HAL open science

A trajectorial interpretation of the dissipations of entropy and Fisher information for stochastic differential equations

Joaquin Fontbona, Benjamin Jourdain

To cite this version:

Joaquin Fontbona, Benjamin Jourdain. A trajectorial interpretation of the dissipations of entropy and Fisher information for stochastic differential equations. 2013. hal-00608977v2

HAL Id: hal-00608977

https://hal.science/hal-00608977v2

Preprint submitted on 2 Jul 2013 (v2), last revised 5 Feb 2016 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A trajectorial interpretation of the dissipations of entropy and Fisher information for stochastic differential equations

J.Fontbona* B.Jourdain†

July 2, 2013

Abstract

We introduce and develop a pathwise description of the dissipation of general convex entropies for continuous time Markov processes, based on simple backward martingales and convergence theorems with respect to the tail sigma field. The entropy is in this setting the expected value of a backward submartingale. In the case of (non necessarily reversible) Markov diffusion processes, we use Girsanov theory to explicit its Doob-Meyer decomposition, thereby providing a stochastic analogue of the well known entropy dissipation formula, valid for general convex entropies (including total variation). Under additional regularity assumptions, and using Itô's calculus and ideas of Arnold, Carlen and Ju [2], we obtain a new Bakry Emery criterion which ensures exponential convergence of the entropy to 0. This criterion is non-intrisic since it depends on the square root of the diffusion matrix, and cannot be written only in terms of the diffusion matrix itself. Last, we provide examples where the classic Bakry Emery criterion fails, but our non-intrisic criterion ensuring exponential convergence to equilibrium applies without modifying the law of the diffusion process.

Keywords : long-time behaviour, stochastic differential equations, time reversal, Girsanov theory, Bakry Emery criterion, convex Sobolev inequalities

AMS 2010 subject classifications : 60H10 60H30 37A35 26D10 35B40

Introduction

We are interested in the long-time behaviour of solutions to the stochastic differential equation

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt \tag{0.1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\otimes d'}$ and $W = (W_t, t \geq 0)$ is a standard Brownian motion in $\mathbb{R}^{d'}$.

We consider a convex function $U : [0, \infty) \to \mathbb{R}$ bounded from below and define the $U-$entropy of a probability measure p in $\mathbb{R}^d$ with respect to a probability measure q in $\mathbb{R}^d$ by

$$H_U(p|q) = \begin{cases} \int_{\mathbb{R}^d} U\left(\frac{dp}{dq}(x)\right) dq(x) & \text{if } p \ll q \\ +\infty \text{ otherwise.} \end{cases}$$

*DIM-CMM, UMI(2807) UCHILE-CNRS, Universidad de Chile, Casilla 170-3, Correo 3, Santiago-Chile, e-mail:fontbona@dim.uchile.cl. Partially supported by Fondecyt Grant 1110923 and BASAL-Conicyt.

†Université Paris-Est, CERMICS, 6-8 av Blaise Pascal, Cité Descartes, Champs sur Marne, 77455 Marne-la-Vallée Cedex 2, France - e-mail : jourdain@cermics.enpc.fr. Supported by ANR-09-BLAN-0216-01 MEGAS.

The particular cases $U(x) = \mathbf{1}_{x>0} x \ln(x)$ and $U(x) = (x-1)^2$ respectively correspond to the usual entropy and the χ^2-distance. For $U(x) = |x-1|$, $H_U(p|q)$ coincides with the total variation distance but only when $p \ll q$.

In case (0.1) admits a reversible probability measure, the celebrated Bakry Emery curvature dimension criterion which involves the generator, the carré du champs and the iterated carré du champs is a sufficient condition for this reversible measure to satisfy a Poincaré inequality and a logarithmic Sobolev inequality. From these inequalities, one can respectively deduce exponential convergence to 0 as $t \to \infty$ of the chi-square distance or the relative entropy between the marginal at time t of the process and its reversible measure. This criterion has been generalized to entropy functions U more general than $U(r) = (r-1)^2$ and $U(r) = r\ln(r)$ (see for instance [1]).

In general, even when the stochastic differential equation (0.1) admits an invariant probability measure, this measure is not reversible. It is well known both from a probabilistic point of view [13] and the point of view of partial differential equations [2] that the presence of a contribution antisymmetric with respect to the invariant measure in the drift may accelerate convergence to this invariant measure as $t \to \infty$.

The primal goal of this work is to recover the results of [2] and [1] about the long-time behaviour of the U-entropy of the law of X_t with respect to the invariant measure by arguments based on Itô's stochastic calculus. To achieve this goal, we introduce and develop in the first section of the paper a pathwise description of the dissipation of general convex entropies for continuous time non-homogeneous Markov processes, based on simple backward martingales and convergence theorems with respect to the tail sigma field. Given two different initial laws, the U-entropy of the marginal at time t of the Markov process under the first initial law with respect to its marginal at time t under the second initial law is the expected value of a backward submartingale. This implies that this U-entropy is non-increasing with t and permits to characterize its limit as $t \to \infty$. To our knowledge, this simple result does not seem to have been used in the study of the trend to equilibrium of Markov processes. Conditions ensuring that the U-entropy of general Markov processes converges to 0 are discussed at the end of the section (the case of Markov diffusion processes is studied in more details in the second section).

From the second section of the paper on, we only deal with Markov diffusions given by

$$dX_t = \sigma(t, X_t)dW_t + b(t, X_t)dt \tag{0.2}$$

where $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d \otimes d'}$. Under assumptions that guarantee that for both initial laws, the time-reversed processes are still diffusions, we use Girsanov theory to explicit the Doob-Meyer decomposition of the submartingale obtained in the first section. In this way, we obtain a stochastic analogue of the well known entropy dissipation formula, valid for general convex entropies (including total variation). Taking expectations in this formula, we recover the well known fact that the U-entropy dissipation is equal to the U-Fisher information. Proofs of the main results of this section are given in Appendix A.

It should be noticed that the idea of considering a trajectorial interpretation of entropy to obtain functional inequalities is not new, at least for reversible diffusions (see e.g. the work of Cattiaux [5] whose results are nevertheless of quite different nature). However, even in the reversible case, time reversal of a diffusion starting out of equilibrium modifies the dynamics of the diffusion. The simple martingales introduced in the first section take this fact into account and moreover permit the use of Itô's calculus under less regularity than *a priori* needed when working in the forward time direction. Their interest thus goes beyond the treatment of non-reversible situations.

