A formal* view on 2.5 large deviations and fluctuation relations

A C Barato\textsuperscript{1} and R Chetrite\textsuperscript{2}

\textsuperscript{1} II. Institut für Theoretische Physik, Universität Stuttgart, Stuttgart 70550, Germany
\textsuperscript{2} Laboratoire J. A. Dieudonné, UMR CNRS 6621, Université de Nice Sophia-Antipolis, France

August 22, 2014

Abstract

We obtain the rate function for the level 2.5 of large deviations for pure jump and diffusion processes. This result is proved with two methods: tilting, for which a tilted processes 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 \cite{20, 21, 23} 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 \cite{35, 40} and in modern probability \cite{15, 16, 17, 19, 41}, as evidenced by the Abel Prize awarded to S.R.S Varadhan in 2007.

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 $A$

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

(1)

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

$$\mu_T(dx) \sim \exp(-TI(x)) \, dx.$$  (2)

Historically, large deviation theory originated in the nineteenth century with pioneering works in statistical mechanics \cite{7}. One of the most important contributions to large deviation

\begin{footnotesize}
*“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)
\end{footnotesize}
theory was the general approach for Markov processes developed by Donsker and Varadhan \[18\]. 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 posess 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 \((X_t, X_{t+1}, ..., X_{t+k-1})\) 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 \[26\], 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 \[14\], 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 \[33, 32\], and by Chernyak et al. \[8\], respectively. Finally, Bertini et al. \[4\] succeeded in proving rigorously the level 2.5 for a pure jump processes in a countable space. For diffusion processes no rigorous proof is available.

The purpose of our contribution is to present the level 2.5 of large deviations for continuous time processes. Moreover, we discuss its connection with fluctuation relations. Results derived here concerning level 2.5 of large deviations can be found in \[14, 33, 32, 4\] and in \[30, 10, 9\] for fluctuation relations. However, our presentation and some of the proofs for the level 2.5 are 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 subsection 2.1 we recall basic concepts of Markov processes like transition probability, generator, stationary and equilibrium states, and trajectorial measure. In subsection 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 section 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 subsection 4.1 we use the tilting method to obtain the rate function characterizing the level 2.5. This proof is related to results from \[33, 32\], but the presentation given here is original and shows the generality of this method. Subsection 4.2 contains the spectral method. While the proof for pure jump processes using this spectral method is original to our knowledge, for diffusion processes it can
be found in \[8\]. In comparison to this reference, we expurgate the field theoretical language by using the Girsanov lemma. Finally, in section \[5\] 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 \[12, 36, 37, 39\], considering continuous time Markov processes \(X_t\) taking values in a state space \(E\), which can be continuous, as for example \(\mathbb{R}^d\), or a counting space.

2.1.1 Elements of 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 Chapman-Kolmogorov rule

\[
\int_E P_t(x, dy) P_t(y, dz) = P_{t+t}(x, dz),
\]

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

\[
P_t[f](x) \equiv \int_E P_t(x, dy) f(y).
\]

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

\[
\partial_t P_t = P_t \circ L \quad \text{and} \quad \partial_t P_t = L \circ P_t,
\]

respectively. The symbol \(\circ\) means composition of operators and the initial condition is \(P_0 = I\), where \(I\) 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 \(E\).

The time evolution of the instantaneous one point measure \(\mu_t(dx) = \int_E \mu_0(dx_0) P_t(x_0, dy)\) 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 formal adjoint of \(L\) with respect to the Lebesgue or counting measure. Another fundamental hypothesis is that there exists a unique invariant probability measure \(\mu_{inv}\) satisfying

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

The process is said to be in equilibrium if \(\mu_{inv}\) satisfies the detailed balance relation

\[
\mu_{inv}(dx) P_t(x, dy) = \mu_{inv}(dy) P(y, dx).
\]

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 \[24, 34\] for a diffusion process, leading to the Fokker-Planck equation \(\partial_t \mu_t = L^\dagger[\mu_t]\), where \(L^\dagger\) is the formal adjoint of \(L\) with respect to the Lebesgue or counting measure. Another fundamental hypothesis is that there exists a unique invariant probability measure \(\mu_{inv}\) satisfying

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

The process is said to be in equilibrium if \(\mu_{inv}\) satisfies the detailed balance relation

\[
\mu_{inv}(dx) P_t(x, dy) = \mu_{inv}(dy) P(y, dx).
\]

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 \[24, 34\] for a diffusion process, leading to the Fokker-Planck equation \(\partial_t \mu_t = L^\dagger[\mu_t]\), where \(L^\dagger\) is the formal adjoint of \(L\) with respect to the Lebesgue or counting measure. Another fundamental hypothesis is that there exists a unique invariant probability measure \(\mu_{inv}\) satisfying

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

The process is said to be in equilibrium if \(\mu_{inv}\) satisfies the detailed balance relation

\[
\mu_{inv}(dx) P_t(x, dy) = \mu_{inv}(dy) P(y, dx).
\]

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 \[24, 34\] for a diffusion process, leading to the Fokker-Planck equation \(\partial_t \mu_t = L^\dagger[\mu_t]\), where \(L^\dagger\) is the formal adjoint of \(L\) with respect to the Lebesgue or counting measure. Another fundamental hypothesis is that there exists a unique invariant probability measure \(\mu_{inv}\) satisfying

