CrossMark

A Formal View on Level 2.5 Large Deviations and Fluctuation Relations

Andre C. Barato[1] · **Raphael Chetrite[2]**

© Springer Science+Business Media New York 2015

Abstract We obtain the rate function for the level 2.5 of large deviations for pure jump and diffusion processes. This result is proved by two methods: tilting, for which a tilted process with an appropriate typical behavior is considered, and a spectral method, for which the scaled cumulant generating function is used. We also briefly discuss fluctuation relations, pointing out their connection with large deviations at the level 2.5.

1 Introduction

An important recent progress in nonequilibrium statistical physics was the discovery of various fluctuation relations, which are identities involving the statistics of a fluctuating entropy. In particular, the Gallavotti–Cohen–Evans–Morriss (GCEM) relation [25,26,29] imposes a peculiar symmetry related to the rare events associated with this fluctuating entropy. The appropriate theory to describe such rare events is large deviation theory, which is a very fashionable subject in statistical physics [42,49] and in modern probability [16–18,23,50], as evidenced by the Abel Prize awarded to S. R. S Varadhan in 2007.

"I would like to offer some remarks about the word "formal". For the mathematician, it usually means "according to the standard of formal rigor, of formal logic". For the physicists, it is more or less synonymous with "heuristic" as opposed to "rigorous"". Pierre Cartier. Mathemagics (A Tribute to L. Euler and R. Feynman). Seminaire Lotharingien de Combinatoire 44 (2000), Article B44d.

✉ Andre C. Barato
  barato@theo2.physik.uni-stuttgart.de

  Raphael Chetrite
  raphael.chetrite@unice.fr

[1] II. Institut für Theoretische Physik, Universität Stuttgart, 70550 Stuttgart, Germany

[2] Laboratoire J. A. Dieudonné, UMR CNRS 6621, Université de Nice Sophia-Antipolis, Nice, France

We recall that a time dependent measure $\mu_T(dx)$ satisfies the large deviation principle if at large times it takes an exponential decreasing form. This exponential decay is characterized by a lower semi-continuous positive function $I(x)$, which is called the rate function. This function is such that for any set A

$$-\inf_{x \in A^0} I(x) \leq \liminf_{T \to +\infty} \frac{1}{T} \ln \mu_T(A) \leq \limsup_{T \to +\infty} \frac{1}{T} \ln \mu_T(A) \leq -\inf_{x \in \overline{A}} I(x), \qquad (1)$$

where A^0 is the interior of A and $\overline{A}$ is the closure of A. This can be stated less formally as

$$\mu_T(dx) \sim \exp(-T I(x)) \, dx. \qquad (2)$$

Historically, large deviation theory originated in the nineteenth century with pioneering works in statistical mechanics [8]. One of the most important contributions to large deviation theory was the general approach for Markov processes developed by Donsker and Varadhan [19–22]. In this series of papers, they identified three levels of large deviations:

– Level 1, which is the study of fluctuations of additive observables with respect to the mean.
– Level 2, related to fluctuations of the fraction of time spent in each state.
– Level 3, concerning fluctuations on the statistics of infinite trajectories.

The ranking of these levels establishes a hierarchy in which a lower level can be deduced from a higher one by contraction. Donsker and Varadhan proved the large deviation principle for Markov processes at the level 3 by studying random probability measures on infinite trajectories. This queen large deviation result posses an explicit rate function, which is the relative entropy density. Moreover, they proved the large deviation principle at the level 2 for the empirical density, defined as the fraction of time spent in each state up to time T. Contrary to level 3, the rate function for level 2 admits a variational representation, which is in general not explicit. Hence, the explicit character of level 3 disappears after contracting to level 2. At discrete time a more detailed picture is available: it is possible to investigate the large deviation of the k symbol empirical measure and prove that the rate function can be obtained explicitly if $k \geq 2$. Thus filling the gap between level 2 and 3.

However, in discrete time the extended process $(X_t, X_{t+1}, \ldots, X_{t+k-1})$ is itself a Markov chain and therefore the intermediate level can be derived from the level 2. This magnification trick is no longer possible in continuous time. Until recently, no result existed in the literature to fill this level 2–3 gap for continuous time. The first study of this gap in the continuous time setting was by Kesidis and Walrand [32], for pure jump processes with two states. They obtained explicitly the rate function for the joint probability of the empirical density and the empirical flow counting the number of jumps between pair of states up to time T. This intermediate level was then called 2.5. This issue was later studied by De La Fortelle [15], who obtained a weak large deviation in the same context but for countable space.

Somewhat in parallel, in nonequilibrium statistical physics, it has been found that the empirical density at level 2 is not sufficient to study fluctuations of the entropy production and of currents. This also motivated the search of an intermediate level for pure jump and diffusion processes, by Maes and collaborators [39,40], and by Chernyak et al. [9]. Finally, Bertini et al. [5] succeeded in proving rigorously the level 2.5 for pure jump processes in a countable space. Rigorous results for diffusion processes have been obtained in [36].

The purpose of our contribution is to present the level 2.5 of large deviations for continuous time processes and discuss its connection with fluctuation relations. Whereas the explicit rate functions for the level 2.5 of large deviations calculated here have been obtained in [5,9,15,39,40], our presentation unifies the proofs for pure jump and diffusion processes, and clearly compares the two different methods used to obtain these rate functions, namely,

tilting and a spectral method. Moreover, some of the proofs presented here are completely original.

The organization of the paper is as follows. Section 2 sets the stage with the definition of Markov processes, which include jump and diffusion processes. Particularly, in Sect. 2.1 we recall basic concepts of Markov processes like transition probability, generator, stationary and equilibrium states, and trajectorial measure. In Sect. 2.2, we introduce the empirical density, empirical flow, empirical current, and the action functional, which are the fluctuating observables studied in the paper. In Sect. 3 we obtain the finite time fluctuation relation, which results as a tautology from the definition of the action functional. Section 4 is the cornerstone of the paper and deals with the level 2.5 of large deviations. In Sect. 4.1 we use the tilting method to obtain the rate function characterizing the level 2.5. This proof is related to results from [39,40], but the presentation given here is original. Section 4.2 contains the spectral method. In this case, for pure jump processes our proof is original. For diffusion processes the spectral method has been used in [9], in comparison to this reference we expurgate the field theoretical language by using the Girsanov lemma. Finally, in Sect. 5 we obtain a stationary fluctuation relation at the level 2.5 and, by contraction, the GCEM symmetry for the fluctuating entropy.

2 Models and Observables

2.1 Homogeneous Ergodic Markov Processes

We start with a brief overview of homogeneous Markov processes [13,44,46,48], considering continuous time Markov processes X_t taking values in a state space $\mathcal{E}$, which can be continuous, as for example $\mathbb{R}^d$, or a counting space.

2.1.1 Elements of Ergodic Markov Processes

A time-homogeneous Markov process can be defined by a family of transitions kernel $P_t(x, dy)$, which is the conditional probability that $X_{t+t'} \in [y, y + dy]$ given that $X_{t'} = x$. This conditional probability satisfies the Chapmann-Kolmogorov rule

$$\int_{\mathcal{E}} P_s(x, dy) P_t(y, dz) \, = \, P_{s+t}(x, dz), \tag{3}$$

where the measure dy means the Lebesgue measure or the counting measure, depending on $\mathcal{E}$. The semi-group associated with the transition kernel is defined by its action on a bounded measurable function f in $\mathcal{E}$,

$$P_t[f](x) \equiv \int_{\mathcal{E}} P_t(x, dy) f(y). \tag{4}$$

The infinitesimal generator L, formally defined as $P_t \equiv \exp(tL)$, leads to the forward and backward Kolmogorov equations,

$$\partial_t P_t = P_t \circ L \quad \text{and} \quad \partial_t P_t = L \circ P_t, \tag{5}$$

respectively. The symbol $\circ$ means composition of operators and the initial condition is $P_0 = \mathcal{I}$, where $\mathcal{I}$ is the identity kernel. Conservative processes (without death or explosion), for which the normalization condition $\int P_t(x, dy) = 1$ holds, are often considered in Physics. The generator must then obey $L[1] = 0$, where 1 is the function which is equal to 1 on $\mathcal{E}$.

