit solution

We will first show how this variational approach allows one to recover the results of Sect. 5. As before, we limit our discussions to the case where $F^2(x)-2\kappa\, f(x)$ has a single global

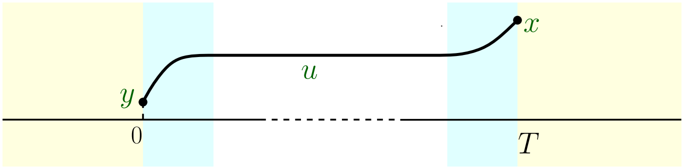

Fig. 3 A schematic of the optimal path for the variational problem in Sect. 6.1

minimum at $x = u$. It will be clear shortly, that in the variational formulation, this condition ensures a single time independent optimal path.

Using variational calculus we get from (56–57) that the optimal path follows

$$\ddot{z} = \frac{d}{dz}\left[\frac{F(z)^2}{2} - \kappa f(z)\right]$$

Multiplying the above equation with $2\dot{z}$ and integrating we get

$$\dot{z}^2 = F(z)^2 - 2\kappa f(z) + K$$

where K is an integration constant. We see the similarity with the trajectory of a mechanical particle of constant energy $\frac{1}{2}K$ in a potential $\kappa f(z) - \frac{F(z)^2}{2}$ which has a single global maximum at $x = u$. The trajectory has to cover a finite distance from the point y to the point x in a very large time T. The only possible way this could happen if the trajectory passes arbitrarily close to u which is a repulsive fixed point of the mechanical dynamics. This requires an energy almost equal to the maximum of the mechanical potential, with the difference vanishing as T grows. This gives $K = 2\kappa f(u) - F(u)^2$ and the optimal path

$$\dot{z}^2 = F(z)^2 - 2\kappa f(z) + 2\kappa f(u) - F(u)^2 \tag{59}$$

Such a trajectory spends most of its time in the position u, and deviates from it only near the boundary to comply with the condition $z(0) = y$ and $z(T) = x$, as sketched in Fig. 3. Then, we can write the optimal path (59), for large T, as

$$\dot{z}(t) = \begin{cases} \text{sgn}(u-y)\sqrt{F(z)^2 - 2\kappa f(z) + 2\kappa f(u) - F(u)^2}, & \text{for } 0 \le t \ll T, \\ 0, & \text{for } 1 \ll t \text{ and } T - t \gg 1, \\ \text{sgn}(x-u)\sqrt{F(z)^2 - 2\kappa f(z) + 2\kappa f(u) - F(u)^2}, & \text{for } 0 \le T - t \ll T. \end{cases}$$

To use this in the variational formula (58) we substitute $F(z)^2$ from (59) in the expression (57) and get

$$\max_{z(t)} S_T^{(\kappa)}[z(t)] = T\left[\kappa\, f(u) - \frac{1}{2}F(u)^2\right] + \int_0^{t_0} dt\, \dot{z}\, [F(z) - \dot{z}] + \int_{t_0}^{T} dt\, \dot{z}\, [F(z) - \dot{z}]$$

where $t_0 \in [0, T]$. We see that, the integration variable can be changed to z, and when $1 \ll t_0$ and $T - t_0 \gg 1$, we can use $z(t_0) = u$, in addition to the boundary condition $z(0) = y$ and $z(T) = x$. Using the explicit solution of $\dot{z}(t)$, given above, we get

$$\max_{z(t)} S_T^{(\kappa)}[z(t)] = T\left[\kappa\, f(u) - \frac{1}{2}F(u)^2\right] - \int_u^y dz\left[F(z) + \text{sgn}(y-u)\sqrt{F(z)^2 - 2\kappa f(z) + 2\kappa f(u) - F(u)^2}\right]$$

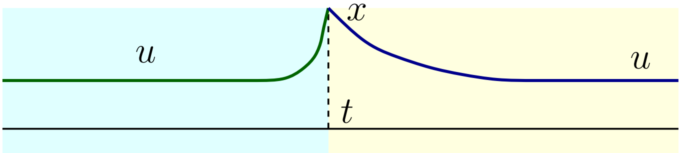

Fig. 4 A schematic of a path leading to a fluctuation x at time t, and subsequent relaxation to the quasi-stationary value u in region III

$$-\int_u^x dz \left[-F(z) + \mathrm{sgn}(x-u)\sqrt{F(z)^2 - 2\kappa f(z) + 2\kappa f(u) - F(u)^2}\right]$$

When we use this result in the variational formula (58) for large T, we get $\chi(\kappa) = \left[\kappa f(u) - \frac{1}{2}F(u)^2\right]$, in agreement with our earlier result in (51). Moreover, we see that the second and third term gives $\psi^{(\kappa)}_{\text{left}}(y)$ and $\psi^{(\kappa)}_{\text{right}}(x)$ in (52a–52b).

6.2 Large deviation function

One could write a similar variational formula for $\psi^{(\kappa)}_t(x)$, defined in (47), at an arbitrary time t. For large T,

$$\psi^{(\kappa)}_t(x) \simeq \max_z A^{(\kappa)}_T[z(\tau)] - \max_{z(t)=x} A^{(\kappa)}_T[z(\tau)] \tag{60a}$$

where the action

$$A^{(\kappa)}_T[z(\tau)] = \int_{-\infty}^{\infty} d\tau \left\{a(\tau)\, f(z) - \frac{\dot{z}^2}{2} + \dot{z}\, F(z) - \frac{F(z)^2}{2}\right\} \tag{60b}$$

with $a(\tau) = \kappa$ for $\tau \in [0, T]$ and $a(\tau) = 0$ elsewhere. The first maximization in (60a) is over all paths, whereas the second maximization is over paths, which are conditioned to be at $z(\tau) = x$ for $\tau = t$.

One may understand the formula (60a) as an optimal contribution from an ensemble of paths with probability weight $e^{\frac{1}{\epsilon}A^{(\kappa)}_T[z]}$ conditioned to pass through x at time t; the first term in (60a) is due to normalization.

Here, we show how one can use this variational approach to derive $\psi^{(\kappa)}_t(x)$ at an arbitrary time. For this we impose as in Sect. 6.1 that $F(x)^2 - 2\kappa\, f(x)$ has a single global minimum such that the most likely position in the quasi-stationary regime is time independent, $z(\tau) = u$.

Quasi-stationary regime

Among all the five regions in Fig. 1, the simplest is to analyze the quasi-stationary regime where $1 \ll t$ and $T - t \gg 1$. Here, for the optimization in (60a), one essentially needs to consider paths which asymptotically reach u, both at small t, as well as when t is close to T. A schematic such path is given in Fig. 4.

The analysis is quite similar to that in Sect. 6.1. We get that the optimal path follows

$$\frac{dz(\tau)}{d\tau} = \begin{cases} \mathrm{sgn}(x-u)\sqrt{F(z)^2 - 2\kappa f(z) + 2\kappa f(u) - F(u)^2}, & \text{for } \tau < t \\ \mathrm{sgn}(u-x)\sqrt{F(z)^2 - 2\kappa f(z) + 2\kappa f(u) - F(u)^2}, & \text{for} \tau > t, \end{cases} \tag{61}$$

and using this in (60a) we get

$$\psi_t^{(\kappa)}(x) = \int_0^t d\tau\, \dot z\,[\dot z - F(z)] + \int_t^T d\tau\, \dot z\,[\dot z - F(z)]$$

Changing the integration variable to z and using the solution (61) with the asymptotics sketched in Fig. 4, we get

$$\psi_t^{(\kappa)}(x) = \int_u^x dz\left[-F(z) + \text{sgn}(x-u)\sqrt{F(z)^2 - 2\kappa f(z) + 2\kappa f(u) - F(u)^2}\right] + \int_u^x dz\left[F(z) + \text{sgn}(x-u)\sqrt{F(z)^2 - 2\kappa f(z) + 2\kappa f(u) - F(u)^2}\right]$$

Comparing with the expression in (52a–52b) we see that $\psi_t^{(\kappa)}(x) = \psi_{\text{right}}^{(\kappa)}(x) + \psi_{\text{left}}^{(\kappa)}(x)$, in agreement with our earlier result (48) and (53).