In the third section, we further suppose that the stochastic differential equation is time-homogeneous (i.e. of the form (0.1)) and that it admits an invariant probability distribution, that is chosen as one of the two initial laws. Under additional regularity assumptions, and using Itô's calculus and some ideas close to Arnold, Carlen and Ju [2], we obtain a new Bakry Emery criterion which ensures exponential convergence of the U-Fischer information to 0 and therefore exponential convergence of the U-entropy to 0. In addition, under this criterion, the invariant measure satisfies a U-convex Sobolev inequality. This criterion is non-intrisic : it depends on the square root σ of the diffusion matrix $a = \sigma\sigma^*$ and cannot be written only in terms of the diffusion matrix itself whereas, under mild regularity assumptions on b and a, the law of $(X_t)_{t\geq 0}$ solving (0.1) is characterized by the associated martingale problem only written in terms of a and b. The main results of this section are proved in Appendix B. We point out that our approach also allows us to recover the results and criterion provided in [2]. The way in which this can be done is described in the Appendix C, where also the difference between the arguments leading to each of the two criteria is highlighted. Additionally, we provide therein a combined criterion based on both the one of [2] and ours.

Last, we provide in Section 4 two examples where the classic Bakry Emery criterion fails, but our non-intrisic criterion ensures exponential convergence to equilibrium without modifying the law of the diffusion process.

As future work, we plan to investigate how to choose the square root σ of the diffusion matrix in order to maximize the rate of exponential convergence to equilibrium given by our non-intrisic Bakry Emery criterion.

Throughout this work, we make the following assumption on U:

$H0)$ $U : [0, \infty) \to \mathbb{R}$ is a convex function such that $\inf U > -\infty$.

Notice that U is then continuous on $(0, +\infty)$ and such that $U(0) \geq \lim_{x\to 0^+} U(x)$.

Moreover, we use the convention of summation over repeated indexes.

Acknowledgements : We thank Tony Lelièvre (CERMICS) for pointing out to us the paper of Arnold, Carlen and Ju [2] at an early stage of this research. We also thank Anton Arnold (TU Wien) for suggesting an improvement of our non-intrinsic Bakry-Emery criterion (see Remark 3.3 below). The first author last thanks the hospitality and partial support of CERMICS.

1 Entropy decrease for general continuous-time Markov processes

In this section $(X_t : t \geq 0)$ is an arbitrary continuous-time Markov process with values in a measurable space $(E, \mathcal{E})$.

Let P_0, Q_0 be probability measures on E. We will use the following notation:

- P_t and Q_t denote the time marginal laws of X_t when the initial laws are P_0 and Q_0, respectively.
- $(X_t^P, t \geq 0)$ and $(X_t^Q, t \geq 0)$ denote realizations of the process (X_t) with X_0 respectively distributed according to P_0 and Q_0.

Proposition 1.1 *The function $t \in \mathbb{R}_+ \mapsto H_U(P_t|Q_t) \in \mathbb{R} \cup \{+\infty\}$ is non-increasing.*

Moreover, if for some $t \geq 0$, $P_t \ll Q_t$, then the law of $(X_r^P)_{r\geq t}$ is absolutely continuous with respect to the one of $(X_r^Q)_{r\geq t}$ with density $\frac{dP_t}{dQ_t}(X_t^Q)$, it holds for all $s \geq t$ that $P_s \ll Q_s$, and $\left(\frac{dP_s}{dQ_s}(X_s^Q)\right)_{s\geq t}$ is a backward martingale with respect to the filtration $\mathcal{F}_s = \sigma(X_r^Q, r \geq s)$.

Last, if $H_U(P_t|Q_t) < +\infty$ for some $t \geq 0$, then $\left(U(\frac{dP_s}{dQ_s}(X_s^Q))\right)_{s\geq t}$ is a backward submartingale with respect to $\mathcal{F}_s$ with expectation $\mathbb{E}\left(U(\frac{dP_s}{dQ_s}(X_s^Q))\right) = H_U(P_s|Q_s)$.

Corollary 1.2 *If $H_U(P_t|Q_t) < +\infty$ for some $t \geq 0$, then*

$$\lim_{s\to\infty} H_U(P_s|Q_s) = \mathbb{E}\left(U\left(\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^Q)\middle| \cap_{s\geq 0}\mathcal{F}_s\right)\right)\right) < \infty.$$

In particular, if the tail σ-field $\cap_{s\geq 0}\mathcal{F}_s$ is trivial a.s. w.r.t. the law of $(X_t^Q)_{t\geq 0}$ then $\lim_{s\to\infty} H_U(P_s|Q_s) = U(1)$.

Proof of Proposition 1.1. Let $s \geq t \geq 0$. If P_t is not absolutely continuous with respect to Q_t, then $+\infty = H_U(P_t|Q_t) \geq H_U(P_s|Q_s)$. Otherwise, for $f : E^{\mathbb{R}_+} \to \mathbb{R}$ measurable with respect to the product sigma-field, using the Markov property for the second equality, one has

$$\begin{aligned}\mathbb{E}(f(X_r^P, r \geq t)) &= \int_{\mathbb{R}^d} \mathbb{E}(f(X_r^P, r \geq t)|X_t^P = x)dP_t(x) = \int_{\mathbb{R}^d} \mathbb{E}(f(X_r^Q, r \geq t)|X_t^Q = x)\frac{dP_t}{dQ_t}(x)dQ_t(x) \\ &= \mathbb{E}\left(f(X_r^Q, r \geq t)\frac{dP_t}{dQ_t}(X_t^Q)\right). \end{aligned} \tag{1.1}$$

The law of $(X_r^P)_{r\geq t}$ is thus absolutely continuous with respect to the one of $(X_r^Q)_{r\geq t}$ with density $\frac{dP_t}{dQ_t}(X_t^Q)$ and $\forall r \geq t$, $P_r \ll Q_r$. Now, for $s \geq t$,

$$\mathbb{E}\left(f(X_r^Q, r \geq s)\frac{dP_t}{dQ_t}(X_t^Q)\right) = \mathbb{E}\left(f(X_r^P, r \geq s)\right) = \mathbb{E}\left(f(X_r^Q, r \geq s)\frac{dP_s}{dQ_s}(X_s^Q)\right)$$

where we used (1.1) with t replaced by s. This ensures that $\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^Q)\middle|\mathcal{F}_s\right) = \frac{dP_s}{dQ_s}(X_s^Q)$. By Jensen's inequality, since U is a convex function bounded from below,

$$\mathbb{E}\left(U\left(\frac{dP_t}{dQ_t}(X_t^Q)\right)\middle|\mathcal{F}_s\right) \geq U\left(\frac{dP_s}{dQ_s}(X_s^Q)\right). \tag{1.2}$$