The time evolution of the instantaneous one point measure $\mu_t(dy) = \int_{\mathcal{E}} \mu_0(dx_0) P_t(x_0, dy)$ can be deduced from the Kolmogorov equation (5), leading to the Fokker–Planck equation $\partial_t \mu_t = L^\dagger[\mu_t]$, where $L^\dagger$ is the adjoint of L with respect to the Lebesgue or counting measure. Since we are considering ergodic Markov processes, there is a unique invariant probability measure μ_{inv} satisfying

$$L^\dagger[\mu_{inv}] = 0. \tag{6}$$

The process is said to be in *equilibrium* w.r.t μ_{inv} if the detailed balance relation is satisfied, i.e.,

$$\mu_{inv}(dx) P_t(x, dy) = \mu_{inv}(dy) P_t(y, dx). \tag{7}$$

In the following it is assumed that the one point measure is smooth with respect to the Lebesgue measure, for example with the conditions of the Hormander theorem [30,41] for a diffusion process, leading to $\mu_t(dx) \equiv \rho_t(x)dx$. With $\mu_{inv}(dx) \equiv \rho_{inv}(x)dx$, the detailed balance condition (7) can be written as[1]

$$\rho_{inv} \circ L \circ \rho_{inv}^{-1} = L^\dagger. \tag{8}$$

In addition to the characterization by the semi-group or the generator, a Markov process can be characterized by its trajectorial measure. The sample path of the process up to time T is the random function $X_0^T : t \in [0, T] \to X_t$. It is a random variable in the space of trajectories $D([0, T], \mathcal{E})$. This trajectorial measure $d\mathbb{P}_{L,\mu_0,T}[x_0^T]$, where μ_0 is the initial measure, is roughly the probability that the trajectory X_0^T equals x_0^T. The expectation of an arbitrary functional $F[X_0^T]$ of the trajectories is then written as,

$$\mathbb{E}_{L,\mu_0}[F] = \int F[x_0^T] \, d\mathbb{P}_{L,\mu_0,T}[x_0^T]. \tag{9}$$

The finite time distributions are sufficient to characterize $d\mathbb{P}_{L,\mu_0,T}$, more precisely, Eq. (9) may be rewritten as

$$\mathbb{E}_{L,\mu_0}[F] = \int_{\mathcal{E}^{n+1}} F(x_0, x_1, \ldots, x_{n-1}, x_n) \mu_0(dx_0) \exp(t_1 L)(x_0, dx_1) \tag{10}$$
$$\times \exp((t_2 - t_1) L)(x_1, dx_2) \cdots \exp((T - t_{n-1}) L)(x_{n-1}, dx_n),$$

for the cylindrical functional

$$F[X] = F(X_0, X_{t_1}, X_{t_2}, \ldots, X_{t_{n-1}}, X_T), \tag{11}$$

with $0 \leq t_1 \leq t_2 \leq \cdots \leq t_{n-1} \leq T$. In the following we consider the two most prominent classes of Markov processes: jump and diffusion processes.

2.1.2 Pure Jump Processes

A Markov process is called a pure jump process if after "arriving" into a state the system stays there for a random exponentially distributed time interval and then jumps to another state. The transition rates $W(x, y)$ give the probability per unit of time for the transition $x \to y$. Moreover, with regularity conditions (see [24, Chap. 8] for example), it is possible to

---

[1] The expression $\rho_{inv} \circ L \circ \rho_{inv}^{-1}$ must be understood as the composition of three operators, first the operator multiplication by ρ_{inv}^{-1}, second the operator L and last the operator multiplication by ρ_{inv}. This type of notation is recurrently used in the article.

prove that for pure jump possesses the generator acting on the bounded measurable function $h : \mathcal{E} \to \mathbb{R}$ is

$$L[h](x) = \int_{\mathcal{E}} W(x, y)(h(y) - h(x))\,dy, \tag{12}$$

for all $x \in \mathcal{E}$. The detailed balance condition (7) with respect to the density ρ_{inv} takes the form

$$\rho_{inv}(x)W(x, y) = \rho_{inv}(y)W(y, x). \tag{13}$$

A relevant quantity in this paper is the current associated with the density ρ_t,

$$J_{\rho_t}(x, y) \equiv \rho_t(x)W(x, y) - W(y, x)\rho_t(y). \tag{14}$$

From equation (6), the current associated with the invariant density is conserved,

$$\int dy\, J_{\rho_{inv}}(x, y) = 0. \tag{15}$$

At the trajectory level it is possible to compare the trajectorial measure (9) of two processes with different transition rates, with the condition that they both have the same set of non vanishing rates. To this end, we introduce the non conservative Markovian generator[2]

$$L_{V_1,V_2}[h](x) \equiv \left(\int_{\mathcal{E}} W(x, y)\left[\exp(V_2(x, y))h(y) - h(x)\right]dy\right) + V_1(x)h(x), \tag{16}$$

for all functions h, with $V_1 : \mathcal{E} \to \mathbb{R}$ and $V_2 : \mathcal{E}^2 \to \mathbb{R}$. We call this generator the twisted generator. From the Girsanov lemma [34, Proposition 2.6] and the Feynamn Kac relation [44,46], it follows that $d\mathbb{P}_{L_{V_1,V_2},\mu_0,T}$ is absolutely continuous w.r.t. $d\mathbb{P}_{L,\mu_0,T}$, and the explicit Radon Nykodym derivative is given by

$$\frac{d\mathbb{P}_{L_{V_1,V_2},\mu_0,T}}{d\mathbb{P}_{L,\mu_0,T}}\left[x_0^T\right] = \exp\left(\sum_{0 \le s \le T / x_{s-} \ne x_{s+}} V_2(x_{s-}, x_{s+}) + \int_0^T ds\,V_1(x_s)\right), \tag{17}$$

where $x_{s-} \equiv \lim_{\delta \to 0} x_{s-\delta}$ and $x_{s+} \equiv \lim_{\delta \to 0} x_{s+\delta}$. Hence, the sum $\sum_{0 \le s \le T / x_{s-} \ne x_{s+}}$ is over all jumps in the trajectory x_0^T. In particular, for two conservative jump processes, one with rates W and the other with rates $W_{V_2}(x, y) = W(x, y)\exp(V_2(x, y))$ relation (17) becomes

$$\frac{d\mathbb{P}_{L_{V_2},\mu_0,T}}{d\mathbb{P}_{L,\mu_0,T}}[x] = \exp\left(\sum_{0 \le s \le T / x_{s-} \ne x_{s+}} V_2(x_{s-}, x_{s+}) - \int_0^T ds\,(W\exp(V_2) - W)[1](x_s)\right),$$

$$\tag{18}$$

where L_{V_2} is the conservative generator obtained from (16) by setting

$$V_1 = (W)[1] - (W\exp(V_2))[1] = \int W(x, y)dy - \int W(x, y)\exp(V_2(x, y))dy. \tag{19}$$

2.1.3 Diffusion Processes

A diffusion process X_t in a d-dimensional manifold is described by the differential equation

$$dX = A_0(X)dt + \sum_\alpha A_\alpha(X) \circ dW_\alpha(t). \tag{20}$$

---

[2] In operational notation $L_{V_1,V_2} = W\exp(V_2) - W[1] + V_1$.

where the drift A_0 and the diffusion coefficient A_α are arbitrary smooth vector fields on $\mathcal{E}$, W_α are independent Wiener processes, and the range of α is model dependent. The symbol $\circ$ indicates that the Stratonovich convention is used. The explicit form of the generator related to (20) is

$$L = A_0 \cdot \nabla + \sum_\alpha \frac{1}{2} (A_\alpha . \nabla)^2 = \widehat{A_0} \cdot \nabla + \frac{1}{2} \nabla \cdot D \cdot \nabla, \tag{21}$$

with the modified drift and covariance

$$\widehat{A_0}(x) = A_0(x) - \frac{1}{2} \sum_\alpha (\nabla \cdot A_\alpha)(x) A_\alpha(x) \quad \text{and} \quad D^{ij}(x) = \sum_\alpha A_\alpha^i(x) A_\alpha^j(x), \tag{22}$$

respectively, where $i = 1, \ldots, d$ and $j = 1, \ldots, d$. It is assumed that D is strictly positive. The detailed balance relation (7) with respect to the invariant measure $\mu_{inv}(dx) = \rho_{inv}(x)dx$ is then equivalent to $\widehat{A_0} = \frac{D}{2} \nabla (\ln \rho_{inv})$.