Remark From (61) one could see that the optimal path leading to a fluctuation in the quasi-stationary regime and subsequent relaxation follows a deterministic evolution in the potential landscape of $\psi_{\text{right}}^{(\kappa)}$ and $\psi_{\text{left}}^{(\kappa)}$.

$$\frac{dz(\tau)}{d\tau} = F(z) + \frac{d}{dz}\psi_{\text{right}}^{(\kappa)}(z) = -\frac{d}{dz}\left[U(z) - \psi_{\text{right}}^{(\kappa)}(z)\right] \quad \text{for } \tau < t, \tag{62a}$$

$$\frac{dz(\tau)}{d\tau} = F(z) - \frac{d}{dz}\psi_{\text{left}}^{(\kappa)}(z) = -\frac{d}{dz}\left[U(z) + \psi_{\text{left}}^{(\kappa)}(z)\right] \quad \text{for } \tau > t. \tag{62b}$$

Region II ($0 \le t \ll T$)

The calculation of $\psi_t^{(\kappa)}(x)$ in other regions of time is quite similar. For example, in region II, in the variational formula (60a), one essentially needs to consider paths which started at the minimum of $U(x)$ (with $F(x) = -U'(x)$) when $\tau \to -\infty$, pass through $z = x$ at $\tau = t \ge 0$, and asymptotically reach the quasi-stationary value u for large time $\tau \gg 1$, as illustrated in Fig. 5.

Following an analysis similar to that in Sect. 6.1 it is straightforward to show that the optimal path in this case

$$\dot z(\tau) = \begin{cases} -F(z), & \text{for } \tau \le 0 \\ \text{sgn}(x-y)\sqrt{F(z)^2 - 2\kappa f(z) + K_1}, & \text{for } 0 \le \tau \le t \\ \text{sgn}(u-x)\sqrt{F(z)^2 - 2\kappa f(z) + K_2}, & \text{for } \tau \ge t, \end{cases} \tag{63}$$

where K_1 and K_2 are integration constants, and the optimal path passes through $z(0) = y$ (say) when $\tau = 0$. The solution for $\tau \le 0$ is easy to see from the condition that at $\tau \to -\infty$ the system started at the minimum of the potential $U(z)$ with $F(z) = -U'(z)$. Similar asymptotics that for large time the system relaxes to the quasi-stationary position $z = u$ gives the constant $K_2 = 2\kappa f(u) - F(u)^2$. In addition, we have the condition

$$t = \int_0^t d\tau = \int_y^x \frac{dz}{\dot z} = \int_y^x \frac{dz}{\text{sgn}(x-y)\sqrt{F(z)^2 - 2\kappa f(z) + K_1}} \tag{64}$$

where we used the solution (63) and this fixes the constant K_1.

When we use the solution (63) to write $F(z)^2$ in the expression (60b), we get

$$\max_{z(t)=x} A_T^{(\kappa)}[z(\tau)] = (T-t)\left[\kappa f(u) - \frac{F(u)^2}{2}\right] + t\,\frac{K_1}{2} - \int_{-\infty}^T d\tau\, \dot z\left[\dot z - F(z)\right]$$

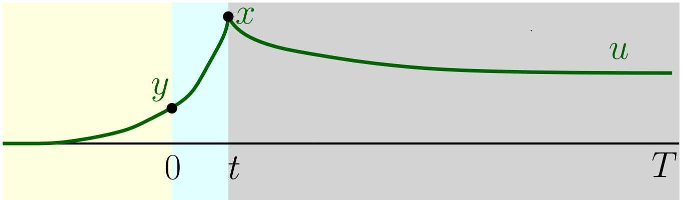

Fig. 5 A schematic of a path leading to a fluctuation x at t in region II, and subsequent relaxation to the quasistationary position u

Using this in (60a) and the result that $\max_{z(t)} A_T^{(\kappa)}[z(\tau)] = T\left[\kappa f(u) - \frac{F(u)^2}{2}\right]$, we get

$$\psi_t^{(\kappa)}(x) = t\left[\kappa f(u) - \frac{F(u)^2}{2}\right] - t\,\frac{K_1}{2} + \int_{-\infty}^{T} d\tau\, \dot{z}\big[\dot{z} - F(z)\big]$$

In this expression, the integration variable can be changed from τ to z, and then using the explicit solution (63), we get

$$\psi_t^{(\kappa)}(x) = t\left[\kappa f(u) - \frac{F(u)^2}{2}\right] + \psi_{\text{left}}^{(\kappa)}(x) + \widehat{B}_t^{(\kappa)}(x, y) + \mathcal{F}(y) \tag{65a}$$

where $\psi_{\text{left}}^{(\kappa)}(x)$ is given in (52b), $\mathcal{F}(y) = -2\int_0^y dz\, F(z)$ and

$$\widehat{B}_t^{(\kappa)}(x, y) = -t\,\frac{K_1}{2} + \int_y^x dz\left[-F(z) + \operatorname{sgn}(x-y)\sqrt{F(z)^2 - 2\kappa f(z) + K_1}\right] \tag{65b}$$

We note that the condition (64) is equivalent to $\partial_{K_1}\widehat{B}_t^{(\kappa)}(x, y) = 0$, which relates K_1 to y. In addition, the solution (65a) must be optimal over a variation in y. These two conditions together leads to $\partial_y \widehat{B}_t^{(\kappa)}(x, y) = 2F(y)$, which with the formula (65b) gives $K_1 = 2\kappa\, f(y)$. We note that this is equivalent of continuity of $\dot{z}(\tau)$ at $\tau = 0$ in the solution (63). This result for K_1, along with (64) and (65a–65b) gives a parametric solution of $\psi_t^{(\kappa)}(x)$ in region II.

We have checked that the same result could be derived using the eigenfunction of the tilted Fokker–Planck operator discussed earlier in Sect. 4.

6.3 The Hamilton–Jacobi equations from the variational approach

In Sect. 4 we have shown how one can write the large deviation function in terms of a solution of the Hamilton–Jacobi equations (43a, 43b) derived from the tilted Fokker–Planck operator. In this section, we describe how the same equations can be obtained using the variational formulation in (58). The advantage is that this variational approach can be extended to more general problems (see our future publication [52], where the approach becomes a generalization of the one by Bertini et al. [79]).

We start with a derivation of (43a). Using the definition (6) one can write for the Langevin equation

$$G_T^{(\lambda)}(x|y) = \int dz\, G_{T-t}^{(\lambda)}(x|z) G_t^{(\lambda)}(z|y) \tag{66}$$

A schematic illustrating this integration is shown in Fig. 6.