Taking expectations one concludes that $H_U(P_t|Q_t) \geq H_U(P_s|Q_s)$. ∎

Proof of Corollary 1.2. If $H_U(P_t|Q_t) < +\infty$ then $P_t \ll Q_t$ and the $\mathcal{F}_s$ backward martingale $(\frac{dP_s}{dQ_s}(X_s^Q))_{s\geq t}$ converges a.s. to $\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^Q)\middle| \cap_{s\geq 0}\mathcal{F}_s\right)$ when $s \to \infty$. By the backward martingale property, for $r \geq t$,

$$\mathbb{E}\left(\frac{dP_r}{dQ_r}(X_r^Q)1_{\{\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^Q)|\cap_{s\geq 0}\mathcal{F}_s\right)=0\}}\right) = \mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^Q)1_{\{\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^Q)|\cap_{s\geq 0}\mathcal{F}_s\right)=0\}}\right) = 0.$$

Hence $\frac{dP_r}{dQ_r}(X_r^Q) = 0$ a.s. on the set $\left\{\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^Q)\middle| \cap_{s\geq 0}\mathcal{F}_s\right) = 0\right\}$. With the continuity of U on $(0,+\infty)$, one deduces that the random variables $U\left(\frac{dP_s}{dQ_s}(X_s^Q)\right)$, $s \geq 0$, converge a.s. to $U\left(\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^Q)\middle| \cap_{s\geq 0}\mathcal{F}_s\right)\right)$ as $s \to +\infty$. Since they are uniformly integrable, convergence also takes place in L^1, and we conclude that $H_U(P_s|Q_s) = \mathbb{E}\left(U\left(\frac{dP_s}{dQ_s}(X_s^Q)\right)\right)$ converges as $s \to \infty$ to the asserted limit. When the tail σ-field is trivial a.s. w.r.t. the law of $(X_t^Q)_{t\geq 0}$, then the limit of the backward martingale is equal to $\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^Q)\right) = 1$ and $U\left(\frac{dP_s}{dQ_s}(X_s^Q)\right)$ converges a.s. and in L^1 to $U(1)$. ∎

The next proposition gives a framework in which the tail σ-fied is trivial.

Proposition 1.3 *Assume that E is a locally compact metric space and that the Markov process $(X_t)_{t\geq 0}$ is Feller and with càdlàg paths. Assume moreover that there is a sigma-finite Borel measure μ with full support in E with respect to which the transition semigroup of $(X_t)_{t\geq 0}$ has a strictly positive transition density $\varphi_t(x,y)$, continuous in (x,y) for each $t > 0$, and that $(X_t)_{t\geq 0}$ has an invariant distribution. Then the tail σ-field $\cap_{t\geq 0}\sigma(X_r, r \geq t)$ is trivial a.s. w.r.t. the law of $(X_t^Q)_{t\geq 0}$ for any choice of initial law Q_0.*

Proof . Notice on one hand that if $(P^{(t)})_{t\geq 0}$ denotes the semigroup of $(X_t)_{t\geq 0}$, by continuity of $x \mapsto \varphi_t(x,y)$ and Fatou's Lemma the functions $P^{(t)}f(x)$ and $P^{(t)}(1-f)(x)$ summing 1 are both l.s.c. if f is any measurable function $f : E \to [0,1]$. This implies that $(X_t)_{t\geq 0}$ actually is strongly Feller. On the other hand, the positivity of the transition density implies that any invariant probability measure is equivalent to μ. Therefore there is a unique (thus ergodic) invariant distribution, which is of the form $p_\infty(x)\mu(dx)$ with $p_\infty(x) > 0$, $\mu(dx)$ a.e..

Now let $\mathbb{P}_\infty$ and $\mathbb{P}_x$ denote the laws of $(X_t)_{t\geq 0}$ on the canonical space $E^{[0,+\infty)}$ when respectively starting from an initial condition distributed according to $p_\infty\mu$ and to δ_x, and write $(Y_t)_{t\geq 0}$ for the canonical process. By the ergodic theorem and the strict positivity of p_∞, we have $\int_0^\infty \mathbf{1}_A(Y_t)dt = +\infty$, $\mathbb{P}_\infty-$ a.s. for each Borel set A in E such that $\mu(A) > 0$. If $\tilde{A} = \{y \in E^{[0,+\infty)} : \int_0^\infty \mathbf{1}_A(y_s)ds = \infty\}$, we deduce that $\mathbb{P}_x(\tilde{A}) = 1$ for $\mu(dx)-$ almost every x. But $\tilde{A}$ is a tail event, and by the Markov property, for any $t > 0$, one has $\mathbb{P}_x(\tilde{A}) = \mathbb{E}_x(\mathbb{P}_{Y_t}(\tilde{A}))$. The strong Feller property then yields $\mathbb{P}_x(\tilde{A}) = 1$ for all $x \in E$. That is, X is positive Harris recurrent with maximal irreducibility measure μ.

By Theorem 1.3.9 in [15] (and noting that its proof uses only continuity of $\varphi_t(x,y)$ in (x,y) for each $t > 0$ but not continuity in (t,x,y)), we get that any tail event B is such that $\mathbb{P}_x(B) = 1$ for all $x \in E$ or $\mathbb{P}_x(B) = 0$ for all $x \in E$, which concludes the proof.

∎

Remark 1.4 *In the case of Markov diffusion processes studied below, conditions ensuring the joint continuity in (x,y) of the transition density with respect to Lebesgue measure can be found in [10] Chapter 9 under uniform ellipticity, and in [16] Theorem 4.5 under hypoellipticity. Alternative conditions for the tail sigma field to be trivial are also given at the end of the next section.*

2 Entropy dissipation for diffusion processes

From now on we assume that $(X_t, t \geq 0)$ is a Markov diffusion process, solution to the stochastic differential equation

$$dX_t = \sigma(t, X_t)dW_t + b(t, X_t)dt \tag{2.1}$$

where $W = (W_t, t \geq 0)$ is a standard Brownian motion in $\mathbb{R}^{d'}$ and $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d \otimes d'}$ are mesurable coefficients satisfying conditions that will be specified below.

Our next goal is to explicitly describe the process $U\left(\frac{dP_s}{dQ_s}(X_s^Q)\right)$ when $P_0 \ll Q_0$ and, as a byproduct, the decrease of its expectation $H_U(P_s|Q_s)$.