A central quantity for diffusion processes is the hydrodynamic probability current [45]

$$J_{\rho_t} = \widehat{A_0} \rho_t - \frac{D}{2} (\nabla \rho_t). \tag{23}$$

The conservation of the current associated with the invariant density then reads

$$\nabla \cdot J_{\rho_{inv}} = 0. \tag{24}$$

Similar to jump processes, the trajectorial measure of two diffusion processes can be compared with a generator corresponding to a non-conservative process, which in the present case is defined as

$$L' \equiv L + B_2 \cdot \nabla + B_1, \tag{25}$$

where B_2 and B_1 are arbitrary vector field and scalar, respectively. Combining the Cameron–Martin–Girsanov lemma [46,48] and the Feynamm–Kac relation [44,46], it follows that

$$\frac{d\mathbb{P}_{L',\mu_0,T}[x]}{d\mathbb{P}_{L,\mu_0,T}[x]} = \exp(V_T[x]), \tag{26}$$

where

$$V_T = \int_0^T \left[D^{-1}(x_u) B_2(x_u) \circ dx_u \right. \\ \left. + \left(B_1(x_u) - D^{-1}(x_u) B_2(x_u) \left(\widehat{A_0} + \frac{B_2}{2} \right)(x_u) - \frac{1}{2} (\nabla \cdot B_2)(x_u) \right) du \right]. \tag{27}$$

Choosing

$$B_2 = DV_2 \quad \text{and} \quad B_1 = V_2 \cdot \left(\widehat{A_0} + \frac{DV_2}{2} \right) + \frac{1}{2} \nabla \cdot (DV_2) + V_1, \tag{28}$$

we obtain

$$V_T = \int_0^T dt \, [V_1(X_t) + V_2(X_t) \circ dX_t]. \tag{29}$$

Equation (26) then becomes

$$\frac{d\mathbb{P}_{L_{V_1,V_2},\mu_0,T}}{d\mathbb{P}_{L,\mu_0,T}}[X] = \exp \left(\int_0^T dt \, [V_1(X_t) + V_2(X_t) \circ dX_t] \right), \tag{30}$$

where the twisted generator reads

$$
\begin{aligned}
L_{V_1,V_2} = L' &= L + DV_2\nabla + V_2 \cdot \left(\widehat{A_0} + \frac{DV_2}{2}\right) + \frac{1}{2}\nabla \cdot (DV_2) + V_1 \\
&= \widehat{A_0} \cdot (\nabla + V_2) + (\nabla + V_2) \circ \frac{D}{2} \circ (\nabla + V_2) + V_1.
\end{aligned}
\tag{31}
$$

2.2 Empirical Observables and Ergodic Behavior

2.2.1 Empirical Density, Flow and Current

The set of functional observables that define the level 2.5 of large deviations depend of the type of Markov processes considered. For pure jump processes the set of observables is the empirical density ρ_T^e and empirical flow C_T^e. They are given by

$$
\rho_T^e(x) = \frac{1}{T}\int_0^T \delta(X_t - x)\,dt \quad \text{and}
$$

$$
C_T^e(x, y) = \frac{1}{T}\sum_{0 \le s \le T/X_{s-} \ne X_{s+}} \delta(X_{t-} - x)\,\delta(X_{t+} - y).
\tag{32}
$$

The empirical density $\rho_T^e(x)$[3] can be understood as the fraction of time spent in x over $[0, T]$ and the empirical flow $C_T^e(x, y)$ as the number of jumps from x to y (times $1/T$) during the trajectory. Another functional of central interest is the empirical current

$$
J_T^e(x, y) = C_T^e(x, y) - C_T^e(y, x).
\tag{33}
$$

Since we assume the system to be ergodic, the law of large numbers for the empirical density and flow becomes

$$
\rho_T^e \to \rho_{inv} \quad \text{and} \quad C_T^e \to C_{\rho_{inv}},
\tag{34}
$$

where

$$
C_{\rho_{inv}}(x, y) = \rho_{inv}(x)W(x, y).
\tag{35}
$$

Moreover, the finite time Kirchoff's law [33] reads

$$
\begin{aligned}
&\int dy\, C_T^e(x, y) - \int dy\, C_T^e(y, x) \\
&= \frac{1}{T}\sum_{0 \le s \le T/X_{s-} \ne X_{s+}} \delta(X_{t-} - x) - \frac{1}{T}\sum_{0 \le s \le T/X_{s-} \ne X_{s+}} \delta(X_{t+} - x) \\
&= \frac{\delta(X_0 - x) - \delta(X_T - x)}{T} = O(1/T).
\end{aligned}
\tag{36}
$$

In the following we will show that the large deviation rate function of C_T^e is infinite for any untypical C not fulfilling

$$
\int dy\, C(x, y) = \int dy\, C(y, x).
\tag{37}
$$

For diffusion processes, the set of observables is composed by the empirical density ρ_T^e and the empirical current j_T^e, which read

---

[3] Rigorously, we should instead define the empirical measure $\mu_T^e = \frac{1}{T}\int_0^T \delta_{X_t}\,dt$.

$$\rho_T^e(x) = \frac{1}{T} \int_0^T \delta(X_t - x) \, dt \quad \text{and} \quad j_T^e(x) = \frac{1}{T} \int_0^T \delta(X_t - x) \circ dX_t. \tag{38}$$

Roughly speaking, the empirical current (see [28] for a rigorous definition) is the sum of the displacements that the system makes if it is at x. For diffusion processes, with the ergodic assumption the law of large numbers takes the form

$$\rho_T^e \to \rho_{inv} \quad \text{and} \quad j_T^e \to J_{\rho_{inv}}. \tag{39}$$

where the current $J_{\rho_{inv}}$ is defined in relation (23). From the definition (38), we obtain the pathwise constraint[4]

$$\nabla \cdot j_T^e(x) = \frac{1}{T} \left(\delta(X_0 - x) - \delta(X_t - x) \right). \tag{40}$$

Hence, analogously to (37) the large deviation rate function of j_T^e is infinite at any j not fulfilling

$$\nabla . j = 0. \tag{41}$$

2.2.2 Action Functional and Fluctuating Entropy

For time-homogeneous processes, the action functional $\mathbb{W}_T$ is obtained by comparing the trajectorial measure of X_t with the time-reversed trajectorial measure. At the level of trajectories, we introduce the path-wise time inversion[5] R acting on the space of trajectories as

$$R\left[X_0^T\right]_t \equiv \left[X_0^T\right]_{T-t}, \tag{42}$$

where $\left[X_0^T\right]_t \equiv X_t$.

The action functional is defined by the relation

$$\exp(-\mathbb{W}_T) \equiv \frac{R_\star \left(d\mathbb{P}_{L,\mu_0^b,T} \right)}{d\mathbb{P}_{L,\mu_0,T}}. \tag{43}$$

where μ_0^b is the arbitrary initial measure of the reversed trajectory and the push-forward measure can be loosely written as $R_\star \left(d\mathbb{P}_{L,\mu_0^b,T} \right)[x_0^T] = d\mathbb{P}_{L,\mu_0^b,T} \left[R\left[x_0^T\right] \right]$. Due to the freedom in choosing μ_0 and μ_0^b, it is possible to identify the action functional $\mathbb{W}_T$ with different quantities. It becomes the fluctuating total entropy production σ_T for $\mu_0^b(dx) = \mu_T(dx) \equiv \int dy \rho_0(y) P_T^T(y, x) dx$ and the fluctuating entropy increase of the external environment $\mathbb{J}_T$ for $\mu_0(dx) = \mu_0^b(dx) = dx$. The difference between σ_T and $\mathbb{J}_T$ is the boundary term $\ln(\rho_0(x_0)) - \ln(\rho_T(x_T))$, which is the variation of the entropy of the system. We note that names like total entropy production or entropy increase of the external environment become meaningful only if a Markov process is given a clear physical interpretation. In this case these functionals are related to key thermodynamic quantities [47].

---

[4]
$$\int_{\mathcal{E}} dx g(x) \nabla \cdot j_T^e(x) = -\int_{\mathcal{E}} dx j_T^e(x) \cdot \nabla g(x) = -\frac{1}{T} \int_0^T \nabla g(X_t) \circ dX_t$$
$$= \frac{1}{T} \left(g(X_0) - g(Xt) \right), \quad \text{for all functions } g.$$

[5] Here, we do not consider the case where the time inversion acts non-trivially on the space $\mathcal{E}$. For example, such a situation takes place for the non-over-damped Kramers equation [10].