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

The expression \(\rho_{inv} \circ L \circ \rho_{inv}^{-1}\) must be understood as the composition of three operators, first the operator multiplication by \(\rho_{inv}^{-1}\), second the operator \(L\) and last the operator multiplication by \(\rho_{inv}\). This type of notation is recurrently used in the article.
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 $X_t^T: t \in [0,T] \to X_1$. It is a random variable in the space of trajectories $D(\mathbb{R}^+, \mathcal{E})$. This trajectorial measure $dP_{L,\mu_0,T}[x_0^T]$, where $\mu_0$ is the initial measure, is roughly the probability of the trajectory $X_t^T = x_0^T$. The expectation of an arbitrary functional $F[x_0^T]$ of the trajectories is then written as,

$$E_{L,\mu_0} [F] = \int F[x_0^T] dP_{L,\mu_0,T}[x_0^T].$$

The finite time distribution $P_1$ is sufficient to characterize $dP_L$, more precisely, equation (9) may be rewritten as

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

for the cylindrical functional

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

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

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

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

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

A relevant quantity in this paper is the current associated with the density $\rho_t$,

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

From equation (10), the current associated with the invariant density is conserved,

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

At the trajectory level it is possible to compare the trajectorial measure [6] 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\(^2\)

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

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

4
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 [25, Proposition 2.6] and the Feynmann-Kac relation [36, 37] it then follows that $dP_{L_{V_1,V_2,\mu_0,T}}$ and $dP_{L,\mu_0,T}$ are absolutely continuous, and the explicit Radon Nykodym derivative is given by
\begin{equation}
\frac{dP_{L_{V_1,V_2,\mu_0,T}}}{dP_{L,\mu_0,T}}[x] = \exp \left( \sum_{0 \leq s \leq T} V_2(x_{s-}, x_{s+}) + \int_0^T ds V_1(x_s) \right),
\end{equation}
where $x_{s-} \equiv \lim_{s \to 0} x_{s-}$ and $x_{s+} \equiv \lim_{s \to 0} x_{s+}$. Hence, the sum $\sum_{0 \leq s \leq T} V_2(x_{s-}, x_{s+})$ is over all jumps in the trajectory $x^T$. 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 \cite{22} becomes
\begin{equation}
\frac{dP_{L_{V_2,\mu_0,T}}}{dP_{L,\mu_0,T}}[x] = \exp \left( \sum_{0 \leq s \leq T} V_2(x_{s-}, x_{s+}) - \int_0^T ds (W \exp (V_2) - W)[1](x_s) \right),
\end{equation}
where $L_{V_2}$ is the conservative generator obtained from \cite{15} by setting $V_1 = (W)[1] - (W \exp (V_2))[1]$.

2.1.3 Diffusion processes
A diffusion process $X_t$ in a $d$-dimensional manifold is described by the differential equation
\begin{equation}
\begin{split}
dX = A_0(X)dt + \sum_{\alpha} A_{\alpha}(X) \circ dW_{\alpha}(t),
\end{split}
\end{equation}
where the drift $A_0$ and the diffusion coefficient $A_{\alpha}$ are arbitrary smooth vector fields on $\mathcal{E}$, $W_{\alpha}$ are independent Wiener processes, and the range of $\alpha$ is model dependent. The symbol $\circ$ indicates that the Stratonovich convention is used. The explicit form of the generator related to \cite{19} is
\begin{equation}
L = A_0 \nabla + \sum_{\alpha} \frac{1}{2} (A_{\alpha}\nabla)^2 = \widehat{A}_0 \nabla + \frac{1}{2} \nabla D \nabla,
\end{equation}
with the modified drift and covariance
\begin{equation}
\widehat{A}_0(x) = A_0(x) - \frac{1}{2} \sum_{\alpha} (\nabla A_{\alpha})(x)A_{\alpha}(x) \quad \text{and} \quad D^{ij}(x) = \sum_{\alpha} A_{\alpha i}(x)A_{\alpha j}(x),
\end{equation}
respectively, where $i = 1, \ldots, d$ and $j = 1, \ldots, d$. The detailed balance relation \cite{21} with respect to the invariant measure $\mu_{inv}(dx) = \rho_{inv}(x)dx$ is then equivalent to $\widehat{A}_0 = \frac{D}{2} \nabla (\ln \rho)$.

A central quantity for diffusion processes is the hydrodynamic probability current $J_{\rho_1}$
\begin{equation}
J_{\rho_1} = \widehat{A}_0 \rho_1 - \frac{D}{2} (\nabla \rho_1).
\end{equation}
The conservation of the current associated with the invariant density then reads
\begin{equation}
\nabla J_{\rho_{inv}} = 0.
\end{equation}

Similar to jump processes, the trajectoryal 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
\begin{equation}
L' \equiv L + B_2 \nabla + B_1,
\end{equation}
with \( B_2 \) and \( B_1 \) arbitrary vector field and scalar, respectively. Combining the Cameron-Martin-Girsanov lemma \cite{37, 39} and the Feynman-Kac relation \cite{36, 37}, it follows that

\[
\frac{dP_{L', \mu_0, T}[x]}{dP_{L, \mu_0, T}[x]} = \exp(V_T[x]),
\]

where

\[
V_T = \int_0^T \left[ D^{-1}(x_u)B_2(x_u) \circ dx_u + \left( B_1(x_u) - D^{-1}(x_u)B_2(x_u) \left( \tilde{A}_0 + \frac{B_2}{2} \right)(x_u) - \frac{1}{2}(\nabla.B_2)(x_u) \right) du \right].
\]

Choosing \( B_2 = DV_2 \) and \( B_1 = V_2 \left( \tilde{A}_0 + \frac{DV_2}{2} \right) + \frac{1}{2}\nabla.(DV_2) + V_1 \), we obtain

\[
V_T = \int_0^T dt \left[ V_1(X_t) + V_2(X_t) \circ dX_t \right].
\]