We fix a finite time-horizon $T \in (0, +\infty)$ in order to work with standard (forward) filtrations by time reversal in $[0, T]$. Let us introduce some notation:

- $\mathbb{Q}^T$ (resp. $\mathbb{P}^T$) will denote the law of the time reversed processes $(X_{T-t}^Q, t \leq T)$ (resp. $(X_{T-t}^P, t \leq T)$) on the canonical space $C([0, T], \mathbb{R}^d)$.
- $(Y_t)_{t \leq T}$ stands from now on for the canonical process on $C([0, T], \mathbb{R}^d)$ and $\mathcal{G}_t = \sigma(Y_s, 0 \leq s \leq t)$ denotes its natural (complete, right continuous) filtration.
- In all the sequel, $\mathbb{E}^T$ will denote the expectation under the law $\mathbb{Q}^T$.

Remark 2.1 *a) Whenever $P_0 \ll Q_0$, we obtain by obvious adaptations of the results and proofs in the previous section that $\mathbb{P}^T \ll \mathbb{Q}^T$, with $\frac{d\mathbb{P}^T}{d\mathbb{Q}^T} = \frac{dP_0}{dQ_0}(Y_T)$, and that*

$$D_t \stackrel{\text{def}}{=} \left.\frac{d\mathbb{P}^T}{d\mathbb{Q}^T}\right|_{\mathcal{G}_t} = \frac{dP_{T-t}}{dQ_{T-t}}(Y_t), \quad 0 \leq t \leq T \tag{2.2}$$

is a $\mathbb{Q}^T - \mathcal{G}_t$ martingale. Moreover, $H_U(P_s|Q_s) < +\infty$ for $s \in [0, T]$ if and only if $(U(D_t))_{0 \leq t \leq T-s}$ is a uniformly integrable $\mathbb{Q}^T - \mathcal{G}_t$ submartingale, in which case one has

$$H_U(P_t|Q_t) = \mathbb{E}^T\left(U\left(D_{T-t}^T\right)\right) \text{ for all } t \in [s, T].$$

b) If $\mathbb{H}_U(\mathbb{P}_1|\mathbb{P}_2)$ denotes the pathwise $U-$entropy between two probability measures $\mathbb{P}_1, \mathbb{P}_2$ on $C([0, T], \mathbb{R}^d)$, that is

$$\mathbb{H}_U(\mathbb{P}_1|\mathbb{P}_2) = \begin{cases} \int_{C([0,T],\mathbb{R}^d)} U\left(\frac{d\mathbb{P}_1}{d\mathbb{P}_2}(w)\right) d\mathbb{P}_2(w) & \text{if } \mathbb{P}_1 \ll \mathbb{P}_2, \\ +\infty & \text{otherwise,} \end{cases}$$

one immediately gets that $\mathbb{H}_U\left(law\left(X_t^P : t \geq 0\right) \middle| law\left(X_t^Q : t \geq 0\right)\right) = H_U(P_0|Q_0)$ since X^P and X^Q have the same (forward) generator. The identities in a) written for $t = T$ imply that also

$$H_U(P_0|Q_0) = \mathbb{H}_U(\mathbb{P}^T|\mathbb{Q}^T).$$

The stability of the usual pathwise entropy ($U(r) = r \ln r$) under time reversal was remarked by Föllmer [8] and used therein to study time reversal of diffusion processes (see also [9] for an extension to infinite dimension where the situation is more complicated). In the converse sense, we will use here time reversal to study the $U-$entropy.

In order to use Itô calculus to obtain the explicit form of the Girsanov density D_t as a $\mathbb{Q}^T - \mathcal{G}_t$ martingale, and then deduce the Doob-Meyer decomposition of the submartingale $U(D_t)$, we will assume that the Markov processes $(X^Q_{T-t}, t \leq T)$ and $(X^P_{T-t}, t \leq T)$ are again diffusion processes. Conditions ensuring this fact have been studied in the aforementioned paper [8], and in Hausmann and Pardoux [11], Pardoux [18] and Millet *et. al* [17] among others, who in particular provide the semimartingale decomposition of $(X^Q_{T-t}, t \leq T)$ in its filtration, for suitable Q_0. We recall in Theorem 2.2 below the general results in [17] in a slightly more restrictive setting. The following conditions are needed:

$H1)$ For each $T > 0$, $\sup_{t\in[0,T]}(|b(t,0)|+|\sigma(t,0)|) < +\infty$ and for every $A > 0$ there is a constant $K_{T,A} > 0$ such that

$$|b(t,x) - b(t,y)| + \sum_{i=1}^{d'} |\sigma_{\bullet i}(t,x) - \sigma_{\bullet i}(t,y)| \leq K_{T,A}|x-y|, \ \forall t \in [0,T], \ \forall x,y \in B(0,A),$$

where $\sigma_{\bullet i}$ denotes the i-th column of the matrix σ and $B(0,A)$ is the ball of radius $A > 0$ centered at the origin in $\mathbb{R}^d$. Moreover,

$H1)'$ the constants $K_{T,A}$ do not depend on A, or

$H1)''$ for each $s \geq 0$, equation (2.1) starting at time s is strictly conservative, and for any bounded open set $D \subset \mathbb{R}^d$,

$$\sup_{x\in D} \sup_{s\in[0,T]} \mathbb{E}\left\{ \exp\left[\int_s^T \left[4B_{s,t}(x) + 8\sum_j |A^j_{s,t}(x)|^2 \right] dt \right] \right\} < \infty,$$

where $B_{s,t}(x) = \left[\sum_{i,k=1}^d \partial_i b_k((t, X_{s,t}(x))^2\right]^{\frac{1}{2}}$, $A^j_{s,t}(x) = \left[\sum_{i,k=1}^d \partial_i \sigma_{kj}(t, X_{s,t}(x))^2\right]^{\frac{1}{2}}$ and $X_{s,t}(x)$ denotes the solution to (2.1) starting from x at time $s < t$.

$H2)_Q$ For each $t > 0$, the law $Q_t(dx)$ of X^Q_t has a density $q_t(x)$ with respect to Lebesgue measure.