For pure jump processes this action functional is [37,38]

$$\mathbb{W}_T = \ln\left(\rho_0(X_0)\right) - \ln\left(\rho_0^b(X_T)\right) + \sum_{0 \leq s \leq T / X_{s^-} \neq X_{s^+}} \ln\left[\frac{W(X_{t^-}, X_{t^+})}{W(X_{t^+}, X_{t^-})}\right]. \quad (44)$$

For diffusion processes it reads [37]

$$\mathbb{W}_T = \ln\left(\rho_0(X_0)\right) - \ln\left(\rho_0^b(X_T)\right) + 2\int_0^T dt\, \widehat{A_0}\left(X_t\right) \cdot D^{-1}\left(X_t\right) \circ dX_t. \quad (45)$$

3 Transient Fluctuation Relation

Before obtaining the rate function at the level 2.5, let us briefly discuss the transient fluctuation relation. From the definition of the action functional (43) it follows that for all functionals $F_{[0,T]}$,

$$\mathbb{E}_{\mu_0^b, L}\left[F_{[0,T]} \circ R\right] = \mathbb{E}_{\mu_0, L}\left[F_{[0,T]} \exp\left(-\mathbb{W}_T\right)\right]. \quad (46)$$

The backward action functional is defined as

$$\exp\left(-\mathbb{W}_T^b\right) \equiv \frac{R_\star\left(d\mathbb{P}_{L,\mu_0,T}\right)}{d\mathbb{P}_{L,\mu_0^b,T}}. \quad (47)$$

Comparing (43) and (47) we obtain the antisymmetric relation

$$\mathbb{W}_T^b = -\mathbb{W}_T \circ R. \quad (48)$$

For the special case $F_{[0,T]} = \delta(\mathbb{W}_T - W)$, with δ denoting the indicator function, relation (46) becomes the generalized Crooks relation [10,14,37,38,47]

$$\mathbb{P}_{\mu_0^b, L}(\mathbb{W}_T^b = -W) = \exp\left(-W\right) \mathbb{P}_{\mu_0, L}(\mathbb{W}_T = W). \quad (49)$$

From (46), we also deduce the Jarzynski equality [14,31]

$$\mathbb{E}_{\mu_0, L}\left[\exp(-\mathbb{W}_T)\right] = 1. \quad (50)$$

This relation implies two important results. First, (50) and Jensen's inequality gives the second law of thermodynamics $\mathbb{E}_{\mu_0, L}[\mathbb{W}_T] \geq 0$. Second, (50) and the Markov inequality $\mathbb{P}_{\mu_0, L}\left(\exp\left(-\mathbb{W}_T\right) \geq \exp(L)\right) \leq \frac{\mathbb{E}_{\mu_0, L}[\exp(-\mathbb{W}_T)]}{\exp(L)}$ gives an upper bound[6] on the probability of "transient deviations" from the second law, i.e., $\mathbb{P}_{\mu_0, L}\left(\mathbb{W}_T \leq -L\right) \leq \exp\left(-L\right)$.

4 Heuristic Proof for 2.5 Large Deviations

In this section we demonstrate that the joint fluctuation of empirical density and empirical flow for jump processes, and the joint fluctuation of empirical density and empirical current for diffusion processes admit a large deviation regime with an explicit rate function. For jump processes this rate function reads [40]

---

[6] A better upper bound has been obtained in [11] using the classical Martingale inequality.

$$I[\rho, C]$$
$$= \begin{cases} \int dx\, dy \left(\begin{array}{c} -C(x,y) + \rho(x)W(x,y) \\ + C(x,y) \ln \frac{C(x,y)}{\rho(x)W(x,y)} \end{array} \right) & \text{if } \forall x \in \mathcal{E}: \int dy\, C(x,y) = \int dy\, C(y,x) \\ \infty & \text{otherwise,} \end{cases}$$
(51)

while for diffusion processes it is [9,39]

$$I[\rho, j] = \begin{cases} \frac{1}{2} \int dx (j - J_\rho)(\rho D)^{-1}(j - J_\rho) & \text{if } \nabla \cdot j = 0 \\ \infty & \text{otherwise.} \end{cases}$$
(52)

Note that the constraints $\int dy\, C(x, y) = \int dy\, C(y, x)$ and $\nabla \cdot j = 0$ come from (36) and (40), respectively. Formally, by contraction we can obtain the Donsker–Varadhan variational expression for the rate function for the level 2 of large deviations from the level 2.5 rate function. Explicitly, for pure jump processes $I(\rho) = \min_C [I(\rho, C)]$, whereas for diffusion processes $I(\rho) = \min_j [I(\rho, j)]$. These relations lead to

$$I[\rho] = - \inf_{[h]>0} \left[\int dx \rho(x) h^{-1}(x) L[h](x) \right],$$
(53)

where the minimization is over strictly positive functions h. A rigorous proof of this contraction for pure jump processes can be found in [6]. Similarly, a formal contraction implies that the action functional (44) (or (45) for diffusion processes) fulfills a Large Deviation principle. It is also possible to obtain the rate function related to the joint probability of the empirical density $\rho_T^e(x, y)$ and the empirical current $J_T^e(x, y)$ by contraction from (51) [40].

We present two methods to prove (51) and (52): tilting and a spectral method. The proof for jump processes using the spectral method is original. Proofs using tilting for pure jump processes can be found in [40] and for diffusion processes in [39]. Another proof for diffusion processes using the spectral method was obtained in [9]. The novelty in these cases is in our presentation, which highlight the generality of both methods. A third method, which is totally rigorous, for pure jump processes in a countable space related to the contraction of the rate function of the level 3 of large deviations has been recently obtained in [5]. Whereas the proof using the tilting method is more direct, in the spectral method a connection between the rate function and the maximum eigenvalue of a modified generator is established. This connection is often useful for numerical calculations of rate functions.

4.1 Tilting

We consider, for general stochastic processes X_t,[7] the joint large deviation of N observables $\overrightarrow{\omega_t^e} \equiv \left\{ \omega_{t,1}^e, \omega_{t,2}^e, \ldots, \omega_{t,N}^e \right\}$. The trajectorial measure is denoted by $d\mathbb{P}_{\mu_0,T}$ and $\overrightarrow{\omega_{inv}} \equiv \left\{ \omega_{inv,1}, \omega_{inv,2}, \ldots, \omega_{inv,N} \right\}$ represents the typical behavior of $\overrightarrow{\omega_t^e}$, with typical behavior meaning almost sure convergence. If the following two conditions are satisfied then the family of probability measures $\left(\mathbb{P}_{\mu_0,T} \circ \left\{ \overrightarrow{\omega_t^e} \right\}^{-1} \right)_{t \geq 0}$, or equivalently $\overrightarrow{\omega_t^e}$, satisfies a large deviation principle with rate function $I(\overrightarrow{\omega})$, where $\overrightarrow{\omega} = \{\omega_1, \omega_2, \ldots, \omega_N\}$ is the desired untypical behavior.

– *Condition 1* There exists an ergodic tilted process X_t', with trajectorial measure $d\mathbb{P}_{\mu_0,T}'$, such that its typical behavior is $\overrightarrow{\omega_t^e}$.

---

[7] X_t does not need to be Markovian here.

– *Condition 2* For this tilted process, there exists a function I defined by the asymptotic relation $\frac{d\mathbb{P}_{\mu_0,T}}{d\mathbb{P}'_{\mu_0,T}}[X] \sim \exp\left(-TI\left(\overrightarrow{\omega_T^e}\right)\right)$. This means that asymptotically the Radon-Nykodym derivative can be expressed in terms of the N observables $\omega_{t,1}^e, \omega_{t,2}^e, \ldots, \omega_{t,N}^e$.

Note that larger N makes the fulfillment of the first condition harder, while the fulfillment of second condition becomes easier. For a fixed process X_t and a fixed observable $\overrightarrow{\omega_t^e}$, we postulate that the process X_t' exists.