Equation (25) then becomes

\[
\frac{dP_{L_{V_1, V_2, \mu_0, T}}[X]}{dP_{L, \mu_0, T}}[X] = \exp \left( \int_0^T dt \left[ V_1(X_t) + V_2(X_t) \circ dX_t \right] \right),
\]

where the twisted generator reads

\[
L_{V_1, V_2} = L' = L + DV_2 \nabla + V_2 \left( \tilde{A}_0 + \frac{DV_2}{2} \right) + \frac{1}{2}\nabla.(DV_2) + V_1.
\]

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 on the type of Markov processes considered. For pure jump processes the set of observables is the empirical density \( \rho_{e,T} \) and empirical flow \( C_{e,T} \). They are given by

\[
\rho_{e,T}(x) = \frac{1}{T} \int_0^T \delta(X_t - x) \, dt \quad \text{and} \quad C_{e,T}(x,y) = \frac{1}{T} \sum_{0 \leq s \leq T/X_s \neq X_s} \delta(X_{s+} - x) \delta(X_s+ - y).
\]

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

\[
J_{e,T}(x,y) = C_{e,T}(x,y) - C_{e,T}(y,x).
\]

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

\[
\rho_{e,T} \to \rho_{\text{inv}} \quad \text{and} \quad C_{e,T} \to C_{\rho_{\text{inv}}}.
\]

\(^3\)Rigorously, we should instead define the empirical measure \( \mu_{e,T} = \int_0^T \delta_X dt. \)
where

\[ C_{\rho_{\text{inv}}}(x, y) = \rho_{\text{inv}}(x)W(x, y). \]  

(34)

Moreover, the finite time Kirchhoff’s law \[27\] reads

\[
\int dy C_T^x (x, y) - \int dy C_T^y (y, x) = \frac{1}{T} \sum_{0 \leq s \leq T} \delta (X_{t-} - x) - \frac{1}{T} \sum_{0 \leq s \leq T} \delta (X_{t+} - x)
\]

\[ = \frac{\delta (X_0 - x) - \delta (X_T - x)}{T} = O(1/T). \]  

(35)

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

\[
\int dy C(x, y) = \int dy C(y, x).
\]  

(36)

For diffusion processes, the set of observables is composed by the empirical density \( \rho^e_T \) and the empirical current \( j^e_T \), which read

\[
\rho^e_T (x) = \frac{1}{T} \int_0^T \delta (X_t - x) \, dt \quad \text{and} \quad j^e_T (x) = \frac{1}{T} \int_0^T \delta (X_t - x) \circ dX_t.
\]  

(37)

Roughly speaking, the empirical current (see \[22\] 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

\[
\rho^e_T \to \rho_{\text{inv}} \quad \text{and} \quad j^e_T \to J_{\rho_{\text{inv}}},
\]  

(38)

where the current \( J_{\rho_{\text{inv}}} \) is defined in relation \([22]\). From the definition \[32\], we obtain the pathwise constraint

\[
\nabla . j^e_T(x) = \frac{1}{T} \left( \delta (X_0 - x) - \delta (X_T - x) \right).
\]  

(39)

Hence, analogously to \[32\] the large deviation rate function of \( j^e_T \) is infinite at any \( j \) not fulfilling

\[
\nabla . j = 0.
\]  

(40)

### 2.2.2 Action functional and fluctuating entropy

For time-homogeneous processes, the action functional \( W_T \) 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 \[4\] \( R \) acting on the space of trajectories as

\[
R \left[ X^T_0 \right]_{j} = \left[ X^T_0 \right]_{T - j},
\]  

(41)

where \( \left[ X^T_0 \right]_j \equiv X_t \).

The action functional is defined by the relation

\[
\exp (-W_T) \equiv \frac{R_* \left( dP_{L,\mu,0}T \right)}{dP_{L,\mu,0}T}.
\]  

(42)

\[
\int d\xi g(x) \nabla j^e_T(x) = - \int d\xi j^e_T(x) \nabla g(x) = - \frac{1}{T} \int_0^T \nabla g(X_t) \circ dX_t = \frac{1}{T} \left( g(X_0) - g(X_T) \right), \quad \text{for all functions } g.
\]

\[5\] 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 \[9\].
where \( \mu_0^b \) is the arbitrary initial measure of the reversed trajectory and the push-forward measure can be loosely written as \( R_\star \left( d\mathbb{P}_{L,\mu_0^b,T} \right) \left[ x_0^b \right] = d\mathbb{P}_{L,\mu_0^b,T} \left[ R \left[ x_0^b \right] \right] \). Due to the freedom in choosing \( \mu_0 \) and \( \mu_0^b \), it is possible to identify the action functional \( \mathbb{W}_T \) with different quantities. It becomes the fluctuating total entropy production \( \sigma_T \) for \( \mu_0^b(dx) = \mu_T(dx) = \int dy \rho_0(y) P_T^b(y,x)dx \) and the fluctuating entropy increase of the external environment \( J_T \) for \( \mu_0(dx) = \mu_0^b(dx) = dx \). The difference between \( \sigma_T \) and \( J_T \) is the boundary term \( \ln (\rho_0(x_0)) - \ln (\rho_T(x_T)) \), which is the variation of the entropy of the system. Depending on the physical interpretation of the Markov process, these functionals can be related to key thermodynamic quantities \( [38, 9] \).