$H3)_Q$ Denoting $a_{ij} = (\sigma\sigma^*)_{ij}$, for each $i = 1, \ldots, d$ the distributional derivative $\partial_j(a_{ij}(t,x)q_t(x))$ is a locally integrable function on $[0,T] \times \mathbb{R}^d$:

$$\int_0^T \int_D |\partial_j(a_{ij}(t,x)q_t(x))| dx dt < \infty \text{ for any bounded open set } D \subset \mathbb{R}^d.$$

We denote for $(t,x) \in [0,T] \times \mathbb{R}^d$

- $\bar{a}_{ij}(t,x) := a_{ij}(T-t,x), i,j = 1,\ldots,d,$
- $\bar{b}^Q_i(t,x) = -b_i(T-t,x) + \frac{\partial_j(a_{ij}(T-t,x)q_{T-t}(x))}{q_{T-t}(x)}$ (with the convention that the term involving $\frac{1}{q_{T-t}(x)}$ is zero when $q_{T-t}(x)$ is zero)

and notice that $\bar{b}^Q(t,x)$ is defined $dt \otimes dx$ a.e. on $[0,T] \times \mathbb{R}^d$ under assumption $H3)_Q$.

Theorem 2.2 *Assume that $H1)$ and $H2)_Q$ hold.*

a) Suppose moreover that $H3)_Q$ holds. Then, $\mathbb{Q}^T$ is a solution to the martingale problem:

$$(MP)_Q: \quad M_t^f := f(Y_t) - f(Y_0) - \int_0^t \frac{1}{2}\bar{a}_{ij}(s,Y_s)\partial_{ij}f(Y_s) + \bar{b}_i^Q(s,Y_s)\partial_i f(Y_s)ds, \ t \in [0,T]$$

is a continuous martingale with respect to the filtration $(\mathcal{G}_t)$ for all $f \in C_0^\infty(\mathbb{R}^d)$.

b) Let $\tilde{b} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\tilde{\sigma} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\otimes d'}$ be measurable functions such that $\int_0^T \int_D |\tilde{a}_{ij}(t,x)| + |\tilde{b}_i(t,x)| q_{T-t}(x)dxdt < \infty$ for any bounded open set $D \subset \mathbb{R}^d$. Assume moreover that $\mathbb{Q}^T$ is a solution to the martingale problem w/r to $(\mathcal{G}_t)$ for the generator $\mathcal{L}_t f(x) = \frac{1}{2}\tilde{a}_{ij}(t,x)\partial_{ij}f(x) + \tilde{b}_i(t,x)\partial_i f(x)$. Then $H3)_Q$ holds, $\tilde{b} = \bar{b}$ and $\tilde{a} = \bar{a}$.

Proof . According to Theorem 3.3 [17], under $H1)$, $H2)_Q$ and $H3)_Q$, (M_t^f) is a continuous $\mathcal{G}_t$-martingale in the interval $[0,T)$ under $\mathbb{Q}^T$. When f is C^∞ on $\mathbb{R}^d$ and vanishes outside $B(0,A)$, we have

$$\begin{aligned}
&\mathbb{E}^T\left(\int_0^T |\bar{b}_i^Q(s,Y_s)||\partial_i f(Y_s)|ds\right) \\
&\qquad \le \sup_{B(0,A)} |\nabla f| \left(T \sup_{[0,T]\times B(0,A)} |b(s,x)| + \int_{[0,T]\times B(0,A)} \sum_{i=1}^d |\partial_j(a_{ij}(s,x)q_s(x))|\, dsdx\right)
\end{aligned} \tag{2.3}$$

where the right-hand-side is finite under $H1)$ and $H3)_Q$. This implies that $\mathbb{E}^T(|M_T^f|) < +\infty$, and together with $H1)$, that $t \mapsto M_T^f$ is continuous on $[0,T]$ and a continuous $\mathcal{G}_t$-martingale on $[0,T]$ under $\mathbb{Q}^T$. Part b) follows from Theorem 3.3 in [17]. ∎

Thus, under $(MP)_Q$ and $(MP)_P$ the process Y_t is respectively a weak solution to the SDEs

$$dX_t = \bar{\sigma}(t,X_t)d\bar{W}_t + \bar{b}^Q(t,X_t)dt, \quad t \in [0,T] \tag{2.4}$$

and

$$dX_t = \bar{\sigma}(t,X_t)d\tilde{W}_t + \bar{b}^P(t,X_t)dt, \quad t \in [0,T],$$

where $\bar{\sigma}(t,x) = \sigma(T-t,x)$ and $\bar{W}$ and $\tilde{W}$ are d' dimensional Brownian motions in a possibly enlarged probability space. If uniqueness in law holds for both SDEs, P_0 and Q_0 (and then $\mathbb{P}^T$ and $\mathbb{Q}^T$ by Remark 2.1) are equivalent, and $p_t(x)$ and $q_t(x)$ are strictly positive and differentiable, then the difference between the drift terms of the two equations is

$$\begin{aligned}
\bar{b}_i^P(t,x) - \bar{b}_i^Q(t,x) =& \bar{a}_{ij}(t,\cdot)\partial_j \ln p_{T-t}(x) - \bar{a}_{ij}(t,\cdot)\partial_j \ln q_{T-t}(x) \\
=& \bar{a}_{ij}(t,x)\partial_j\left[\ln \frac{p_{T-t}}{q_{T-t}}(x)\right],
\end{aligned}$$

and the simplest form of Girsanov theorem allows us to deduce that

$$D_t = \exp\left\{\int_0^t \nabla^*\left[\ln\frac{p_{T-t}}{q_{T-t}}(x)\right]\bar{\sigma}(t,X_t)d\bar{W}_t - \frac{1}{2}\int_0^t \left(\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)ds\right\}.$$

However, in the general case when $q_t(x)$ or $p_t(x)$ may vanish and are possibly not differentiable, it is not clear what sense should be given to the derivatives above. If the diffusion matrix is singular, neither is it clear that difference of drift terms $\bar{b}^Q$ and $\bar{b}^P$ (defined by means of distributional derivatives) is in the range of the diffusion matrix, which is required to use Girsanov theorem.

This question is reminiscent and, somehow, reciprocal to the stochastic construction of Nelson processes where $\mathbb{Q}^T$ and the possibly singular difference of drift terms are given, and one tries to construct $\mathbb{P}^T$ (see for instance [6]). The following technical lemma answers the question in the most general situations covered by Theorem 2.2. Its proof, not hard but lengthy, relies on Girsanov theory in the absolutely continuous setting and is differed to the Appendix A.1 section. Recall that an element $\mathbb{P}_0 \in \mathcal{M}$ of a given set $\mathcal{M}$ of probability measures in $C([0,T],\mathbb{R}^d)$ is said to be *extremal* if $\mathbb{P}_0 = \alpha\mathbb{P}_1 + (1-\alpha)\mathbb{P}_2$ for some $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{M}$ and $\alpha \in (0,1)$ implies $\mathbb{P}_0 = \mathbb{P}_1 = \mathbb{P}_2$.