Formal proof From the second condition it follows that

$$
\begin{aligned}
\mathbb{P}_{\mu_0,T}\left[\overrightarrow{\omega_T^e} \simeq \overrightarrow{\omega}\right] &= \int d\mathbb{P}_{\mu_0,T}[X]\,\delta(\overrightarrow{\omega_T^e} - \overrightarrow{\omega}) \\
&= \int d\mathbb{P}'_{\mu_0,T}[X].\frac{d\mathbb{P}_{\mu_0,T}}{d\mathbb{P}'_{\mu_0,T}}[X]\,\delta(\overrightarrow{\omega_T^e} - \overrightarrow{\omega}) \\
&\sim \int d\mathbb{P}'_{\mu_0,T}[X].\exp\left(-TI\left(\overrightarrow{\omega_T^e}\right)\right)\delta(\overrightarrow{\omega_T^e} - \overrightarrow{\omega}) \\
&\sim \exp\left(-TI\left(\overrightarrow{\omega}\right)\right)\int d\mathbb{P}'_{\mu_0,T}[X]\,\delta(\overrightarrow{\omega_T^e} - \overrightarrow{\omega}).
\end{aligned}
\tag{54}
$$

Since the process X_t' is assumed to be ergodic, with the first condition, we obtain

$$
\int d\mathbb{P}'_{\mu_0,T}[X]\,\delta(\overrightarrow{\omega_T^e} - \overrightarrow{\omega}) = \mathbb{P}'_{\mu_0}\left[\overrightarrow{\omega_T^e} \simeq \overrightarrow{\omega}\right] \to 1,
\tag{55}
$$

which, with (54), gives the required Large deviation rate function I. Rigorously, following the same procedure for $\mathbb{P}_{\mu_0}\left[\overrightarrow{\omega_T^e} \in B\left(\overrightarrow{\omega}, \epsilon\right)\right]$, where $B\left(\overrightarrow{\omega}, \epsilon\right)$ an open ball of radius ϵ, the lower bound of the rate function (1) is obtained [5]. We note that these two conditions are not enough for a rigorous proof, which requires a lower and an upper bound on the rate function [5,36].

Examples:

– If X_t is a Markov process and $\overrightarrow{\omega_t^e} \equiv \{\rho_t^e\}$, from the Girsanov relation (18) (or (26) for diffusion processes), we obtain that it is not possible to find a process fulfilling the second condition. The solution to find an explicit rate function is then to increase N.

– If X_t is a pure jump process and $\overrightarrow{\omega_t^e} = \{\rho_t^e, C_t^e\}$, by choosing X' with the transition rates

$$
W'(x, y) = \frac{C(x, y)}{\rho(x)},
\tag{56}
$$

the ergodic behavior of X_t' becomes $\rho'_{inv} = \rho$ and $C_{\rho'_{inv}} = C$, which implies the fulfillment of condition 1. The process X_t' also obeys the conservation law (15), leading to the constraint on the marginal of C in the rate function (51). The Girsanov relation (18) with $V_2(x, y) = \ln\left(\frac{C(x,y)}{\rho(x)W(x,y)}\right)$ becomes

$$
\begin{aligned}
\frac{d\mathbb{P}_{L_{V_2},T}}{d\mathbb{P}_{L,T}}[X] = \exp\Bigg[&\sum_{0 \leq s \leq T / X_{s^-} \neq X_{s^+}} \ln\left(\frac{C(X_{s^-}, X_{s^+})}{\rho(X_{s^-})W(X_{s^-}, X_{s^+})}\right) \\
&- \int_0^T ds \int_{\mathcal{E}} dy \left(\frac{C(X_s, y)}{\rho(X_s)} - W(X_s, y)\right)\Bigg]
\end{aligned}
$$

$$= \exp\left[T\int_{\mathcal{E}^2} dydx \left[C_T^e(x,y)\ln\left(\frac{C(x,y)}{\rho(x)W(x,y)}\right)\right.\right.$$
$$\left.\left. -\rho_T^e(x)\left(\frac{C(x,y)}{\rho(x)} - W(x,y)\right)\right]\right]. \tag{57}$$

Hence, condition 2 is exactly verified at finite time with the rate function I given by (51).

– If X_t is a diffusion process and $\overrightarrow{\omega_t^e} = \{\rho_t^e, j_t^e\}$, condition 1 is fulfilled by choosing X_t' with drift and diffusion coefficient

$$A_0' = \frac{j + \frac{D}{2}\nabla\rho}{\rho} \quad\text{and}\quad A_\alpha' = A_\alpha. \tag{58}$$

This can be shown with the ergodic law (38), which implies

$$\rho_{inv}' = \rho \text{ and } J_{\rho_{inv}'} = j, \tag{59}$$

where ρ_{inv}' is the invariant density of the process X_t'. From the Girsanov relation (26), condition 2 is verified with I given by (52).

– It is possible to apply the tilting method to find the rate function of more informative quantities, e.g., the m-words generalization of empirical flow associated with a pure jump process [12]. The method can also be used to obtain the rate function of the empirical density and flow of pure jump processes that are non-homogeneous and periodic in time [7].

4.2 Spectral Method

4.2.1 Generating Function

The scaled cumulant generating function associated with the vector $\overrightarrow{\omega_t^e}$ is defined as

$$\Lambda[V_1, V_2, \ldots, V_N] = \lim_{T\to\infty}\frac{1}{T}\ln\left(\mathbb{E}_{\mu_0, L}\left[\exp\left(T\sum_{i=1}^{N}\langle\omega_{t,i}^e, V_i\rangle\right)\right]\right) \tag{60}$$

where V_i are objects having the same tensorial nature as $\omega_{t,i}^e$ and $\langle., .\rangle$ denotes the associated canonical scalar product. Assuming that the Gärtner-Ellis theorem [16,17] is still valid in this functional form[8], then if Λ exist and is differentiable for all V_i, the family of probability measures $\left(\mathbb{P}_{\mu_0, T}\circ\left\{\overrightarrow{\omega_t^e}\right\}^{-1}\right)_{t\geq 0}$ satisfies a large deviation principle with rate function

$$I[\omega_1, \omega_2, \ldots, \omega_N] = \sup_{\overrightarrow{V}}\left\{\sum_{i=1}^{N}\langle\omega_i, V_i\rangle - \Lambda[V_1, V_2, \ldots, V_N]\right\}. \tag{61}$$

For pure jump processes, with $\overrightarrow{\omega_t^e} = \{\rho_t^e, C_t^e\}$, the scaled cumulant generating function becomes

$$\Lambda[V_1, V_2]$$
$$= \lim_{T\to\infty}\frac{1}{T}\ln\left(\mathbb{E}_{\mu_0, L}\left[\exp\left(\int_0^T dt V_1(X_t) + \sum_{0\leq s\leq T/X_{s-}\neq X_{s+}} V_2(X_{t-}, X_{t+})\right)\right]\right). \tag{62}$$

---

[8] For a theoretical Physicist point of view, this theorem is a functional Laplace transform followed by a saddle point approximation.

For diffusion processes, with $\overrightarrow{\omega_t^e} = \{\rho_t^e, j_t^e\}$ we obtain

$$\Lambda [V_1, V_2] = \lim_{T \to \infty} \frac{1}{T} \ln \left(\mathbb{E}_{\mu_0, L} \left[\exp \left(\int_0^T dt \, [V_1(X_t) + V_2(X_t) \circ dX_t] \right) \right] \right). \quad (63)$$

4.2.2 Twisted Process

Defining

$$A_T^e \equiv \frac{1}{T} \left(\int_0^T dt \, V_1(X_t) + \sum_{0 \leq s \leq T / X_{s-} \neq X_{s+}} V_2 \left(X_{t-}, X_{t+} \right) \right), \quad (64)$$

relation (17), which is valid for pure jump processes, is equivalent to

$$\mathbb{E}_{L, \mu_0} \left[\exp \left(T A_T^e \right) F \right] = \mathbb{E}_{L_{V_1, V_2}, \mu_0} [F], \quad (65)$$

where F is a generic functional and L_{V_1, V_2} is defined in (16) for pure jump processes. For diffusion processes

$$A_T^e \equiv \frac{1}{T} \left(\int_0^T dt \, [V_1(X_t) + V_2(X_t) \circ dX_t] \right), \quad (66)$$

and relation (30) is equivalent to (65), with L_{V_1, V_2} defined in (31).