Three important results. First, Jensen’s inequality gives the second law of thermodynamics \( \mathbb{E}_{\mu_0,L}[\mathbb{W}_T] \geq 0 \). Second, the Markov inequality \( \mathbb{P}_{\mu_0,L}(\mathbb{W}_T \geq \mathbb{E}(\mathbb{W}_T)) \leq \frac{\mathbb{E}(\mathbb{W}_T)}{\exp(\mathbb{W}_T)} \) gives an upper bound \( \mathbb{P}_{\mu_0,L}(\mathbb{W}_T \leq -L) \leq \exp(-L) \).

\(^6\) A better upper bound has been obtained in \([10]\) using the classical Martingale inequality.
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

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

while for diffusion processes it is [32] [8]

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

Note that the constraints $\int dy C(x, y) = \int dy C(y, x)$ and $\nabla \cdot j = 0$ come from [35] and [59], 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 $I(\rho) = \min_{i} [I(\rho, C)]$, whereas for diffusion processes $I(\rho) = \min_{j} [I(\rho, j)]$. These relations lead to

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

where the minimization is over strictly positive functions $h$. A rigorous proof of this contraction for pure jump processes can be found in [5]. Similarly, a formal contraction implies that the action functional [33] (or [44] 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 $\rho^T_j(x, y)$ and the empirical current $J^T_j(x, y)$ by contraction from [30] [33].

We present two methods to prove [50] and [51]: 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 [33] and for diffusion processes in [32]. Another proof for diffusion processes using the spectral method was obtained in [8]. 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 [4].

4.1 Tilting

We consider, for general stochastic processes $X_t$, the joint large deviation of $N$ observables $\omega^T \equiv \{ \omega^T_{1, j}, \omega^T_{2, j}, \ldots, \omega^T_{N, j} \}$. The trajectory measure is denoted by $d\tilde{P}_{\mu_0, T}$ and $\omega^T_{inv} \equiv \{ \omega_{inv, 1}, \omega_{inv, 2}, \ldots, \omega_{inv, N} \}$ represents the typical behavior of $\omega^T_t$.

If the following two conditions are satisfied then the family of probability measures $\left( P_{\mu_0, T} \circ \left\{ \omega^T_t \right\}^{-1} \right)_{t \geq 0}$, or equivalently $\omega^T_t$, satisfies a large deviation principle with rate function $I(\omega)$, where $\omega = \{ \omega_1, \omega_2, \ldots, \omega_N \}$ is the desired untypical behavior.

- **Condition 1**: There exists a tilted process $X'_t$, with trajectory measure $d\tilde{P'}_{\mu_0, T}$, such that its typical behavior is $\omega^T_{1, j}$.

---

$^7$ $X_t$ does not need to be Markovian here.
• **Condition 2:** For this tilted process, there exists a function $I$ defined by the asymptotic relation
\[
\frac{d\mathbb{P}_{\mu_0,T}}{d\mathbb{P}_{\mu_0,T}}[X] \sim \exp \left(-TI\left(\overrightarrow{\omega}_T\right)\right).
\]
This means that asymptotically the Radon-Nykodym derivative can be expressed in terms of the $N$ observables $\omega_{t,1}^x, \omega_{t,2}^x, \ldots, \omega_{t,N}^x$.

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 $\overrightarrow{\omega}_T$, we postulate that the process $X'$ verifying these two properties is unique, if it exists.

**Formal proof:** From the second condition it follows that
\[
\mathbb{P}_{\mu_0}\left[\overrightarrow{\omega}_T \simeq \overrightarrow{\omega}\right] = \int d\mathbb{P}_{\mu_0,T}[X] \cdot \delta(\overrightarrow{\omega}_T - \overrightarrow{\omega}) = \int d\mathbb{P}_{\mu_0,T}[X] \cdot \frac{d\mathbb{P}_{\mu_0,T}}{d\mathbb{P}_{\mu_0,T}}[X] \cdot \delta(\overrightarrow{\omega}_T - \overrightarrow{\omega})
\]

\[
\sim \int d\mathbb{P}_{\mu_0,T}[X] \cdot \exp \left(-TI\left(\overrightarrow{\omega}_T\right)\right) \cdot \delta(\overrightarrow{\omega}_T - \overrightarrow{\omega}) \sim \exp \left(-TI\left(\overrightarrow{\omega}\right)\right) \int d\mathbb{P}_{\mu_0,T}[X] \cdot \delta(\overrightarrow{\omega}_T - \overrightarrow{\omega}).
\]

(53)

Since the process $X'_t$ is assumed to be ergodic, with the first condition, we obtain
\[
\int d\mathbb{P}_{\mu_0,T}[X] \cdot \delta(\overrightarrow{\omega}_T - \overrightarrow{\omega}) = \mathbb{P}_{\mu_0}\left[\overrightarrow{\omega}_T \simeq \overrightarrow{\omega}\right] \to 1,
\]

(54)

which, with (53), gives the required Large deviation rate function $I$. Rigorously, following the same procedure for $\mathbb{P}_{\mu_0}\left[\overrightarrow{\omega}_T \in B(\overrightarrow{\omega},\epsilon)\right]$, where $B(\overrightarrow{\omega},\epsilon)$ an open ball of radius $\epsilon$, the lower bound of the rate function \[1\] is obtained \[1\].