Lemma 2.3 *Assume that $H1)$, $H2)_Q$, $H3)_Q$ and $H3)_P$ hold, with $P_0 \ll Q_0$, and let $\frac{p_t}{q_t}(x)$ be the Radon-Nikodyn derivative of $p_t(x)dxdt$ w.r.t $q_t(x)dxdt$ on $[0,T]\times\mathbb{R}^d$ (the existence of which follows from Proposition 1.1). Then,*

a) there exists a measurable function in $[0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ denoted $(t,x)\mapsto \nabla \ln[\frac{p_t}{q_t}](x)$ such that

$$\bar{b}^P(t,x) - \bar{b}^Q(t,x) = \bar{a}(t,x)\nabla\left[\ln\frac{p_{T-t}}{q_{T-t}}(x)\right], \qquad p_{T-t}(x)dx\ dt\, a.e..$$

b) Define $q_t(x)dx\ dt$ a.e. in $[0,T]\times\mathbb{R}^d$ the function $(t,x)\mapsto\nabla[\frac{p_t}{q_t}](x)$ by

$$\nabla\left[\frac{p_t}{q_t}\right](x) := \frac{p_t}{q_t}(x)\nabla\left[\ln\frac{p_t}{q_t}\right](x)$$

and assume moreover that $\mathbb{Q}^T$ is an extremal solution to the martingale problem $(MP)_Q$. Then, the $\mathbb{Q}^T$-$(\mathcal{G}_t)$ martingale $(D_t)_{t\in[0,T]}$ introduced in (2.2) has a continuous version (denoted in the same way) satisfying

$$\begin{aligned} D_t =& \frac{p_T}{q_T}(Y_0) + \int_0^t D_s\nabla\ln\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R}\cdot dM_s \\ =& \frac{p_T}{q_T}(Y_0) + \int_0^t \nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{\{\frac{p_{T-s}}{q_{T-s}}(Y_s)>0\}}\cdot dM_s \end{aligned}$$

where $M_t = (M_t^i)_{i=1}^d$ are the continuous local martingales w.r.t. $\mathbb{Q}^T$ and $(\mathcal{G}_t)$ defined by

$$M_t^i := Y_t^i - Y_0^i - \int_0^t \bar{b}_i^Q(s,Y_s)ds,\ t\in[0,T]$$

and R is the $(\mathcal{G}_t)$-stopping time $R := \inf\{s\in[0,T] : D_s = 0\}$. Moreover, $\mathbb{Q}^T$ a.s., one has

$$\langle D\rangle_t = \int_0^t \left(\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R}\ ds\,,\ \forall t\in[0,T].$$

From the proof of Lemma 2.3 it will be clear that if p_t and q_t are everywhere strictly positive and of class C^1, $(t,x)\mapsto\nabla[\frac{p_t}{q_t}](x)$ and $(t,x)\mapsto\nabla\ln[\frac{p_t}{q_t}](x)$ can be respectively taken to be the usual gradient and gradient of the logarithm of $\frac{p_t}{q_t}$ (see Remark A.4 for details). An exponential form for D_t can also be given.

We introduce now the notation U'_- and $U''(dy)$ for respectively the left-hand derivative of the restriction of the convex function $U : [0,\infty)\to\mathbb{R}$ to $(0,+\infty)$ and the non-negative measure on $(0,+\infty)$ equal to the second order distribution derivative of this restriction.

We are ready to state the main result of this section:

Theorem 2.4 *(Stochastic U-Entropy dissipation) Let Q_0 and P_0 be probability measures on $\mathbb{R}^d$ such that*

$$H_U(P_0|Q_0) < \infty$$

and assume that $H1)$, $H2)_Q$, $H3)_Q$ and $H3)_P$ hold. Suppose moreover that $\mathbb{Q}^T$ is an extremal solution to the martingale problem $(MP)_Q$.

Then, the submartingale $(U(D_t))_{t\in[0,T]}$ has the Doob-Meyer decomposition

$$\begin{aligned}\forall t \in [0,T],\ U(D_t) =& U(D_0) + \int_0^t U'_-(D_s)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R}\cdot dM_s \\ &+ \frac{1}{2}\int_{(0,+\infty)} L_t^r(D)U''(dr) - 1_{\{0<R\leq t\}}\Delta U(0),\end{aligned} \tag{2.5}$$

where $R := \inf\{s \in [0,T] : D_s = 0\}$, $\Delta U(0) = \lim_{x\to 0^+} U(x) - U(0) \leq 0$ and $L_t^r(D)$ denotes the local time at level $r \geq 0$ and time t of the continuous version of the martingale $(D_s)_{s\in[0,T]}$.

In particular, if U is continuous on $[0,+\infty)$ and C^2 on $(0,+\infty)$, one has

$$\begin{aligned}\forall t \in [0,T],\ U(D_t) =& U(D_0) + \int_0^t U'(D_s)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R}\cdot dM_s \\ &+ \frac{1}{2}\int_0^t U''\left(\frac{p_{T-s}}{q_{T-s}}(Y_s)\right)\left(\nabla^*\left[\frac{p_{T-s}}{q_{T-s}}\right]\bar{a}(s,\cdot)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right]\right)(Y_s)\mathbf{1}_{s<R}ds.\end{aligned} \tag{2.6}$$

In the above equation and from now on, we denote by ∇^* the transpose of the gradient. The proof of Theorem 2.4 is given in the Appendix A section. We next briefly discuss some of its assumptions and then state some consequences.