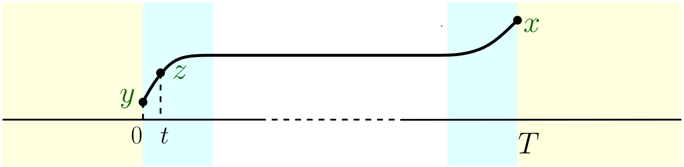

Fig. 6 A schematic of the sample of paths contributing in the time convolution in (66)

Using the large deviation form (44) and the path integral representation (56), for small ϵ, it is straightforward to write

$$t\chi(\kappa) - \psi_{\text{left}}^{(\kappa)}(y) \simeq \max_z \left\{ S_t^{(\kappa)}(z, y) - \psi_{\text{left}}^{(\kappa)}(z) \right\} \tag{67}$$

where, from the Action (57), we get for small t,

$$S_t^{(\kappa)}(z, y) = t\,\kappa\, f(y) - \frac{1}{2}\left[\frac{(z-y)}{t} - F(y)\right]^2 + \cdots$$

Expanding (67) around y we get

$$\chi(\kappa) \simeq \kappa\; f(y) - \frac{F(y)^2}{2} + \frac{1}{t}\max_z \left\{ (z-y)[F(y) - \partial_y \psi_{\text{left}}^{(\kappa)}(y)] - \frac{(z-y)^2}{2\,t} \right\}$$

Higher order terms in the expansion are negligible in the small t limit.

In this expression, the maximum is for

$$\frac{(z-y)}{t} = F(y) - \partial_y \psi_{\text{left}}^{(\kappa)}(y)$$

Substituting this in the above expression for $\chi(\kappa)$ and taking $t \to 0$ limit we recover the Hamilton–Jacobi equation (43a) for $h(x) = 0$. One can similarly derive the Hamilton–Jacobi equation (43b) for $\psi_{\text{right}}^{(\kappa)}(x)$. The analysis could be extended for $h(x) \neq 0$, as well.

7 The effect of conditioning on the noise

In Sect. 3, we found (shown earlier in [44–47]) that the Langevin dynamics (28) conditioned on Q_T, for large T, can be effectively described by another Langevin equation with a modified force (39a–39d), and still a white noise. Here, we show that the noise realizations in the original Langevin equation (28), which are compatible with the condition on (30), are colored.

The effective dynamics

In (39a–39d) we have seen that the Langevin dynamics conditioned on Q_T in (30) can be described, in the large T limit, by another Langevin dynamics with an effective force $F_t^{(\lambda)}(x)$ and a Gaussian white noise $\widetilde{\eta}_t$ with mean zero and covariance $\langle \widetilde{\eta}_t \widetilde{\eta}_{t'} \rangle = \epsilon\,\delta(t - t')$. In the weak noise limit, the effective force in the quasi-stationary regime ($t \gg 1$ and $T - t \gg 1$) is given by (50) with (53), when $h(x) = 0$ in (30). So the effective dynamics, for large T and small ϵ, is

$$\dot{X}_t = -\text{sgn}(X_t - u)\sqrt{F(X_t)^2 - F(u)^2 - 2\kappa[f(X_t) - f(u)]} + \widetilde{\eta}_t. \tag{68}$$

In this quasi-stationary regime, the most probable position $X_t = u$ is time independent (under the condition that $F(x)^2 - 2\kappa f(x)$ has a single global minimum at $x = u$). Writing small fluctuations $r_t = X_t - u$ around u, we get from (68)

$$\dot{r}_t = -\Gamma_u r_t + \widetilde{\eta}_t, \qquad \text{with} \qquad \Gamma_u = \sqrt{F'(u)^2 + F(u)F''(u) - \kappa f''(u)}.$$

The solution

$$r_t = \int_{-\infty}^{t} dt'\, e^{-\Gamma_u (t-t')}\, \widetilde{\eta}_{t'}$$

leads to the following correlation

$$\langle X_t X_{t'} \rangle_c = \langle r_t r_{t'} \rangle = \frac{\epsilon}{2\Gamma_u} e^{-\Gamma_u |t-t'|}. \tag{69}$$

So, for small ϵ, a typical trajectory of the effective dynamics in the quasi-stationary state has small fluctuations around $X_t = u$ with correlation (69).

The conditioned dynamics

If we come back to the original Langevin equation (28),

$$\dot{Y}_t = F(Y_t) + \eta_t \tag{70}$$

then, η_t is a priori delta-correlated in time. We are now going to show that conditioning on a value of Q_T for large T, induces correlations of the noise η_t. To do so we use the fact that at least for small ϵ the trajectories of the dynamics (70) when conditioned on Q_T for large T are the same as for the effective dynamics (68) and therefore

$$\langle Y_t | Q_T \rangle = \langle X_t \rangle \quad \text{and} \quad \langle Y_t Y_{t'} | Q_T \rangle_c = \langle X_t X_{t'} \rangle_c \tag{71}$$

Small fluctuations $s_t = Y_t - u$ in the quasi-stationary regime are generated by a noise realization η_t in (70) given by

$$\eta_t \simeq -F(u) + \dot{s}_t - F'(u)\, s_t \tag{72}$$

Then, using (69), (71), and (72) one gets

$$\langle Y_t\, \eta_{t'} | Q_T \rangle_c = \langle s_t\, \eta_{t'} | Q_T \rangle = \begin{cases} g_R(t'-t), & \text{for } t' > t \\ g_F(t-t'), & \text{for } t' < t \end{cases} \tag{73a}$$

where

$$g_F(t) = \frac{-F'(u) + \Gamma_u}{2\Gamma_u} e^{-\Gamma_u t} \qquad \text{and} \qquad g_R(t) = \frac{-F'(u) - \Gamma_u}{2\Gamma_u} e^{-\Gamma_u t} \tag{73b}$$

In this description (73a), we see that the fluctuation s_t is correlated not only to the noise in the past, but also to the noise in the future. Of course, when one removes the conditioning, i.e. for $\kappa = 0$, and using $F(0) = 0$ (assuming $x = 0$ is the stable fixed point for the unconditioned case), one has $\Gamma_0 = -F'(0)$ and $g_R = 0$, as one would expect in a Markovian process. One can also see, using (69), (71), and (72) that

$$\langle \eta_t | Q_T \rangle = -F(u) \qquad \text{and} \qquad \langle \eta_t \eta_{t'} | Q_T \rangle_c = \epsilon\, \frac{F'(u)^2 - \Gamma_u^2}{2\Gamma_u}\, e^{-\Gamma_u |t-t'|} \tag{74}$$

This means that in the conditioned ensemble, the original white noise η_t in the Langevin equation (70) becomes colored due to the conditioning on Q_T, even in the large T limit.

8 Summary

In this work we studied how a stochastic system adapts its dynamics when it is conditioned on a certain value of an empirical observable Q_T of the form (4). This problem has been studied earlier in [32,44–47]. The constrained dynamics in the large T limit is described by an effective Markov process [see (21, 26)] if the original process is itself Markovian. In the case of the Langevin dynamics, the conditioning modifies the effective force [see (39a–39d)]. The description in terms of the effective dynamics in the large T limit comes from an equivalence of ensembles between the microcanonical ensemble [where conditioning is on a fixed value of Q_T, defined in (4) and (30)] and the canonical ensemble (where the dynamics is weighted by $e^{\lambda Q_T}$). This is similar to the equivalence of thermodynamic ensembles in equilibrium when volume is large. The equivalence of ensembles and several of the expressions obtained in Sects. 2 and 3 were already known [32,44–48,59], mostly in the quasi-stationary regime. Here, we extend them to all regions of time.

In the weak noise limit of the Langevin dynamics, one can introduce large deviation functions which characterize fluctuations in the conditioned dynamics, for large T. Using a WKB solution we showed in Sect. 4.1 that these large deviation functions can be expressed in terms of the solution of the Hamilton–Jacobi equations (43a–43b). The same result can also be derived (see Sect. 6) using a variational formulation, where the large deviation functions are related to the minimum of the Action that characterizes the path-space probability. Within this variational approach, one can calculate the optimal trajectory, which describes how atypical fluctuations are generated and how they relax (61, 63). A similar approach to our variational formulation was used recently [50,51] in the quasi-stationary regime of a Langevin dynamics in a periodic potential.