**Examples:**

- If $X_t$ is a Markov process and $\overrightarrow{\omega}_t \equiv \{\rho_t^x\}$, from the Girsanov relation \[18\] (or \[24\] 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 $\overrightarrow{\omega}_t = \{\rho_t^x, C_t^x\}$, by choosing $X'$ with the transition rates
\[
W'(x,y) = \frac{C(x,y)}{\rho(x)},
\]

(55)

the ergodic behavior of $X'_t$ becomes $\rho'_t = \rho$ and $C'_t = C$, which implies the fulfillment of condition 1. The process $X'_t$ also obeys the conservation law \[15\], leading to the constraint on the maximal of $C$ in the rate function \[50\]. The Girsanov relation \[18\] with $V_2(x,y) = \ln \left(\frac{C(x,y)}{\rho(x)W(x,y)}\right)$ becomes
\[
\frac{d\mathbb{P}_{V_2,T}}{d\mathbb{P}_{L,T}}[X] = \exp \left[\sum_{0 \leq s \leq T, X_s \neq X_s^+} \ln \left(\frac{C(X_{s^-},X_{s^+})}{\rho(X_{s^-})W(X_{s^-},X_{s^+})}\right) - \int_0^T ds \int dy \left(\frac{C(x,y)}{\rho(x)} - W(x,y)\right)\right]
\]

(56)

Hence, condition 2 is exactly verified at finite time with the rate function $I$ given by \[50\].
If \( X_t \) is a diffusion process and \( \overrightarrow{\omega}^t \) = \( \{\rho^t, j^t\} \), condition 1 is fulfilled by choosing \( X'_t \) with drift and diffusion coefficient

\[
A'_0 = j + \frac{D}{\rho} \nabla \rho \quad \text{and} \quad A'_\alpha = A_\alpha.
\]  

(57)

This can be shown with the ergodic law (37), which implies

\[
\rho'_{inv} = \rho \quad \text{and} \quad J'\rho_{inv} = j,
\]  

(58)

where \( \rho'_{inv} \) is the invariant density of the process \( X'_t \). From the Girsanov relation (25), condition 2 is verified with \( I \) given by (51).

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 [11]. 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 [6].

4.2 Spectral method

4.2.1 Generating function

The scaled cumulant generating function associated with the vector \( \overrightarrow{\omega}^t \) is defined as

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

(59)

where \( V_i \) are objects having the same tensorial nature as \( \omega^t \) and \( \langle \ldots \rangle \) denotes the associated canonical scalar product. Assuming that the Gartner-Ellis theorem [15, 16] is still valid in this functional form, then if \( \Lambda \) exist and is differentiable for all \( V_i \), the family of probability measures \( \left\{ P_{\mu_0, T} \circ \left\{ \overrightarrow{\omega}^t \right\}^{-1} \right\}_{t \geq 0} \) satisfies a large deviation principle with rate function

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

(60)

For pure jump processes, with \( \overrightarrow{\omega}^t = \{\rho^t, C^t\} \), the scaled cumulant generating function becomes

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

(61)

For diffusion processes, with \( \overrightarrow{\omega}^t = \{\rho^t, j^t\} \) we obtain

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

(62)

*For a theoretical Physicist point of view, it is a functional Laplace transform followed by a saddle point approximation.*
where multiplicative factors are fixed by normalization, i.e., function of \( A \) from relations (59), (66) and the Krein-Rutman theorem, the scaled cumulant generating function (68), relation (60), with \( r \) eigenvector \( R \) of Perron-Frobenius type: there exists a positive principal eigenvalue with maximal real part \( \lambda \) [\( V_1, V_2 \)], which follows from the Krein-Rutman theorem [29]. Moreover, up to multiplicative factors, there is an unique positive right eigenvector \( r \) [\( V_1, V_2 \)] and an unique positive left eigenvector \( l \) [\( V_1, V_2 \)] related to \( \lambda \) [\( V_1, V_2 \)]. The multiplicative factors are fixed by normalization, i.e.,

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

From relations (69), (66) and the Krein-Rutman theorem, the scaled cumulant generating function of \( A_T^r \) is

\[
\Lambda \left[ V_1, V_2 \right] = \lambda \left[ V_1, V_2 \right].
\]

We are now ready to prove that (60) allows us to obtain the explicit forms (60) and (61).

4.2.3 Level 2.5 for jump processes

Using (68), relation (60), with \( \omega_T^2 \equiv \{ \rho^2, C^2 \} \), becomes

\[
I \left[ \rho, C \right] = \sup_{\mathcal{V}_1, \mathcal{V}_2} \left\{ \int_\mathcal{E} dx \rho(x) V_1(x) + \int_\mathcal{E} \int_\mathcal{E} dx dy C(x, y) V_2(x, y) - \lambda \left[ V_1, V_2 \right] \right\}.
\]

The functions \( V_1^* \) and \( V_2^* \) extremizing the above expression are then obtained by solving the equations

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

Furthermore, the normalization (69) and \( L_{V_1, V_2} \) [\( r \left[ V_1, V_2 \right] \)] (x) = \( \lambda \) [\( V_1, V_2 \)] \( r \left[ V_1, V_2 \right] \) (x), lead to