The special functional $F = \delta(X_T - y)$ gives the Feynamn–Kac type relation

$$\mathbb{E}_{L, \mu_0} \left[\exp \left(T A_T^e \right) \delta(X_T - y) \right] = \int_{\mathcal{E}} \mu_0(dx_0) \exp \left(T L_{V_1, V_2} \right) (x_0, y). \quad (67)$$

We assume that the twisted operator L_{V_1, V_2} is of Perron–Frobenius type, i.e., there exists a positive gaped principal eigenvalue with maximal real part $\lambda [V_1, V_2]$ related to a unique positive right eigenvector $r [V_1, V_2]$ and a unique positive left eigenvector $l [V_1, V_2]$[9]. Multiplicative factors are fixed by normalization as

$$\int_{\mathcal{E}} l [V_1, V_2] (x) dx = 1 \quad \text{and} \quad \int_{\mathcal{E}} l [V_1, V_2] (x) r [V_1, V_2] (x) dx = 1. \quad (68)$$

It is also assumed that the initial measure fulfills

$$\int_{\mathcal{E}} \mu_0(dx) r [V_1, V_2] (x) < \infty. \quad (69)$$

With this principal eigenvalue and its associated eigenvectors, the semi-group generated by L_{V_1, V_2} can be expanded as

$$\exp \left(T L_{V_1, V_2} \right) (x, y) = \exp \left(T \lambda_{V_1, V_2} \right) \left(r [V_1, V_2] (x) l [V_1, V_2] (y) + O \left(\exp \left(-t \Delta_{V_1, V_2} \right) \right) \right), \quad (70)$$

---

[9] These properties follow from the Krein–Rutman theorem [35], which, however, requires that the operator L_{V_1, V_2} is compact. For a uniformly elliptic operator in divergent form as the generator of a diffusion process, a version of the Krein–Rutman theorem is proven, for example, in [27, Chap. 6.5.2], where the hypothesis are: $\mathcal{E}$ is bounded, open and connected; $\partial \mathcal{E}$ is smooth; D and $\widehat{A_0}$ are smooth; $L_{V_1, V_2} [1] \geq 0$ on $\mathcal{E}$. Strictly speaking, the theorem is not valid if, for example, $\mathcal{E}$ is not bounded, with the extension for an unbounded $\mathcal{E}$ being a difficult and contemporary problem [4]. Even though we are not aware of proof for unbounded $\mathcal{E}$ in the mathematics literature, more sophisticated related results do exist, as for example in [43, Chap. 4.11]. From a physicist perspective, if the drift of the process is sufficiently confining then the result for bounded $\mathcal{E}$ case should also be true for unbounded $\mathcal{E}$.

where Δ_{V_1, V_2} is the spectral gap. Combining this last equation with the Feynman–Kac relation (67) we obtain

$$\mathbb{E}_{\mu_0, L}\left[\exp\left(T A_T^e\right)\delta(X_T - y)\right] = \exp\left(T\lambda_{V_1, V_2}\right)\int_{\mathcal{E}}\mu_0(dx_0)\,(r\,[V_1, V_2]\,(x)l\,[V_1, V_2]\,(y)$$
$$+ O\left(\exp\left(-t\Delta_{V_1, V_2}\right)\right)). \tag{71}$$

Therefore, the scaled cumulant generating function of A_T^e is

$$\Lambda\,[V_1, V_2] = \lambda\,[V_1, V_2]. \tag{72}$$

We are now ready to prove that (61) allows us to obtain the explicit forms (51) and (52).

4.2.3 Level 2.5 for Jump Processes

Using (72), relation (61), with $\overrightarrow{\omega_t^e} \equiv \{\rho_t^e, C_t^e\}$, becomes

$$I\,[\rho, C] = \sup_{V_1, V_2}\left\{\int_{\mathcal{E}}dx\rho(x)V_1(x) + \int\int_{\mathcal{E}^2}dxdyC(x, y)V_2(x, y) - \lambda\,[V_1, V_2]\right\}. \tag{73}$$

The functions $V_1^\star$ and $V_2^\star$ extremizing the above expression are then obtained by solving the equations

$$\left.\frac{\delta\lambda\,[V_1, V_2]}{\delta V_1(x)}\right|_{V_1^\star, V_2^\star} = \rho(x) \quad \text{and} \quad \left.\frac{\delta\lambda\,[V_1, V_2]}{\delta V_2(x, y)}\right|_{V_1^\star, V_2^\star} = C(x, y). \tag{74}$$

Furthermore, the normalization (68) and $L_{V_1, V_2}\,[r\,[V_1, V_2]]\,(x) = \lambda\,[V_1, V_2]\,r\,[V_1, V_2]\,(x)$, lead to

$$\int_{\mathcal{E}}l\,[V_1, V_2]\,(x)L_{V_1, V_2}\,[r\,[V_1, V_2]]\,(x)dx = \lambda\,[V_1, V_2]. \tag{75}$$

From (16), applying functional derivatives to (75) we obtain

$$\begin{cases} l\,[V_1, V_2]\,(x)r\,[V_1, V_2]\,(x) = \frac{\delta\lambda[V_1, V_2]}{\delta V_1(x)} \\ l\,[V_1, V_2]\,(x)W(x, y)\left[\exp\left(V_2(x, y)\right)\right]r\,[V_1, V_2]\,(y) = \frac{\delta\lambda[V_1, V_2]}{\delta V_2(x, y)}, \end{cases} \tag{76}$$

which, with (74), leads to

$$\begin{cases} l\left[V_1^\star, V_2^\star\right](x)r\left[V_1^\star, V_2^\star\right](x) = \rho(x) \\ l\left[V_1^\star, V_2^\star\right](x)W(x, y)\left[\exp\left(V_2^\star(x, y)\right)\right]r\left[V_1^\star, V_2^\star\right](y) = C(x, y). \end{cases} \tag{77}$$

From the definitions of $l\,[V_1, V_2]$ and $r\,[V_1, V_2]$ as the left and right eigenvectors of L_{V_1, V_2}, the second equation in (77) implies

$$\begin{cases} \int dxC(x, y) = \left(\lambda\left[V_1^\star, V_2^\star\right] + W\,[1]\,(y) - V_1(y)\right)l\left[V_1^\star, V_2^\star\right](y)r\left[V_1^\star, V_2^\star\right](y) \\ \int dxC(y, x) = \left(\lambda\left[V_1^\star, V_2^\star\right] + W\,[1]\,(y) - V_1(y)\right)l\left[V_1^\star, V_2^\star\right](y)r\left[V_1^\star, V_2^\star\right](y), \end{cases} \tag{78}$$

where the first (second) line is obtained with an integration in x (y). Hence, the constraint (37) is a necessary condition for the extremization and, moreover, using the first equation in (77) we obtain

$$\lambda\left[V_1^\star, V_2^\star\right] + W\left[1\right](y) - V_1^\star(y) = \frac{\int dx C(x, y)}{\rho(y)}.$$ (79)

Finally, from (73) we obtain the rate function (51) as follows,

$$
\begin{aligned}
I\left[\rho, C\right] &= \int\int_{\mathcal{E}^2} dx dy C(x, y) V_2^\star(x, y) - \int_{\mathcal{E}} dx \rho(x)\left(\lambda\left[V_1^\star, V_2^\star\right] - V_1^\star(x)\right) \\
&= \int\int_{\mathcal{E}^2} dx dy C(x, y) \ln\left[\frac{C(x, y)}{l\left[V_1^\star, V_2^\star\right](x) W(x, y) r\left[V_1^\star, V_2^\star\right](y)}\right] \\
&\quad - \int_{\mathcal{E}} dx \rho(x)\left(\frac{\int dy C(y, x)}{\rho(x)} - W\left[1\right](x)\right) \\
&= \int\int_{\mathcal{E}^2} dx dy C(x, y) \ln\left[\frac{C(x, y)}{l\left[V_1^\star, V_2^\star\right](x) r\left[V_1^\star, V_2^\star\right](x) W(x, y)}\right] \\
&\quad + \int\int_{\mathcal{E}^2} dx dy C(x, y) \ln\left[\frac{r\left[V_1^\star, V_2^\star\right](x)}{r\left[V_1^\star, V_2^\star\right](y)}\right] \\
&\quad - \int_{\mathcal{E}} dx \rho(x)\left(\frac{\int dy C(y, x)}{\rho(x)} - \int dy W(x, y)\right) \\
&= \int\int_{\mathcal{E}^2} dx dy C(x, y) \ln\left[\frac{C(x, y)}{\rho(x) W(x, y)}\right] \\
&\quad - \int_{\mathcal{E}} dx \rho(x)\left(\frac{\int dy C(y, x)}{\rho(x)} - \int dy W(x, y)\right) \\
&\quad + \int_{\mathcal{E}} dx \ln\left[r\left[V_1^\star, V_2^\star\right](x)\right] \int_{\mathcal{E}} dy \left(C(x, y) - C(y, x)\right).
\end{aligned}
$$ (80)

Passing from the first to the second line we used $V_2^\star(x, y) = \ln\left[\frac{C(x,y)}{l[V_1^\star, V_2^\star](x) W(x,y) r[V_1^\star, V_2^\star](y)}\right]$, which follows from (77), and Eq. (79). Moreover, in the last equality we used the first equation in (77) and the last term is zero due to the constraint (37), thus leading to expression (51) for the rate function.