One of the rather surprising aspects in the Langevin dynamics (28) is that the noise realizations, which are compatible with the condition on Q_T in (30) become correlated over time [see (74)]. Moreover, fluctuations of the position at a time become correlated to the noise in the future.

The examples discussed in this paper are simple as they deal with a single degree of freedom. They are part of a theory which is rather general. In a forthcoming publication [52] we shall apply the same ideas for a system with many degrees of freedom [12,24,30], e.g. the symmetric exclusion process. The variational approach discussed here for the Langevin dynamics can be generalized for large systems where the weak noise limit comes from the large volume. Several of the ideas used in this paper will be extended there.

We have seen in (18c) and (38c) that in the quasi-stationary regime the canonical measure is a product of the left and right eigenvectors corresponding to the largest eigenvalue of the tilted matrix. Even in the non-stationary regime [see (25a)] the canonical measure is a product of a left vector and a right vector, which evolve according to linear equations. This is reminiscent of Quantum Mechanics, where probability is expressed as a product of the wave function, as already noted by Schrödinger [84] (see also [85,86] for additional references).

Acknowledgements We acknowledge the hospitality of ICTS-Bengaluru, India, where part of the work was completed during a workshop on Large deviation theory in August, 2017.

A Ensemble equivalence

In this appendix we show that, for large T, the equivalence of ensembles holds for an arbitrary time t. For an earlier derivation of the ensemble equivalence see [47].

As the reasoning is very similar in the five regions of Fig. 1, we will limit our discussion to the case of region II, i.e. for $0 \le t \ll T$. Let $P_t(C_T, C, Q|C_0)$ be the joint probability of configuration C_T at time T, configuration C at time t, and of the observable Q_T to take value Q given its initial configuration C_0 at time 0.

To establish the equivalence of ensembles in (17), we need to show that the conditioned probability in the microcanonical ensemble

$$\mathcal{P}_t(C|Q = qT) = \frac{\sum_{C_T} \sum_{C_0} P_t(C_T, C, Q = qT|C_0) R_0(C_0)}{\sum_{C'} \left[\sum_{C_T} \sum_{C_0} P_t(C_T, C', Q = qT|C_0) R_0(C_0)\right]} \tag{75}$$

and the canonical measure

$$P_t^{(\lambda)}(C) = \frac{\sum_{C_T} \sum_{C_0} \int dQ\, e^{\lambda Q} P_t(C_T, C, Q|C_0) R_0(C_0)}{\sum_{C'} \left[\sum_{C_T} \sum_{C_0} \int dQ\, e^{\lambda Q} P_t(C_T, C', Q|C_0) R_0(C_0)\right]} \tag{76}$$

converge to the same distribution for large T when λ and q are related by (12).

For this, we write, in terms of the probability (5),

$$P_t(C_T, C, Q = qT|C_0) = \int dQ_t\ P_{T-t}(C_T, qT - Q_t|C)\ P_t(C, Q_t|C_0) \tag{77}$$

and use the large T asymptotics (13), which gives

$$P_{T-t}(C_T, qT - Q_t|C) \simeq e^{-(T-t)\phi(q)-(tq-Q_t)\phi'(q)} \sqrt{\frac{\phi''(q)}{2\pi T}} R_{\phi'(q)}(C_T) L_{\phi'(q)}(C)$$

Substituting in (75) and simplifying the expression for large T we get the microcanonical probability

$$P_t(C|Q = qT) \simeq \frac{L_{\phi'(q)}(C) \sum_{C_0} G_t^{(\phi'(q))}(C|C_0)\ R_0(C_0)}{\sum_{C'} L_{\phi'(q)}(C') \sum_{C_0} G_t^{(\phi'(q))}(C'|C_0)\ R_0(C_0)} \tag{78}$$

where $G_t^{(\lambda)}(C|C_0)$ is defined in (6). On the other hand, using (9) for large T we get the canonical probability

$$P_t^{(\lambda)}(C) \simeq \frac{L_\lambda(C) \sum_{C_0} G_t^{(\lambda)}(C|C_0)\ R_0(C_0)}{\sum_{C'} L_\lambda(C') \sum_{C_0} G_t^{(\lambda)}(C'|C_0)\ R_0(C_0)} \tag{79}$$

Clearly the two probabilities in the two ensembles coincide for $\lambda = \phi'(q)$. Replacing $G_t^{(\lambda)}(C|C_0)$ by $M_\lambda^t(C, C_0)$ in (79) leads to the canonical measure (18b).

The same reasoning can be easily adapted in the other regions of Fig. 1.

B Continuous time Markov process

In this Appendix, we describe a continuous time limit of the Markov process, illustrated in Fig. 7. In this, the empirical observable analogous to (4) is the $dt \to 0$ limit of

$$Q_T = dt \sum_{i=0}^{\frac{T}{dt}-1} f(C_i) + \sum_n g(C_n^+, C_n^-) \tag{80}$$

where $t = i\, dt$, and (C_n^-, C_n^+) are the configurations before and after the nth jump during the time interval $[0, T]$.

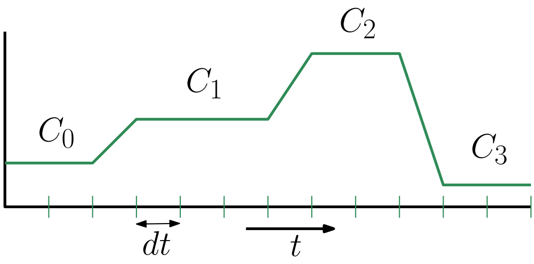

Fig. 7 A schematic of a time evolution in a Markov process with discrete time steps dt. The continuous time limit is obtained by taking $dt \to 0$ limit

From (7) we get

$$G_T^{(\lambda)}(C'|C_0) = \sum_C M_\lambda(C', C)\, G_{T-dt}^{(\lambda)}(C|C_0)$$

where

$$M_\lambda(C', C) = \begin{cases} M_0(C', C)e^{\lambda[dt\, f(C)+g(C',C)]} & \text{for } C' \neq C, \\ M_0(C, C)e^{\lambda\, dt\, f(C)} & \text{for } C' = C. \end{cases}$$

Using the construction (27) for $M_0(C', C)$ we take the continuous time limit $dt \to 0$ and get

$$\frac{d}{dT} G_T^{(\lambda)}(C'|C_0) = \sum_C \mathcal{M}_\lambda(C', C) G_T^{(\lambda)}(C|C_0) \tag{81}$$

where we recover an earlier result [6,47] for the tilted matrix $\mathcal{M}_\lambda$ for the continuous time process, given by

$$\mathcal{M}_\lambda(C', C) = \begin{cases} e^{\lambda g(C',C)}\mathcal{M}_0(C', C) & \text{for } C' \neq C, \\ \lambda f(C) - \sum_{C'' \neq C} \mathcal{M}_0(C'', C) & \text{for } C' = C. \end{cases} \tag{82}$$

This shows that the generating function is the (C, C_0)th element of $e^{T\mathcal{M}_\lambda}$, i.e.

$$G_T^{(\lambda)}(C|C_0) = e^{T\mathcal{M}_\lambda}(C, C_0) \tag{83}$$

Although (81) resembles a Master equation, the tilted matrix $\mathcal{M}_\lambda$ is not a Markov matrix as $\sum_{C'} \mathcal{M}_\lambda(C', C)$ does not necessarily vanish.

For large T, one would get $G_T^{(\lambda)}(C|C_0) \simeq e^{T\mu(\lambda)} R_\lambda(C) L_\lambda(C_0)$ where the cumulant generating function $\mu(\lambda)$ is the largest eigenvalue of $\mathcal{M}_\lambda$ with $L_\lambda(C)$ and $R_\lambda(C)$ being the left and right eigenvectors, respectively. (Note the difference with the discrete time case (9), where $\mu(\lambda)$ is the *logarithm* of the largest eigenvalue of the tilted matrix M_λ in (8).)