\[
\int_\mathcal{E} l \left[ V_1, V_2 \right] (x)L_{V_1, V_2} \left[ r \left[ V_1, V_2 \right] \right] (x) dx = \lambda \left[ V_1, V_2 \right].
\]
From (10), applying functional derivatives to (11) we obtain
\[
\begin{align*}
I [V_1, V_2] (x) r [V_1, V_2] (x) = \frac{\delta [V_1, V_2]}{\delta V_1 (x)} , \\
I [V_1, V_2] (x) W (x, y) \exp (V_2 (x, y)) r [V_1, V_2] (y) = \frac{\delta [V_1, V_2]}{\delta V_2 (x, y)} ,
\end{align*}
\tag{72}
\]
which, with (70), leads to
\[
\begin{align*}
I [V_1^*, V_2^*] (x) r [V_1^*, V_2^*] (x) = \rho (x) \\
I [V_1^*, V_2^*] (x) W (x, y) \exp (V_2^* (x, y)) r [V_1^*, V_2^*] (y) = C (x, y).
\end{align*}
\tag{73}
\]
From the definitions of $I [V_1, V_2]$ and $r [V_1, V_2]$ as the left and right eigenvectors of $L_{V_1, V_2}$, the second equation in (73) implies
\[
\begin{align*}
\int dx C (x, y) = (\lambda [V_1^*, V_2^*] + W [1] (y) - V (y)) I [V_1^*, V_2^*] (y) r [V_1^*, V_2^*] (y) \\
\int dx C (y, x) = (\lambda [V_1^*, V_2^*] + W [1] (y) - V (y)) I [V_1^*, V_2^*] (y) r [V_1^*, V_2^*] (y) ,
\end{align*}
\tag{74}
\]
where the first (second) line is obtained with an integration in $x$ ($y$). Hence, the constraint (66) is a necessary condition for the extremization and, moreover, using the first equation in (73) we obtain
\[
\lambda [V_1^*, V_2^*] + W [1] (y) - V (y) = \frac{\int dx C (x, y)}{\rho (y)} .
\tag{75}
\]
Finally, from (66) we obtain the rate function (60) as follows,
\[
I [\rho, C] = \int \int \varepsilon^2 C (x, y) V_1^* (x, y) - \int \varepsilon^2 dx p (x) (\lambda [V_1^*, V_2^*] - V_1^* (x))
\]
\[
= \int \int \varepsilon^2 dx dy C (x, y) \ln \left( \frac{C (x, y)}{I [V_1^*, V_2^*] (x) W (x, y) r [V_1^*, V_2^*] (y)} \right) - \int \varepsilon^2 dx p (x) \left( \frac{\varepsilon d C (y, x)}{\rho (x)} - W [1] (x) \right)
\]
\[
= \int \varepsilon^2 \varepsilon^2 dx dy C (x, y) \ln \left( \frac{C (x, y)}{I [V_1^*, V_2^*] (x) W (x, y) r [V_1^*, V_2^*] (y)} \right) + \int \varepsilon^2 dx dy C (x, y) \ln \left( \frac{r [V_1^*, V_2^*] (x)}{r [V_1^*, V_2^*] (y)} \right)
\]
\[
= \int \varepsilon^2 dx dy C (x, y) \ln \left( \frac{C (x, y)}{\rho (x) W (x, y)} \right) - \int \varepsilon^2 dx p (x) \left( \frac{\varepsilon d C (y, x)}{\rho (x)} - \varepsilon d W (x, y) \right)
\]
\[
+ \int \varepsilon^2 dx \ln \left( \frac{r [V_1^*, V_2^*] (x)}{r [V_1^*, V_2^*] (y)} \right) \int \varepsilon^2 dy \left( C (x, y) - C (y, x) \right) .
\tag{76}
\]
Passing from the first to the second line we used $V_2^* (x, y) = \ln \left( \frac{C (x, y)}{I [V_1^*, V_2^*] (x) W (x, y) r [V_1^*, V_2^*] (y)} \right)$, which follows from (73), and equation (75). Moreover, in the last equality we used the first equation in (73) and the last term is zero due to the constraint (59).

4.2.4 Level 2.5 for diffusion Processes

Using (65), for diffusion processes (60) becomes
\[
I [\rho, j] = \sup_{V_1, V_2} \left\{ \int \varepsilon^2 dx p (x) V_1 (x) + J (x) V_2 (x) - \lambda [V_1, V_2] \right\} .
\tag{77}
\]
The following three change of variables lead to final the expression (61).
• First, \((V_1, V_2) \rightarrow (V'_1 = \ln (r \left[V_1, V_2\right]), V_2)\), leading to

\[
I[\rho, j] = \sup_{V'_1, V_2} \left\{ \int dx \rho(x) \left( -\exp (-V'_1(x)) L_{0,V_2} \left[ \exp (V'_1) \right](x) \right) + j(x).V_2(x) \right\}. \tag{78}
\]

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

• Second, \((V'_1, V_2) \rightarrow (V'_1, V'_2 = V_2 + \nabla V'_1)\), leading to

\[
I[\rho, j] = -\inf_{\tilde{V}'_1} \left( \int \tilde{E}(x) \tilde{\nabla}\tilde{V}'_1 \right) - \inf_{\tilde{V}'_2} \left( \int dx \left[ \left( V'_2 - (\rho D)^{-1} (j - J_\rho) \right) \frac{D D}{2} \left( V'_2 - (\rho D)^{-1} (j - J_\rho) \right) \right] \right) + \int dx (j - J_\rho) \frac{(\rho D)^{-1}}{2} (j - J_\rho). \tag{79}
\]

This is proved in appendix B.

• Third, \((V'_1, V'_2) \rightarrow (V'_1, V''_2 = V'_2 - (\rho D)^{-1} (j - J_\rho))\), finally gives

\[
I[\rho, j] = -\inf_{\tilde{V}'_1} \left( \int \tilde{E}(x) \tilde{\nabla}\tilde{V}'_1 \right) - \inf_{\tilde{V}'_2} \left( \int dx \tilde{V}'_2(x) \frac{D D}{2} (\tilde{V}'_2(x)) \right) + \int dx (j - J_\rho) \frac{(\rho D)^{-1}}{2} (j - J_\rho). \tag{80}
\]

The first term vanishes with fulfillment of the constraint \(\left[10\right]\) and is \(-\infty\) otherwise, while the second term vanishes. This last equation gives the final form \(\left[31\right]\).