Remark 2.5 *a) By Theorem 3.1 in [11], conditions $H2)_Q$ and $H3)_Q$ hold under condition $H1)'$ if Q_0 has a density q_0 w.r.t. the Lebesgue measure s.t. $\int_{\mathbb{R}^d}\frac{q_0^2(x)dx}{1+|x|^k} < +\infty$ for some $k > 0$ and either*

$$\forall T > 0,\ \exists\varepsilon > 0,\ \forall(t,x) \in [0,T]\times\mathbb{R}^d,\ a(t,x)=\sigma\sigma^*(t,x) \geq \varepsilon I_d,$$

or the second order distribution derivatives $\frac{\partial^2 a^{ij}}{\partial x_i \partial x_j}(t,x)$ are bounded on $[0,T]\times\mathbb{R}^d$ for each $T > 0$ (by Theorem 3.1. in p. 1199 [11], the latter conditions imply that (A)(ii) in p. 1189 and thus Theorem 2.1 therein hold). In particular, under $H1)'$ and the previous conditions, $H2)_P$ and $H3)_P$ also hold if for instance $P_0 \ll Q_0$ and $\frac{dP_0}{dQ_0}$ has polynomial growth.

b) Condition H1)" introduced in [17] allows us to include in our study the fundamental examples of Langevin diffusions with $a(x) = I_d$ and $b(x) = -\nabla V(x)$ for a nonnegative C^2 potential V, possibly superquadratic but satisfying:

$$\limsup_{|x|\to\infty}\frac{-x^*\nabla V(x)}{|x|^2} < +\infty,\ \limsup_{|x|\to\infty}\frac{\Delta V}{|\nabla V|^2}(x) < 2 \text{ and } \limsup_{|x|\to\infty}\frac{\sqrt{\partial_{ik}V\partial_{ik}V}}{V}(x) = 0. \tag{2.7}$$

See the Appendix section A.5 for a proof of this fact.

c) Extremality of the solution $\mathbb{Q}^T$ to the martingale problem $(MP)_Q$ is implied by pathwise uniqueness for the stochastic differential equation (2.4). In the relevant case that σ and

b in (2.1) are time-homogeneous and (0.1) admits an invariant density $p_\infty(x) > 0$, when choosing $Q_0(dx) = p_\infty(x)dx$ equation (2.4) takes the form

$$dX_t = \sigma(X_t)dW_t + \left(\frac{\partial_j(a_{\bullet j}p_\infty)}{p_\infty}(X_t) - b(X_t)\right)dt \quad t \in [0,T].$$

Pathwise uniqueness for this SDE can be proved under H1) by a standard argument using localization, Itô's formula and Gronwall's lemma, whenever the function $-\frac{\partial_j(a_{\bullet j}p_\infty)}{p_\infty}$ is the sum of a locally Lipschitz function and a monotone function. This is for instance the case when $a = I_d$ and $p_\infty(x) = Ce^{-2V(x)}$ for some convex function $V : \mathbb{R}^d \to \mathbb{R}$, or when a and $p_\infty(x) > 0$ have locally Lipschitz derivatives.

The proof of Theorem 2.4 will justify that expectations can be taken in (2.5) and (2.6), and get

Corollary 2.6 *(U-Entropy dissipation) Under the assumptions of Theorem 2.4, we have*

$$\forall t \in [0,T], \quad H_U(P_t|Q_t) = H_U(P_T|Q_T) - \Delta U(0)\mathbb{Q}^T(0 < R \le T-t) + \frac{1}{2}\int_{(0,+\infty)} \mathbb{E}^T\left(L^r_{T-t}(D)\right)U''(dr). \tag{2.8}$$

If U is moreover continuous on $[0,+\infty)$ and C^2 on $(0,+\infty)$, we get the well known expression for the entropy dissipation:

$$\begin{aligned}\forall t \in [0,T], \quad & H_U(P_t|Q_t) = H_U(P_0|Q_0) \\ & - \frac{1}{2}\int_0^t \int_{\{\frac{p_s}{q_s}(x)>0\}} U''\left(\frac{p_s}{q_s}(x)\right)\left(\nabla^*\left[\frac{p_s}{q_s}\right]a(s,\cdot)\nabla\left[\frac{p_s}{q_s}\right]\right)(x)q_s(x)dxds, \end{aligned} \tag{2.9}$$

with $U''(r)$ now standing for the second order derivative of U at $r > 0$.

The particular case $U(x) = |x-1|$ of the total variation is more intricate but we are still able to derive a dissipation formula analogous to (2.9). To our knowledge this formula is new:

Corollary 2.7 *(Dissipation of total variation) Under the assumptions of Theorem 2.4, suppose moreover that for a.e. $t \in [0,T]$, the functions $x \mapsto q_t(x)$ and $x \mapsto \frac{p_t}{q_t}(x)$ are respectively of class C^1 and C^2 and there exists a sequence $(r_n)_n$ of positive numbers tending to $+\infty$ as $n \to \infty$, such that $\lim_{n\to\infty}\frac{1}{r_n}\int_{\{r_n \le |x| < 2r_n\}}\left|a(t,x)\nabla\left[\frac{p_t}{q_t}\right](x)\right|q_t(x)dx = 0$. Furthermore, assume that $\int_0^T\int_{\mathbb{R}^d}\left|\nabla\cdot\left[\bar{a}(s,x)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](x)q_{T-s}(x)\right]\right|dxds < \infty$. Then, $\forall t \in [0,T]$,*

$$\|P_t - Q_t\|_{\mathrm{TV}} = \|P_0 - Q_0\|_{\mathrm{TV}} + \frac{1}{2}\int_0^t\int_{\mathbb{R}^d}\widetilde{sign}\left(\frac{p_s}{q_s} - 1\right)(x)\nabla\cdot\left[a(s,x)\nabla\left[\frac{p_s}{q_s}\right](x)q_s(x)\right]dxds$$

where $\widetilde{sign}(r) = -\mathbf{1}_{(-\infty,0)}(r) + \mathbf{1}_{(0,\infty)}(r)$ and the integral is non-positive for all $t \in [0,T]$.

The proof is given in the Appendix.

Remark 2.8 *a) Denote by $\mathbb{Q}$ the law of $(X^Q_t, t \le T)$ and by $\mathbb{E}$ the corresponding expectation. The following "forward" version of formula (2.8) holds under the assumptions of Theorem 2.4 if moreover $\frac{p_t}{q_t}(Y_t)$ is a continuous $(\mathcal{G}_t)$ semimartingale under $\mathbb{Q}$ (in particular if $(t,x) \mapsto \frac{dP_t}{dQ_t}(x)$ has a version of class $C^{1,2}$):*

$$\forall t \in [0,T], \quad H_U(P_t|Q_t) = H_U(P_0|Q_0) + \Delta U(0)\mathbb{Q}(0 < S \le t) - \frac{1}{2}\int_{(0,+\infty)}\mathbb{E}\left(L^r_t\left(\frac{p_\cdot}{q_\cdot}(Y_\cdot)\right)\right)U''(dr),$$

where $S := \inf\{s \in [0,T] : \frac{p_s}{q_s}(Y_s) > 0\}$. *This follows from the pointwise relation*