In a similar construction, one could get the continuous time limit of the canonical measure (18a–18d) and its time evolution (20a–20d). The analysis is straightforward and we present only the final result.

The time evolution of the canonical measure $P_t^{(\lambda)}(C)$ for a continuous time Markov process is also a Markov process

$$\frac{d}{dt} P_t^{(\lambda)}(C') = \sum_C \mathcal{W}_t^{(\lambda)}(C', C) P_t^{(\lambda)}(C) \tag{84}$$

where $\mathcal{W}_t^{(\lambda)}(C', C)$ is the transition rate from C to C' at time t in the canonical ensemble. The canonical measure and transition rate have different expressions in the five regions indicated in Fig. 1. Their expression is given below, where we use a matrix product notation, e.g. $[L_\lambda \mathcal{M}_\lambda](C) \equiv \sum_{C'} L_\lambda(C')\mathcal{M}_\lambda(C', C)$.

1. Region I.

$$P_t^{(\lambda)}(C) = \frac{[L_\lambda e^{-t\mathcal{M}_0}](C) R_0(C)}{\sum_{C'} L_\lambda(C') R_0(C')} \tag{85a}$$

$$\mathcal{W}_t^{(\lambda)}(C', C) = \frac{[L_\lambda e^{-t\mathcal{M}_0}](C')}{[L_\lambda e^{-t\mathcal{M}_0}](C)} \mathcal{M}_0(C', C) - \frac{\left[L_\lambda e^{-t\mathcal{M}_0}\mathcal{M}_0\right](C)}{[L_\lambda e^{-t\mathcal{M}_0}](C)}\, \delta_{C',C} \tag{85b}$$

2. Region II.

$$P_t^{(\lambda)}(C) = \frac{L_\lambda(C)[e^{t\mathcal{M}_\lambda} R_0](C)}{e^{t\mu(\lambda)} \sum_{C'} L_\lambda(C') R_0(C')} \tag{86a}$$

$$\mathcal{W}_t^{(\lambda)}(C', C) = \frac{L_\lambda(C')}{L_\lambda(C)} \mathcal{M}_\lambda(C', C) - \mu(\lambda)\delta_{C',C} \tag{86b}$$

3. Region III.

$$P_t^{(\lambda)}(C) = L_\lambda(C) R_\lambda(C) \tag{87a}$$

$$\mathcal{W}_t^{(\lambda)}(C', C) = \frac{L_\lambda(C')}{L_\lambda(C)} \mathcal{M}_\lambda(C', C) - \mu(\lambda)\delta_{C',C} \tag{87b}$$

4. Region IV.

$$P_t^{(\lambda)}(C) = \frac{[L_0\, e^{(T-t)\mathcal{M}_\lambda}](C) R_\lambda(C)}{e^{(T-t)\mu(\lambda)} \sum_{C'} R_\lambda(C')} \tag{88a}$$

$$\mathcal{W}_t^{(\lambda)}(C', C) = \frac{[L_0\, e^{(T-t)\mathcal{M}_\lambda}](C')}{[L_0\, e^{(T-t)\mathcal{M}_\lambda}](C)} \mathcal{M}_\lambda(C', C) - \frac{\left[L_0\, e^{(T-t)\mathcal{M}_\lambda}\mathcal{M}_\lambda\right](C)}{\left[L_0\, e^{(T-t)\mathcal{M}_\lambda}\right](C)} \delta_{C',C} \tag{88b}$$

   where the left eigenvector L_0 for the original (unconditioned) evolution is a unit vector such that $[L_0\, \mathcal{M}_\lambda](C) \equiv \sum_{C'} \mathcal{M}_\lambda(C', C)$.
5. Region V.

$$P_t^{(\lambda)}(C) = \frac{[e^{(t-T)\mathcal{M}_0} R_\lambda](C)}{\sum_{C'} R_\lambda(C')} \tag{89a}$$

$$\mathcal{W}_t^{(\lambda)}(C', C) = \mathcal{M}_0(C', C) \tag{89b}$$

These expressions of $P_t^{(\lambda)}$ and $\mathcal{W}_t^{(\lambda)}$, particularly (86a) and (87), have been derived earlier in [47]. The results for $\mathcal{W}_t^{(\lambda)}$ can be viewed as a generalization of the Doob's h-transformation [42,47].

One can verify the property $\sum_{C'} \mathcal{W}_t^{(\lambda)}(C', C) = 0$ in all five regions. Moreover, setting $\lambda = 0$, and $L_0(C) = 1$, gives $\mathcal{W}_t^{(0)}(C', C) = \mathcal{M}_0(C', C)$, as one would expect.

C Langevin dynamics as a limit of a Markov process

In this appendix, we show how the case of Langevin dynamics in Sect. 3 can be obtained as a continuous limit of the discrete time Markov process in Sect. 2.

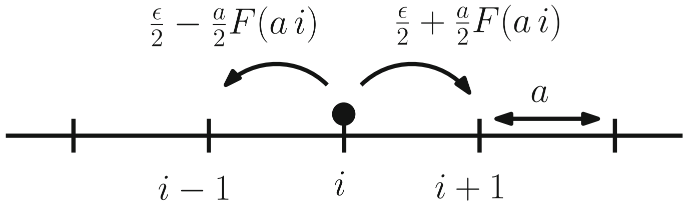

Fig. 8 A jump process on a one-dimensional chain where a particle jumps to its nearest neighbour site with rates indicated in the figure

Let us consider a jump process on a one-dimensional lattice where a configuration C is given by the site index i as indicated in Fig. 8. Only nearest neighbor jumps are allowed with transition rates that we take of the form

$$M_0(i \pm 1, i) = \frac{\epsilon}{2} \pm \frac{a}{2} F(a\,i) \tag{90}$$

with $M_0(i,i) = 1 - \epsilon$, where a is the unit lattice spacing, $\epsilon < 1$ is a fixed parameter, and $F(x)$ is an arbitrary function defined on the lattice.

The probability $P_{t,i}$ of the jump process to be in site i at time t satisfies the Master equation

$$P_{t+1,i} = M_0(i, i+1) P_{t,i+1} + M_0(i, i-1) P_{t,i-1} + M_0(i,i) P_{t,i} \tag{91}$$

Taking the limit $a \to 0$, keeping ϵ arbitrary, one can easily see that $P_{a^2 t}(a\,i) \equiv P_{t,i}$ follows the Fokker–Planck equation (29). This shows that the continuous limit of the jump process is indeed identical to the Langevin dynamics (28).