5 Stationary Fluctuation Relation at the level 2.5

We now consider the fluctuating entropy \(\tilde{J}_T\), which is obtained from the action functional \(\left[92\right]\) setting \(\mu_0(dx) = \mu_0^0(dx) = dx\). We define the function

\[
\tilde{J}_T/T = w(\rho_T, C_T) \quad \text{and} \quad \tilde{J}_T/T = w(\rho_T, j_T), \tag{81}
\]

for pure jump and diffusion processes, respectively. From formulas \(\left[13\right]\) and \(\left[14\right]\), this function reads

\[
w(\rho,C) = \int dx dy \ln \left[ \frac{W(x,y)}{W(y,x)} \right] \quad \text{and} \quad w(\rho, j) = 2 \int dx \tilde{A}_0(x).D^{-1}(x) j(x), \tag{82}
\]

The choice \(F_{0,T} = \delta(\rho_T - \rho, C_T - C)\) for pure jump and \(F_{0,T} = \delta(\rho_T - \rho, j_T - j)\) for diffusion processes in \(\left[45\right]\) gives the finite time relation

\[
\begin{align*}
\hat{P}_{\rho_0,L}(\rho_T = \rho, C_T = C^t) &= \exp(-T w(\rho,C)) P_{\rho_0,L}(\rho_T = \rho, C_T = C) \\
\hat{P}_{\rho_0,L}(\rho_T = \rho, j_T = -j) &= \exp(-T w(\rho,j)) P_{\rho_0,L}(\rho_T = \rho, j_T = C) 
\end{align*} \tag{83}
\]

where we used the general relations \(\rho_T \circ R = \rho_T^T, j_T \circ R = -j_T^T, \) and \(C_T \circ R = (C_T)^T\), 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

\[
I(\rho, C^t) = w[\rho, C] + I(\rho, C) \quad \text{and} \quad I(\rho, -j) = w[\rho, j] + I(\rho, j). \tag{84}
\]
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

$$I(-w) = I(w) + w. \quad (85)$$

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

## Acknowledgments

We thank Krzysztof Gawedzki for helping in the proof presented in section 4.2.4 and Hugo Touchette for carefully reading the manuscript.

### A Proof of (78)

We prove relation $\theta$ from relation $\theta$. Writing

$$(L_0, \nu_2 + V_i) r [V_1, V_2](x) = \lambda [V_1, V_2] r [V_1, V_2](x), \quad (86)$$

we obtain

$$V_1 - \lambda [V_1, V_2] = - (r [V_1, V_2](x))^{-1} L_0, \nu_2 (r [V_1, V_2](x)). \quad (87)$$

With this last equation $\theta$ becomes

$$I[\rho,j] = \sup_{V_1, V_2} \left( \int dx \rho(x) (V_1(x) - \lambda [V_1, V_2]) + j(x).V_2(x) \right) \quad (88)$$

$$= \sup_{V_1, V_2} \left( \int dx \rho(x) \left( - (r [V_1, V_2](x))^{-1} L_0, \nu_2 (r [V_1, V_2](x)) \right) + j(x).V_2(x) \right)$$

$$= \sup_{V_1, V_2} \left( \int dx \rho(x) \left( - \exp (-V_1'(x)) L_0, \nu_2 \left[ \exp (V_1') \right](x) \right) + j(x).V_2(x) \right),$$

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

### B Proof of (79)

The goal here is to prove relation $\theta$ from $\theta$. From a direct calculation we obtain

$$\exp (-V_1') L_0, \nu_2 \left( \exp V_1' \right) = L_0, \nu_2 + \nu V_1'[1]. \quad (89)$$

Relation $\theta$ then becomes

$$I[\rho,j] = \sup_{V_1', V_2} \left( \int dx \left( j(x).V_2(x) - \rho(x)L_0, \nu_2 + \nu V_1'[1](x) \right) \right)$$

$$= \sup_{V_1', V_2} \left( \int dx \left( -j(x).\nabla V_1' + j(x).V_2'(x) - \rho(x)L_0, \nu_2 [1](x) \right) \right)$$

$$= - \inf_{V_1} \left( \int dx j(x).\nabla V_1' \right) + \sup_{V_2} \left( \int dx \left( j(x).V_2'(x) - \rho(x)L_0, \nu_2 [1](x) \right) \right). \quad (90)$$
We obtain the final relation \( \hat{A}_0.V'_2 + \hat{A}_0.V'_2 + \nabla \cdot (\hat{A}_0.D_2.V'_2) \) and the algebraic manipulation

\[
\int_{\xi} dx \left( j(x).V'_2(x) - \rho(x)A_0.V'_2 + \rho(x) \left( \hat{A}_0.D_2.V'_2 + \nabla \cdot (D_2.V'_2) \right) \right)
\]

\[
= \int_{\xi} dx j(x).V'_2(x) - \left[ \rho(x)D_2.V'_2 + \hat{A}_0.D_2.V'_2 + \nabla \cdot (D_2.V'_2) \right]
\]

\[
= -\int_{\xi} dx \left( (V'_2 - (\rho D)^{-1}(j - J_\rho)) \rho D^2 \left( V'_2 - (\rho D)^{-1}(j - J_\rho) \right) - (j - J_\rho) \rho D^2 \right),
\]

which included formal integration by parts.