4.2.4 Level 2.5 for Diffusion Processes

Using (72), for diffusion processes (61) becomes

$$I\left[\rho, j\right] = \sup_{V_1, V_2}\left\{\int_{\mathcal{E}} dx \rho(x) V_1(x) + j(x) . V_2(x) - \lambda\left[V_1, V_2\right]\right\}.$$ (81)

The following three change of variables lead to the final expression (52).

- First, $(V_1, V_2) \to \left(V_1' = \ln\left(r\left[V_1, V_2\right]\right), V_2\right)$, leading to

$$I\left[\rho, j\right] = \sup_{V_1', V_2}\left\{\int_{\mathcal{E}} dx \rho(x)\left(-\exp\left(-V_1'(x)\right) L_{0, V_2}\left[\exp\left(V_1'\right)\right](x)\right) + j(x) . V_2(x)\right\}.$$ (82)

This is proved in Appendix 1. Note that $\ln\left(r\left[V_1, V_2\right]\right)$ is well defined because $r\left[V_1, V_2\right]$ is positive (from the Perron–Frobenius theorem).

- Second, $(V_1', V_2) \to (V_1', V_2' = V_2 + \nabla V_1')$, leading to

$$
I[\rho, j] = -\inf_{V_1'} \left(\int_{\mathcal{E}} dx j(x).\nabla V_1' \right)
$$
$$
-\inf_{V_2''} \left(\int_{\mathcal{E}} dx \left[(V_2' - (\rho D)^{-1}(j - J_\rho)) \frac{\rho D}{2} (V_2' - (\rho D)^{-1}(j - J_\rho)) \right] \right)
$$
$$
+ \int dx (j - J_\rho) \frac{(\rho D)^{-1}}{2} (j - J_\rho). \tag{83}
$$

This is proved in Appendix 2.
- Third, $(V_1', V_2') \to (V_1', V_2'' = V_2' - (\rho D)^{-1}(j - J_\rho))$, finally gives

$$
I[\rho, j] = -\inf_{V_1'} \left(\int_{\mathcal{E}} dx j(x).\nabla V_1' \right)
$$
$$
-\inf_{V_2''} \left(\int_{\mathcal{E}} dx V_2''(x) \frac{\rho D}{2}(x) V_2''(x) \right) + \int dx (j - J_\rho) \frac{(\rho D)^{-1}}{2} (j - J_\rho). \tag{84}
$$

The first term vanishes with fulfillment of the constraint (41) and is $-\infty$ otherwise, while the second term vanishes. This last equation gives the final form (52).

5 Stationary Fluctuation Relation at the Level 2.5

We now consider the fluctuating entropy $\mathbb{J}_T$, which is obtained from the action functional (43) setting $\mu_0(dx) = \mu_0^b(dx) = dx$. We define the function

$$
\mathbb{J}_T/T = w(\rho_T^e, C_T^e) \quad \text{and} \quad \mathbb{J}_T/T = w(\rho_T^e, j_T^e), \tag{85}
$$

for pure jump and diffusion processes, respectively. From formulas (44) and (45), this function reads

$$
w(\rho, C) = \int dx dy C(x, y) \ln \left[\frac{W(x, y)}{W(y, x)} \right] \quad \text{and} \quad w(\rho, j) = 2 \int dx \widehat{A_0}(x).D^{-1}(x) j(x), \tag{86}
$$

The choice $F_{[0,T]} = \delta(\rho_T^e - \rho, C_T^e - C)$ for pure jump and $F_{[0,T]} = \delta(\rho_T^e - \rho, j_T^e - j)$ for diffusion processes in (46) gives the finite time relation

$$
\begin{cases}
\mathbb{P}_{\mu_0, L}(\rho_T^e = \rho, C_T^e = C^t) = \exp(-T w(\rho, C)) \, \mathbb{P}_{\mu_0, L}(\rho_T^e = \rho, C_T^e = C) \\
\mathbb{P}_{\mu_0, L}(\rho_T^e = \rho, j_T^e = -j) = \exp(-T w(\rho, j)) \, \mathbb{P}_{\mu_0, L}(\rho_T^e = \rho, j_T^e = C)
\end{cases}, \tag{87}
$$

where we used the general relations $\rho_T^e \circ R = \rho_T^e$, $j_T^e \circ R = -j_T^e$, and $C_T^e \circ R = (C_T^e)^t$, with the index t indicating transposition. With the rate function for the large deviations at the level 2.5 obtained in the last section, the large time asymptotic of both sides of the previous relation becomes the stationary fluctuation relation at level 2.5

$$
I(\rho, C^t) = w[\rho, C] + I(\rho, C) \quad \text{and} \quad I(\rho, -j) = w[\rho, j] + I(\rho, j). \tag{88}
$$

From this relation, with the contraction $I(w) = \min_{w(\rho, C)=w}[I(\rho, C)]$ (or $I(w) = \min_{w(\rho, j)=w}[I(\rho, j)]$ for diffusion processes), we obtain the stationary fluctuation relation

$$I(-w) = I(w) + w. \tag{89}$$

This symmetry on the rate function of $\mathbb{J}_T$ is the GCEM symmetry. This relation can also be obtained from the transient fluctuation relation (49). We note that currents with such a symmetry in the rate function that are different from the fluctuating entropy $\mathbb{J}_T$ have been found in [1–3]. Investigating, the relation between this symmetric non-entropic currents and large deviations at the level 2.5 would be interesting.

Appendix 1: Proof of (82)

We prove relation (82) from relation (81). Writing

$$\left(L_{0, V_2} + V_1\right) r\left[V_1, V_2\right](x) = \lambda\left[V_1, V_2\right] r\left[V_1, V_2\right](x), \tag{90}$$

we obtain

$$V_1 - \lambda\left[V_1, V_2\right] = -\left(r\left[V_1, V_2\right](x)\right)^{-1} L_{0, V_2}\left(r\left[V_1, V_2\right]\right)(x). \tag{91}$$

With this last equation (81) becomes

$$
\begin{aligned}
I\left[\rho, j\right] &= \sup_{V_1, V_2}\left(\int_{\mathcal{E}} dx \rho(x)\left(V_1(x) - \lambda\left[V_1, V_2\right]\right) + j(x) \cdot V_2(x)\right) \\
&= \sup_{V_1, V_2}\left(\int_{\mathcal{E}} dx \rho(x)\left(-\left(r\left[V_1, V_2\right](x)\right)^{-1} L_{0, V_2}\left(r\left[V_1, V_2\right]\right)(x)\right) + j(x) \cdot V_2(x)\right) \\
&= \sup_{V_1', V_2}\left(\int_{\mathcal{E}} dx \rho(x)\left(-\exp\left(-V_1'(x)\right) L_{0, V_2}\left[\exp\left(V_1'\right)\right](x)\right) + j(x) \cdot V_2(x)\right),
\end{aligned}
\tag{92}
$$

where $V_1' = \ln r(V_1, V_2)$.

Appendix 2: Proof of (83)

The goal here is to prove relation (83) from (82). From a direct calculation we obtain

$$\exp\left(-V_1'\right) L_{0, V_2}\left(\exp V_1'\right) = L_{0, V_2 + \nabla V_1'}[1]. \tag{93}$$

Relation (82) then becomes

$$
\begin{aligned}
I\left[\rho, j\right] &= \sup_{V_1', V_2}\left(\int_{\mathcal{E}} dx\left(j(x) \cdot V_2(x) - \rho(x) L_{0, V_2 + \nabla V_1'}[1](x)\right)\right) \\
&= \sup_{V_1', V_2'}\left(\int_{\mathcal{E}} dx\left(-j(x) \cdot \nabla V_1' + j(x) \cdot V_2'(x) - \rho(x) L_{0, V_2'}[1](x)\right)\right) \\
&= -\inf_{V_1'}\left(\int_{\mathcal{E}} dx j(x) \cdot \nabla V_1'\right) + \sup_{V_2'}\left(\int_{\mathcal{E}} dx\left(j(x) \cdot V_2'(x) - \rho(x) L_{0, V_2'}[1](x)\right)\right).
\end{aligned}
\tag{94}
$$