One can similarly obtain the tilted Langevin dynamics from the continuous limit of the jump process when weighted by $e^{\lambda Q}$ with the observable Q in (24). For this we define

$$f_t(i) = a^2 f(a\,i, a^2 t) \qquad \text{and} \qquad g_t(j,i) = (j-i)\,a\,\left\{\alpha\, h(a\,j, a^2 t) + (1-\alpha)\, h(a\,i, a^2 t)\right\}$$

where α is the parameter, which specifies the prescription (Îto or Stratonovich) as in (31). Then, the continuous limit of (24) corresponds to an observable Q of the Langevin dynamics

$$Q = \int dt\; f(X_t, t) + \int dX_t\; h(X_t, t) \tag{92}$$

In the expression (25a) for the canonical measure if we define $H^{(\lambda)}_{a^2 t}(a\,i) \equiv Z^{(\lambda)}_t(i)$ and $\mathbb{H}^{(\lambda)}_{a^2 t}(a\,i) \equiv \mathbb{Z}^{(\lambda)}_t(i)$, then in the continuous limit $a \to 0$ we get the canonical measure for the Langevin dynamics weighted by $e^{\lambda Q}$ with Q in (92):

$$P^{(\lambda)}_t(x) = \frac{H^{(\lambda)}_t(x)\, \mathbb{H}^{(\lambda)}_t(x)}{\int dy\, H^{(\lambda)}_t(y)\, \mathbb{H}^{(\lambda)}_t(y)} \tag{93}$$

The time evolution of $H^{(\lambda)}_t(x)$ and $\mathbb{H}^{(\lambda)}_t(x)$ are obtained from (25b–25c) for the jump process by taking the $a \to 0$ limit, keeping ϵ fixed. We get

$$\frac{d}{dt} H^{(\lambda)}_t(x) = \lambda f(x,t) H^{(\lambda)}_t(x) - \left(\frac{d}{dx} - \lambda\, h(x,t)\right) F(x) H^{(\lambda)}_t(x) + \frac{\epsilon}{2}\left(\frac{d^2}{dx^2} H^{(\lambda)}_t(x) \right.$$
$$\left. - 2\lambda\, h(x,t) \frac{d}{dx} H^{(\lambda)}_t(x) - 2(1-\alpha)\lambda\, \partial_x h(x,t) H^{(\lambda)}_t(x) + \lambda^2 h(x,t)^2 H^{(\lambda)}_t(x)\right) \tag{94a}$$

$$-\frac{d}{dt}\mathbb{H}_t^{(\lambda)}(x) = \lambda f(x,t)\mathbb{H}_t^{(\lambda)}(x) + F(x)\left(\frac{d}{dx} + \lambda\, h(x,t)\right)\mathbb{H}_t^{(\lambda)}(x) + \frac{\epsilon}{2}\left(\frac{d^2}{dx^2}\mathbb{H}_t^{(\lambda)}(x) + 2\lambda\, h(x,t)\frac{d}{dx}\mathbb{H}_t^{(\lambda)}(x) + 2\alpha\lambda\, \partial_x h(x,t)\mathbb{H}_t^{(\lambda)}(x) + \lambda^2 h(x,t)^2\mathbb{H}_t(x)\right) \tag{94b}$$

Similarly, the continuous limit of (21, 26) gives the Fokker–Planck equation

$$\frac{d}{dt}P_t^{(\lambda)}(x) = -\frac{d}{dx}\left[F_t^{(\lambda)}(x)P_t^{(\lambda)}(x)\right] + \frac{\epsilon}{2}\frac{d^2}{dx^2}P_t^{(\lambda)}(x) \tag{95a}$$

where the modified force

$$F_t^{(\lambda)}(x) = F(x) + \epsilon\left(\lambda\, h(x,t) + \frac{d}{dx}\log\mathbb{H}_t^{(\lambda)}(x)\right) \tag{95b}$$

This gives the time evolution of the Langevin dynamics when it is weighted by the observable (92).

Remarks 1. In the derivation of (95a) we have used that the denominator in (93) is time independent, which can be checked using (94a, 94b).

2. The Fokker–Planck equation (95a) shows that the effect of biasing a Langevin dynamics by $e^{\lambda Q}$ with an arbitrary time dependent observable (92) is described by another Langevin dynamics with a modified force (95b), but the noise strength ϵ remains unchanged. This works even without a large parameter T (see [38,49] for earlier examples of conditioned dynamics).

Our results in Sect. 3 belongs to a particular case, where the observable (92) is defined in a large time interval $[0, T]$. This corresponds to [see (30)]

$$f(x,t) = \begin{cases} f(x) \text{ for } t \in [0,T], \\ 0 \text{ otherwise,} \end{cases} \qquad \text{and} \qquad h(x,t) = \begin{cases} h(x) \text{ for } t \in [0,T], \\ 0 \text{ otherwise.} \end{cases}$$

In this case, (94a, 94b) gives

$$\frac{d}{dt}H_t^{(\lambda)}(x) = \mathcal{L}(t)\cdot H_t^{(\lambda)}(x), \qquad \frac{d}{dt}\mathbb{H}_t^{(\lambda)}(x) = -\mathcal{L}^\dagger(t)\cdot\mathbb{H}_t^{(\lambda)}(x) \tag{96}$$

where $\mathcal{L}(t) = \mathcal{L}_\lambda$ for $t \in [0, T]$ and $\mathcal{L}(t) = \mathcal{L}_0$ outside this time window, with the operators defined in (29) and (33); similar for the conjugate operator $\mathcal{L}^\dagger(t)$.

This gives, for example, for $t \le 0$, $H_t^{(\lambda)}(x) = r_0(x)$ (defined in (35)), whereas $\mathbb{H}_t^{(\lambda)}(x) \sim e^{-t\mathcal{L}_0^\dagger}\cdot e^{T\mathcal{L}_\lambda^\dagger}\cdot \ell_0(x)$, (upto a constant pre-factor) which in the large T limit, gives $\mathbb{H}_t^{(\lambda)}(x) \sim e^{T\mu(\lambda)}\left[e^{-t\mathcal{L}_0^\dagger}\cdot\ell_\lambda\right](x)$. Substituting these in (93) and (95b) we get the expression for the canonical measure (38a) and effective force (39a), respectively, in region I of Fig. 1. Results for rest of the regions in Sect. 3.2 can be obtained similarly.

Lastly, from (25b) one could see that for the observable (4), the generating function $G_T^{(\lambda)}(C|C_0)$ in (6) is identical to $Z_T^{(\lambda)}(C)$ if one sets $Z_0^{(\lambda)}(C) = \delta_{C,C_0}$. Then from the above calculation it is straightforward to show that in the continuous limit one would get (32).

D Path integral formulation

The path integral formulation of a Fokker–Planck equation is standard [83]. The Fokker–Planck equation (29) can be written as

$$\frac{dP_t(x)}{dt} = -\frac{d}{dx}\left[F(x)P_t(x)\right] + \frac{\epsilon}{2}\frac{d^2}{dx^2}P_t(x) \equiv -\mathcal{H}\left(x, -i\frac{d}{dx}\right)P_t(x)$$

such that $H(x, p) = F'(x) + iF(x)p + \frac{\epsilon}{2}p^2$. Considering a small increment dt in time, we get

$$\begin{aligned} P_{t+dt}(x) &\simeq \int dx' \left[1 - dt\,\mathcal{H}\left(x, -i\frac{d}{dx}\right)\right]\delta(x - x')P_t(x') \\ &\simeq \int \frac{dp\,dx'}{2\pi}\left[1 - dt\,\mathcal{H}(x, p)\right]e^{i\,p(x-x')}P_t(x') \end{aligned}$$

where we used the Fourier transform of the Dirac delta function $\delta(x - x')$. Iterating the evolution and taking $dt \to 0$ limit we get a path integral representation

$$P_T(x) = \int_{z(0)=y}^{z(T)=x} \mathcal{D}[z, p]e^{\int_0^T dt\,[ip\dot{z} - H(z,p)]}$$

with an initial condition $P_0(z) = \delta(z - y)$. The $H(z, p)$ is quadratic in p, and the corresponding path integral can be evaluated exactly, giving

$$P_T(x) = \int_{z(0)=y}^{z(T)=x} \mathcal{D}[z]e^{-\frac{1}{2\epsilon}\int_0^T dt\,(\dot{z}-F(z))^2 - \int_0^T dt\,F'(z)}$$

This is the path integral representation of the Fokker–Planck equation (29).