References

[1] A. C. Barato, R. Chetrite, H. Hinrichsen, and D. Mukamel, Entropy production and fluctuation relations for a KPZ interface, J. Stat. Mech.: Theor. Exp., P10008 (2010).
[2] A. C. Barato, R. Chetrite, H. Hinrichsen, and D. Mukamel, A Gallavotti-Cohen-Evans-Morriss like symmetry for a class of Markov jump processes, J. Stat. Phys. 146, 294 (2012).
[3] A. C. Barato and R. Chetrite, On the symmetry of current probability distributions in jump processes, J. Phys. A: Math. Theor. 45, 485002 (2012).
[4] L. Bertini, D. Gabrielli and A. Faggionato, Large deviations of the empirical flow for continuous time Markov chains, arXiv:1210.2004 (2012).
[5] L. Bertini, A. Faggionato, and D. Gabrielli. From level 2.5 to level 2 large deviations for continuous time Markov chains, arXiv:1212.6908 (2012).
[6] L. Bertini, R. Chetrite, A. Faggionato, and D. Gabrielli, Large Deviation of the empirical flow for Markov Chain with Time-periodic rates, In preparation.
[7] L. Boltzmann, Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht, Sitzungsberichte der Akademie der Wissenschaften, Wien, II, 75, 67 (1877).
[8] V. Chernyak, M. Chertkov, S. Malinin, and R. Teodorescu, Non-Equilibrium Thermodynamics and Topology of Currents, J. Stat. Phys. 137, 109-147 (2009).
[9] R. Chetrite and K. Gawedzki, Fluctuation relations for diffusion processes, Commun. Math. Phys. 282, 469-518 (2008).
[10] R. Chetrite and S. Gupta, Two Refreshing Views of Fluctuation Theorems Through Kinematics Elements and Exponential Martingale, J. Stat. Phys. 143, 543-584 (2011).
[11] R. Chetrite, A. Faggionato and D. Gabrielli, Large Deviation principle for the m-order empirical measure and flow of continuous time Markov Chain, In preparation.
[12] K. L. Chung and J. B. Walsh, Markov Processes, Brownian Motion, and Time Symmetry (Springer, Berlin, 2005), 2nd ed.
[13] G.E. Crooks, Path ensembles averages in systems driven far from equilibrium, Phys. Rev. E 61, 2361 (2000).
A. de la Fortelle, *Large Deviation Principle for Markov Chains in Continuous Time*, Prob. Inf. Transm. **37**, 120 (2001).

F. den Hollander, *Large Deviations*, Am. Math. Soc., Providence (2000).

A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, (Springer-Verlag, New York, 1998), 2nd ed.

J. D. Deuschel and D. Stroock, *Large deviations*, (Academic Press, Boston, 1989).

M.D. Donsker and S.R.S. Varadhan, *Asymptotic evaluation of certain Markov process expectation for large time*, Comm. Pure Appl. Math **I 28**, 1 (1975); **II 28**, 279 (1975); **III 29**, 389 (1976); **IV 36**, 183 (1983).

R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, (Springer, New York, 1985).

D. J. Evans, E. G. D. Cohen, and G. P. and Morriss, *Probability of second law violations in shearing steady states*, Phys. Rev. Lett. **71**, 2401 (1993).

D. J. Evans and D.J. Searles, *Equilibrium microstates which generates the second law violating steady state*, Phys. Rev. E **50**, 1645 (1994).

F. Flandoli, M. Giacinti, M. Gubinelli, and V. M. Tortorelli, *Stochastic currents*, Stoch. Proc. Appl. **155**, 1583 (2005).

G. Gallavotti and E.G.D. Cohen, *Dynamical ensemble in a stationary state*, J. Stat. Phys. **80**, 931-970 (1995).

L. Hormander, *Hypoelliptic second order differential equations*, Acta. Math. **119**, 147 (1967).

C. Jarzynski, *A nonequilibrium equality for free energy differences*, Phys. Rev. Lett. **78**, 2690 (1997).

G. Kesidis and J. Walrand, *Relative entropy between Markov transition rate matrices*, IEEE Trans. Info. Theo **39**, 10561057 (1993).

G. Kirchkoff, *Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird*, Ann. Phys. (Berlin) **148**, 497 (1847).

C. Kipnis and C. Landim, *Scaling limits of interacting particle systems*, (Springer-Verlag, Berlin, 1999).

M. G. Krein and M. A. Rutman, *Linear operators leaving a cone in a Banach space*, Uspekhi Matematicheskikh Nauk **3**, 3 (1948).

J. Lebowitz, H. Spohn, *A Gallavotti-Cohen type symmetry in the large deviation functional for stochastic dynamics*, J. Stat. Phys. **95**, 333 (1999).

C. Maes, *The fluctuation theorem as a Gibbs property*, J. Stat. Phys. **95**, 367 (1999).

C. Maes, K. Netocny, and B. Wynants, *Steady state statistics of driven diffusions*, Physica A **387**, 2675 (2008).

C. Maes and K. Netocny, *The canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states*, Europhys. Lett. **82**, 30003 (2008).

P. Malliavin, in Proc. Intern. Symp. SDE (Kyoto): *Stochastic calculus of variations and hypoelliptic operators*, (Wiley, New York, 1978).

Y. Oono, *Large deviation and statistical physics*, Prog. Theor. Phys. Suppl. **99**, 165 (1989).
[36] D. Revuz and M. Yor, *Continuous martingales and Brownian Motion*, (Springer-Verlag, Berlin, 1999), Third ed.

[37] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes, and Martingales*, (Cambridge University Press, 2000).

[38] U. Seifert, *Stochastic thermodynamics, fluctuation theorems, and molecular machines*, Rep. Prog. Phys. 75, 126001 (2012).

[39] D. Stroock and S. R. S. Varadhan, *Multidimensional Diffusion Processes*, (Springer, Berlin, 1979).

[40] H. Touchette, *The large deviation approach to statistical mechanics*, Phys. Rep. 478, 1 (2009).

[41] S.R.S. Varadhan, *Large Deviations and Applications*, (SIAM, Philadelphia, 1984).