We obtain the final relation (83) with $L_{0,V_2'}[1] = \widehat{A_0} \cdot V_2' + V_2' \cdot \frac{D}{2} \cdot V_2' + \nabla \cdot \left(\frac{D}{2} \cdot V_2'\right)$ and the algebraic manipulation

$$
\begin{aligned}
&\int_{\mathcal{E}} dx \left(j(x) \cdot V_2'(x) - \rho(x) L_{0,V_2'}[1](x) \right) \\
&= \int_{\mathcal{E}} dx \left(j(x) \cdot V_2'(x) - \rho(x) \left(\widehat{A_0} \cdot V_2' + V_2' \cdot \frac{D}{2} \cdot V_2' + \nabla \cdot \left(\frac{D}{2} \cdot V_2' \right) \right) \right) \\
&= \int_{\mathcal{E}} dx\, j(x) \cdot V_2'(x) - \left[\rho(x) V_2' \cdot \frac{D}{2} \cdot V_2' + V_2' \cdot \left(\widehat{A_0}\rho(x) - \frac{D}{2} \cdot \nabla\rho + j \right) \right] \\
&= \int_{\mathcal{E}} dx \left[-\rho(x) V_2' \cdot \frac{D}{2} \cdot V_2' + V_2' \cdot \left(j - J_\rho \right) \right] \\
&= -\int_{\mathcal{E}} dx \Bigg[\left(V_2' - (\rho D)^{-1} \left(j - J_\rho \right) \right) \frac{\rho D}{2} \left(V_2' - (\rho D)^{-1} \left(j - J_\rho \right) \right) \\
&\qquad - \left(j - J_\rho \right) \frac{(\rho D)^{-1}}{2} \left(j - J_\rho \right) \Bigg],
\end{aligned}
\tag{95}
$$

which included formal integration by parts.

References

1. Barato, A.C., Chetrite, R., Hinrichsen, H., Mukamel, D.: Entropy production and fluctuation relations for a KPZ interface. J. Stat. Mech.: Theor. Exp. P10008 (2010)
2. Barato, A.C., Chetrite, R., Hinrichsen, H., Mukamel, D.: A Gallavotti–Cohen–Evans–Morriss like symmetry for a class of Markov jump processes. J. Stat. Phys. **146**, 294 (2012)
3. Barato, A.C., Chetrite, R.: On the symmetry of current probability distributions in jump processes. J. Phys. A: Math. Theor. **45**, 485002 (2012)
4. Berestycki, H., Rossi, L.: Generalization and properties of the principal eigeinvalue of elliptic operators in unbounded domains. Commun. Pure Appl. Math. **68**, 1014 (2015)
5. Bertini, L., Gabrielli, D., Faggionnato, A.: Large deviations of the empirical flow for continuous time Markov chains. arXiv:1210.2004 (2012)
6. Bertini, L., Faggionato, A., Gabrielli, D.: From level 2.5 to level 2 large deviations for continuous time Markov chains. Markov Process. Relat. Fields **20**, 545 (2014)
7. Bertini, L., Chetrite, R., Faggionato, A., Gabrielli, D.: Large deviation of the empirical flow for Markov chain with time-periodic rates (in preparation)
8. Boltzmann, L.: Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respecktive den Sätzen über das Wärmegleichgewicht. Sitzungsberichte der Akademie der Wissenschaften, Wien, II **75**, 67 (1877)
9. Chernyak, V., Chertkov, M., Malinin, S., Teodorescu, R.: Non-equilibrium thermodynamics and topology of currents. J. Stat. Phys. **137**, 109–147 (2009)
10. Chetrite, R., Gawedzki, K.: Fluctuation relations for diffusion processes. Commun. Math. Phys. **282**, 469–518 (2008)
11. Chetrite, R., Gupta, S.: Two refreshing views of fluctuation theorems through kinematics elements and exponential martingale. J. Stat. Phys. **143**, 543–584 (2011)
12. Chetrite, R., Faggionato, A., Gabrielli, D.: Large deviation principle for the m-order empirical measure and flow of continuous time Markov chain (in preparation)
13. Chung, K.L., Walsh, J.B.: Markov Processes, Brownian Motion, and Time Symmetry, 2nd edn. Springer, Berlin (2005)
14. Crooks, G.E.: Path ensembles averages in systems driven far from equilibrium. Phys. Rev. E **61**, 2361 (2000)
15. de la Fortelle, A.: Large deviation principle for Markov chains in continuous time. Prob. Inf. Transm. **37**, 120 (2001)
16. den Hollander, F.: Large Deviations. American Mathematical Society, Providence (2009)

17. Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications, 2nd edn. Springer, New York (1998)
18. Deuschel, J.D., Stroock, D.: Large Deviations. Academic Press, Boston (1989)
19. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectation for large time, I. Commun. Pure Appl. Math. **28**, 1 (1975)
20. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectation for large time, II. Commun. Pure Appl. Math. **28**, 279 (1975)
21. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectation for large time, III. Commun. Pure Appl. Math. **29**, 389 (1976)
22. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectation for large time, IV. Commun. Pure Appl. Math. **36**, 183 (1983)
23. Ellis, R.S.: Entropy, Large Deviations, and Statistical Mechanics. Springer, New York (1985)
24. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley, New York (2009)
25. Evans, D.J., Cohen, E.G.D., Morriss, G.P.: Probability of second law violations in shearing steady states. Phys. Rev. Lett. **71**, 2401 (1993)
26. Evans, D.J., Searles, D.J.: Equilibrium microstates which generates the second law violating steady state. Phys. Rev. E **50**, 1645 (1994)
27. Evans, L.C.: Partial Differential Equation. American Mathematical Society, Providence (2010)
28. Flandoli, F., Giaquinta, M., Gubinelli, M., Tortorelli, V.M.: Stochastic currents. Stoch. Process. Appl. **155**, 1583 (2005)
29. Gallavotti, G., Cohen, E.G.D.: Dynamical ensemble in a stationary state. J. Stat. Phys. **80**, 931–970 (1995)
30. Hormander, L.: Hypoelliptic second ordrer differential equations. Acta Math. **119**, 147 (1967)
31. Jarzynski, C.: A nonequilibrium equality for free energy differences. Phys. Rev. Lett. **78**, 2690 (1997)
32. Kesidis, G., Walrand, J.: Relative entropy between Markov transition rate matrices. IEEE Trans. Inf. Theory **39**, 10561057 (1993)
33. Kirchkoff, G.: Über die Auflosung der Gleichungen, auf welche man bei der Untersuchungder der linearen Vertheilung galvanischer Strome gefuhrt wird. Ann. Phys. (Berlin) **148**, 497 (1847)
34. Kipnis, C., Landim, C.: Scaling Limits of Interacting Particle Systems. Springer, Berlin (1999)
35. Krein, M.G., Rutman, M.A.: Linear operators leaving a cone in a Banach space. Uspekhi Matematicheskikh Nauk **3**, 3 (1948)
36. Kusuoka, S., Kuwada, K., Tamura, Y.: Large deviation for stochastic line integrals as L^p-currents. Probab. Theory Relat. Fields **147**, 649 (2010)
37. Lebowitz, J., Spohn, H.: A Gallavotti–Cohen type symmetry in the large deviation functional for stochastic dynamics. J. Stat. Phys. **95**, 333 (1999)
38. Maes, C.: The fluctuation theorem as a Gibbs property. J. Stat. Phys. **95**, 367 (1999)
39. Maes, C., Netocny, K., Wynants, B.: Steady state statistics of driven diffusions. Physica A **387**, 2675 (2008)
40. Maes, C., Netocny, K.: The canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states. Europhys. Lett. **82**, 30003 (2008)
41. Malliavin, P.: Stochastic calculus of variations and hypoelliptic operators. In: Proc. Intern. Symp. SDE (Kyoto), Wiley, New York (1978)
42. Oono, Y.: Large deviation and statistical physics. Prog. Theoret. Phys. Suppl. **99**, 165 (1989)
43. Pinsky, R.G.: Positive Harmonic Functions and Diffusion. Cambridge University Press, Cambridge (1995)
44. Revuz, D., Yor, M.: Continuous Martingles and Brownian Motion, 3rd edn. Springer, Berlin (1999)
45. Risken, H.: The Fokker–Planck Equation: Methods of Solution and Applications. Springer, New York (1996)
46. Rogers, L.C.G., Williams, D.: Diffusions, Markov Processes, and Martingales. Cambridge University Press, Cambridge (2000)
47. Seifert, U.: Stochastic thermodynamics, fluctuation theorems, and molecular machines. Rep. Prog. Phys. **75**, 126001 (2012)
48. Stroock, D., Varadhan, S.R.S.: Multidimensional Diffusion Processes. Springer, Berlin (1979)
49. Touchette, H.: The large deviation approach to statistical mechanics. Phys. Rep. **478**, 1 (2009)
50. Varadhan, S.R.S.: Large Deviations and Applications. SIAM, Philadelphia (1984)