It is straightforward to generalize the above analysis for the generating function (32) and we get

$$G_T^{(\lambda)}(x|y) = \int_{z(0)=y}^{z(T)=x} \mathcal{D}[z]e^{\mathbb{S}_T^{(\lambda)}[z(t)]} \tag{97a}$$

where the Action

$$\mathbb{S}_T^{(\lambda)}[z] = \int_0^T dt\left[\lambda f(z) + \lambda\dot{z}h(z) - \frac{(\dot{z} - F(z))^2}{2\epsilon} - F'(z) - \epsilon\lambda\left(\alpha - \frac{1}{2}\right)h'(z)\right] \tag{97b}$$

Taking small ϵ limit, we get $\mathbb{S}_T^{(\frac{\kappa}{\epsilon})}[z] \simeq \frac{1}{\epsilon}S_T^{(\kappa)}[z]$ with the latter given in (56) where we used $h(x) = 0$.

References

1. Mey, A.S.J.S., Geissler, P.L., Garrahan, J.P.: Rare-event trajectory ensemble analysis reveals metastable dynamical phases in lattice proteins. Phys. Rev. E **89**, 032109 (2014)
2. Delarue, M., Koehl, P., Orland, H.: Ab initio sampling of transition paths by conditioned Langevin dynamics. J. Chem. Phys. **147**, 152703 (2017)
3. Dykman, M.I., Mori, E., Ross, J., Hunt, P.M.: Large fluctuations and optimal paths in chemical kinetics. J. Chem. Phys. **100**, 5735 (1994)
4. Lauri, J., Bouchet, F.: Computation of rare transitions in the barotropic quasi-geostrophic equations. N. J. Phys. **17**, 015009 (2015)

5. Garrahan, J.P., Jack, R.L., Lecomte, V., Pitard, E., van Duijvendijk, K., van Wijland, F.: Dynamical first-order phase transition in kinetically constrained models of glasses. Phys. Rev. Lett. **98**, 195702 (2007)
6. Garrahan, J.P., Jack, R.L., Lecomte, V., Pitard, E., van Duijvendijk, K., van Wijland, F.: Dynamical first-order phase transition in kinetically constrained models of glasses. J. Phys. A **42**, 075007 (2009)
7. Dorlas, T.C., Wedagedera, J.R.: Large deviations and the random energy model. Int. J. Mod. Phys. B **15**, 1 (2001)
8. Varadhan, S.R.S.: Asymptotic probabilities and differential equations. Commun. Pure Appl. Math. **19**, 261 (1966)
9. Varadhan, S.R.S.: The large deviation problem for empirical distributions of Markov processes. In: Large Deviations and Applications, p. 33. SIAM (1984). https://doi.org/10.1137/1.9781611970241.ch9
10. Varadhan, S.R.S.: Large deviations for random walks in a random environment. Commun. Pure Appl. Math. **56**, 1222 (2003)
11. Ellis, R.S.: Entropy, Large Deviations, and Statistical Mechanics. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer, Berlin (1985)
12. Derrida, B.: Non-equilibrium steady states: fluctuations and large deviations of the density and of the current. J. Stat. Mech. P07023 (2007)
13. den Hollander, F.: Large Deviations, Fields Institute Monographs. American Mathematical Society, Providence (2008)
14. Touchette, H.: The large deviation approach to statistical mechanics. Phys. Rep. **478**, 1 (2009)
15. Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications Stochastic Modelling and Applied Probability. Springer, Berlin (2009)
16. Kurchan, J.: Fluctuation theorem for stochastic dynamics. J. Phys. A **31**, 3719 (1998)
17. Gallavotti, G., Cohen, E.G.D.: Dynamical ensembles in nonequilibrium statistical mechanics. Phys. Rev. Lett. **74**, 2694 (1995)
18. Lebowitz, J.L., Spohn, H.: A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics. J. Stat. Phys. **95**, 333 (1999)
19. Freidlin, M.I., Szücs, J., Wentzell, A.D.: Random Perturbations of Dynamical Systems. Grundlehren der mathematischen Wissenschaften. Springer, Berlin (2012)
20. Graham, R., Tél, T.: Weak-noise limit of Fokker–Planck models and nondifferentiable potentials for dissipative dynamical systems. Phys. Rev. A **31**, 1109 (1985)
21. Graham, R.: Statistical theory of instabilities in stationary nonequilibrium systems with applications to lasers and nonlinear optics. In: Springer Tracts in Modern Physics: Ergebnisse der exakten Naturwissenschaftenc, vol. 66, p.1. Springer, Berlin (1973). https://doi.org/10.1007/978-3-662-40468-3_1
22. Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., Landim, C.: Fluctuations in stationary nonequilibrium states of irreversible processes. Phys. Rev. Lett. **87**, 040601 (2001)
23. Derrida, B.: Microscopic versus macroscopic approaches to non-equilibrium systems. J. Stat. Mech. **2011**, P01030 (2011)
24. Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., Landim, C.: Macroscopic fluctuation theory. Rev. Mod. Phys. **87**, 593 (2015)
25. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain markov process expectations for large time, I. Commun. Pure Appl. Math. **28**, 1 (1975)
26. Derrida, B., Lebowitz, J.L.: Exact large deviation function in the asymmetric exclusion process. Phys. Rev. Lett. **80**, 209 (1998)
27. Bodineau, T., Derrida, B.: Current fluctuations in nonequilibrium diffusive systems: an additivity principle. Phys. Rev. Lett. **92**, 180601 (2004)
28. Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., Landim, C.: Current fluctuations in stochastic lattice gases. Phys. Rev. Lett. **94**, 030601 (2005)
29. Hurtado, P.I., Garrido, P.L.: Large fluctuations of the macroscopic current in diffusive systems: a numerical test of the additivity principle. Phys. Rev. E **81**, 041102 (2010)
30. Hurtado, P.I., Espigares, C.P., del Pozo, J.J., Garrido, P.L.: Thermodynamics of currents in nonequilibrium diffusive systems: theory and simulation. J. Stat. Phys. **154**, 214 (2014)
31. Bertini, L., Faggionato, A., Gabrielli, D.: Large deviations of the empirical flow for continuous time Markov chains. Ann. Inst. H. Poincaré Prob. Stat. **51**, 867 (2015)
32. Touchette, H.: Introduction to dynamical large deviations of Markov processes. In: Lecture Notes of the 14th International Summer School on Fundamental Problems in Statistical Physics. Physica A 504, 5 (2018)
33. Maes, C., Netocný, K.: Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states. Europhys. Lett. **82**, 30003 (2008)
34. Maes, C., Netocnný, K., Wynants, B.: Steady state statistics of driven diffusions. Physica A **387**, 2675 (2008)

35. Evans, R.M.L.: Rules for transition rates in nonequilibrium steady states. Phys. Rev. Lett. **92**, 150601 (2004)
36. Evans, R.M.L.: Detailed balance has a counterpart in non-equilibrium steady states. J. Phys. A **38**, 293–313 (2004)
37. Hartmann, C., Schütte, C.: Efficient rare event simulation by optimal nonequilibrium forcing. J. Stat. Mech P11004 (2012)
38. Majumdar, S.N., Orland, H.: Effective Langevin equations for constrained stochastic processes. J. Stat. Mech P06039 (2015)
39. Fleming, W.H.: Stochastic control and large deviations. In: Bensoussan, A., Verjus, J.P. (eds.) Future Tendencies in Computer Science, Control and Applied Mathematics, p. 291. Springer, Berlin (1992)
40. Nemoto, T., Sasa, Si: Thermodynamic formula for the cumulant generating function of time-averaged current. Phys. Rev. E **84**(6), 061113 (2011)
41. Lecomte, V., Appert-Rolland, C., van Wijland, F.: Thermodynamic formalism for systems with Markov dynamics. J. Stat. Phys. **127**, 51 (2007)
42. Strook, D.W.: An Introduction to Markov Processes. Graduate Texts in Mathematics, 2nd edn. Springer, Berlin (2014)
43. Borkar, V.S., Juneja, S., Kherani, A.A.: Peformance analysis conditioned on rare events: an adaptive simulation scheme. Commun. Inf. Syst. **3**, 259–278 (2003)
44. Jack, R.L., Sollich, P.: Large deviations and ensembles of trajectories in stochastic models. Prog. Theor. Phys. Suppl. **184**, 304 (2010)
45. Jack, R.L., Sollich, P.: Effective interactions and large deviations in stochastic processes. Eur. Phys. J. Spec. Top. **224**, 2351 (2015)
46. Chetrite, R., Touchette, H.: Nonequilibrium microcanonical and canonical ensembles and their equivalence. Phys. Rev. Lett. **111**, 120601 (2013)
47. Chetrite, R., Touchette, H.: Nonequilibrium markov processes conditioned on large deviations. Ann. Henri Poincaré **16**, 2005 (2015)
48. Chetrite, R., Touchette, H.: Variational and optimal control representations of conditioned and driven processes. J. Stat. Mech P12001 (2015)
49. Szavits-Nossan, J., Evans, M.R.: Inequivalence of nonequilibrium path ensembles: the example of stochastic bridges. J. Stat. Mech. P12008 (2015)
50. Nyawo, P.T., Touchette, H.: Large deviations of the current for driven periodic diffusions. Phys. Rev. E **94**(3), 032101 (2016)
51. Tizón-Escamilla, N., Lecomte, V., Bertin, E.: Effective driven dynamics for one-dimensional conditioned Langevin processes in the weak-noise limit. J. Stat. Mech. **2019**, 013201 (2019)
52. Derrida, B., Sadhu, T.: Large deviations conditioned on large deviations II: fluctuating hydrodynamics (2019). arXiv:1905.07175
53. Landau, L., Lifshitz, E.: Quantum Mechanics. MIR, Moskow (1967)
54. Derrida, B., Douçot, B., Roche, P.E.: Current fluctuations in the one-dimensional symmetric exclusion process with open boundaries. J. Stat. Phys. **115**, 717 (2004)
55. Hirschberg, O., Mukamel, D., Schütz, G.M.: Density profiles, dynamics, and condensation in the ZRP conditioned on an atypical current. J. Stat. Mech. P11023 (2015)
56. Schütz, G.M.: Duality Relations for the Periodic ASEP Conditioned on a Low Current, p. 323. Springer, Cham (2016)
57. Popkov, V., Schütz, G.M.: Transition probabilities and dynamic structure function in the ASEP conditioned on strong flux. J. Stat. Phys. **142**, 627 (2011)
58. Carollo, F., Garrahan, J.P., Lesanovsky, I., Pérez-Espigares, C.: Making rare events typical in Markovian open quantum systems. Phys. Rev. A **98**, 010103 (2018)
59. Angeletti, F., Touchette, H.: Diffusions conditioned on occupation measures. J. Math. Phys. 57 (2016)
60. Van Kampen, N.: Stochastic Processes in Physics and Chemistry, 3rd edn. North-Holland Personal Library, Elsevier, Amsterdam (2007)
61. Popkov, V., Schütz, G.M., Simon, D.: ASEP on a ring conditioned on enhanced flux. P10007. J. Stat. Mech. (2010)
62. Ellis, R.S.: Large deviations for a general class of random vectors. Ann. Probab. **12**, 1–12 (1984)
63. Bodineau, T., Derrida, B.: Distribution of current in nonequilibrium diffusive systems and phase transitions. Phys. Rev. E **72**, 066110 (2005)
64. Harris, R.J., Rákos, A., Schütz, G.M.: Breakdown of Gallavotti–Cohen symmetry for stochastic dynamics. Eur. Phys. Lett. **75**, 227–233 (2006)
65. Espigares, C.P., Garrido, P.L., Hurtado, P.I.: Dynamical phase transition for current statistics in a simple driven diffusive system. Phys. Rev. E **87**, 032115 (2013)

66. Touchette, H.: Equivalence and nonequivalence of ensembles: thermodynamic, macrostate, and measure levels. J. Stat. Phys. **159**, 987–1016 (2015)
67. McKean, H.P.: Stochastic Integrals. Probability and Mathematical Statistics: A Series of Monographs and Textbooks. Academic Press, Cambridge (1969). https://doi.org/10.1016/B978-1-4832-3054-2.50008-X
68. Mehl, J., Speck, T., Seifert, U.: Large deviation function for entropy production in driven one-dimensional systems. Phys. Rev. E **78**, 011123 (2008)
69. Speck, T., Engel, A., Seifert, U.: The large deviation function for entropy production: the optimal trajectory and the role of fluctuations. J. Stat. Mech. P12001 (2012)
70. Risken, H.: The Fokker–Planck Equation: Methods of Solutions and Applications. Springer Series in Synergetics, 2nd edn. Springer, Berlin (1996)
71. Brownstein, K.R.: Criterion for existence of a bound state in one dimension. Am. J. Phys. **68**, 160–161 (2000)
72. Buell, W.F., Shadwick, B.A.: Potentials and bound states. Am. J. Phys. **63**, 256–258 (1995)
73. Ashbaugh, M.S., Benguria, R.D.: Optimal bounds for ratios of eigenvalues of one-dimensional Schrödinger operators with Dirichlet boundary conditions and positive potentials. Commun. Math. Phys. **124**, 403–415 (1989)
74. Andrews, B., Clutterbuck, J.: Proof fundamental gap conjecture. J. Am. Math. Soc. **24**, 899–916 (2011)
75. Nickelsen, D., Engel, A.: Asymptotics of work distributions: the pre-exponential factor. Eur. Phys. J. B **82**, 207–218 (2011)
76. Engel, A.: Asymptotics of work distributions in nonequilibrium systems. Phys. Rev. E **80**, 021120 (2009)
77. Baule, A., Touchette, H., Cohen, E.G.D.: Stick-slip motion of solids with dry friction subject to random vibrations and an external field. Nonlinearity **24**, 351 (2011)
78. Sadhu, T., Derrida, B.: Correlations of the density and of the current in non-equilibrium diffusive systems. J. Stat. Mech. 113202 (2016)
79. Bertini, L., Sole, A.D., Gabrielli, D., Landim, C.: Macroscopic fluctuation theory for stationary non-equilibrium states. J. Stat. Phys. **107**, 635 (2002)
80. Meerson, B., Zilber, P.: Large deviations of a long-time average in the Ehrenfest urn model. J. Stat. Mech. **2018**, 119901 (2018)
81. Proesmans, K., Derrida, B.: Large-deviation theory for a Brownian particle on a ring: a WKB approach. J. Stat. Mech. **2019**, 023201 (2019)
82. Fischer, L.P., Pietzonka, P., Seifert, U.: Large deviation function for a driven underdamped particle in a periodic potential. Phys. Rev. E **97**, 1–10 (2018)
83. Kubo, R., Matsuo, K., Kitahara, K.: Fluctuation and relaxation of macrovariables. J. Stat. Phys. **9**, 51 (1973)
84. Schrödinger, E.: Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. Ann. Henri Poincaré **2**, 269–310 (1932)
85. Zambrini, J.C.: Euclidean quantum mechanics. Phys. Rev. A **35**(9), 3631–3649 (1987)
86. Cruzeiros, A.B., Zambrini J.C.: Euclidean quantum mechanics. An outline. In: Stochastic Analysis and Applications in Physics, pp. 59–97. Springer Netherlands, Dordrecht (1994). https://doi.org/10.1007/978-94-011-0219-3_4