$$L_T^r\left(\frac{p_{T-\cdot}}{q_{T-\cdot}}(X^Q_{T-\cdot})\right) - L^r_{T-t}\left(\frac{p_{T-\cdot}}{q_{T-\cdot}}(X^Q_{T-\cdot})\right) = L^r_t\left(\frac{p_\cdot}{q_\cdot}(X^Q_\cdot)\right)$$

and the fact that the process $\left(\frac{p_{T-t}}{q_{T-t}}\left(X^Q_{T-t}\right)\right)_{t\in[0,T]}$ *is a.s. stopped upon hitting* 0, *by Lemma 2.3.*

b) *The limit type assumption in Corollary 2.7 is not too stringent. Thanks to* (2.9) *and Cauchy-Schwarz inequality, it holds true for instance if the matrix* a *is uniformly bounded and* $H_U(P_0|Q_0) < \infty$ *for* $U(r) = (r-1)^2$, *since* $\left|a\nabla\left[\frac{p_t}{q_t}\right]\right| = \sup_{|v|\leq 1}(\sigma v)^*(\sigma\nabla\frac{p_t}{q_t}) \leq \sqrt{|a|}\sqrt{\nabla^*\left[\frac{p_t}{q_t}\right]a\nabla\left[\frac{p_t}{q_t}\right]}$.

We end this section providing sufficient conditions in order that $\lim_{t\to\infty} H_U(P_t|Q_t) = 0$.The proof of the following result is also differed to the Appendix A section.

Proposition 2.9 *Let us assume that the coefficients* σ *and* b *are time-homogeneous, globally Lipschitz continuous and that the SDE* (0.1) *admits an invariant density* p_∞ *locally Lipschitz and bounded away from* 0 *and* $+\infty$ *and such that* $\exists k > 0$, $\int_{\mathbb{R}^d}\frac{p^2_\infty(x)dx}{1+|x|^k} < +\infty$ *and that* $-\frac{\partial_j(a_{\bullet j}p_\infty)}{p_\infty}$ *is the sum of a locally Lipschitz function and a monotone function. We also suppose that*

$$\forall A > 0,\ \exists\varepsilon_A > 0,\ \forall|x| \leq A,\ a(x) \geq \varepsilon_A I_d \tag{2.10}$$

with either ε_A *not depending on* A *or the second order distribution derivatives* $\frac{\partial a^{ij}}{\partial x_i\partial x_j}$ *bounded on* $\mathbb{R}^d$. *Then the tail sigma-field* $\cap_{t\geq 0}\sigma(X_r, r \geq t)$ *is trivial a.s. w.r.t. the law of* $(X^Q_t)_{t\geq 0}$ *for any choice of initial law* Q_0. *In particular, if* $U(1) = 0$, *then as soon as* $H_U(P_s|Q_s) < +\infty$ *for some* $s < +\infty$, *one has* $\lim_{t\to\infty} H_U(P_t|Q_t) = 0$.

Remark 2.10 *In the case that* $(X_r)_{r\geq 0}$ *is Feller, the conclusion still holds when the globally Lipschitz condition on* b *is relaxed to locally Lipschitz and* $H1')$ *hold (see the proof of Proposition 2.9).*

3 Dissipation of the Fisher information and non-intrisic Bakry Emery criterion

We will from now on focus in the case when $Q_0(dx) = p_\infty(x)dx$ is a stationary probability law for the time-homogeneous Markov diffusion (0.1) . We denote

$$I_U(p_s|p_\infty) = \frac{1}{2}\int_{\{\frac{p_s}{p_\infty}>0\}} U''\left(\frac{p_s}{p_\infty}\right)\nabla^*\left[\frac{p_s}{p_\infty}\right]a\nabla\left[\frac{p_s}{p_\infty}\right]p_\infty dx$$

the integral that appears in the right-hand-side of (2.9), and we refer to it as the $U-$ Fisher information.

Inspired by the famous Bakry-Emery approach, we want to compute the derivative of $I_U(p_s|p_\infty)$ with respect to the time variable.

In all the sequel, we make the following assumptions :

$H4)$ The drift function b and the matrix σ are time-homogeneous and such that $H1)$ holds. Moreover, b (resp. σ) admits first (resp. second) order derivatives which are locally α-Hölder-continuous on $\mathbb{R}^d$ for some $\alpha > 0$.

$H5)_{p_\infty}$ The Markov process defined by (0.1) has an invariant density $p_\infty(x)$ and $Q_0(dx) = p_\infty(x)dx$. Moreover, p_∞ admits derivatives up to the second order which are locally α-Hölder-continuous on $\mathbb{R}^d$ for some $\alpha > 0$ and $p_\infty(x) > 0$ for all $x \in \mathbb{R}^d$.

$H6)^T_{p_0}$ The initial distribution P_0 admits a probability density p_0 with respect to the Lebesgue measure. Moreover, we assume that $H2)_{p_0}$ holds and that $p_t(x)$ has space derivatives up to the second order for each $t > 0$, which are continuous in $(t, x) \in (0, T] \times \mathbb{R}^d$ and bounded and Hölder continuous in $x \in \mathbb{R}^d$ uniformly on $[\delta, T] \times \mathbb{R}^d$ for each $\delta \in (0, T]$.

Let us also introduce some notations :

- We write $\mathbb{P}^T_\infty := \mathbb{Q}^T$ and $\bar{b}_i := \bar{b}^Q_i$, $i = 1, \dots, d$.
- By possibly enlarging the probability space $\mathcal{G}_t - \mathbb{P}^T_\infty$, we introduce a Brownian motion $\bar{W}$ such that Y_t solves the stochastic differential equation :

$$dY_t = \sigma(Y_t)d\bar{W}_t + \bar{b}(Y_t)dt, \quad t \in [0, T] \text{ where } \bar{b}_i(y) = -b_i(y) + \frac{\partial_j(a_{ij}(y)p_\infty(y))}{p_\infty(y)}. \tag{3.1}$$

  By assumptions $H4)$ and $H5)_\infty$, the coefficients σ and $\bar{b}$ are locally Lipschitz so that trajectorial uniqueness holds for this SDE. By the Yamada-Watanabe theorem, one deduces that uniqueness holds for the martingale problem $(MP)_Q$.
- We write $\rho_t(x) := \frac{p_{T-t}}{p_\infty}(x),\ t \in [0, T]$.

Notice that $H5)_{p_\infty}$ implies $H2)_Q$ for $Q_0(dx) = p_\infty(x)dx$ and combined with $H4)$, it implies $H3)_Q$. Moreover $H6)^T_{p_0}$ implies $H2)_P$ and $H3)_P$. Therefore the hypotheses of Theorem 2.4 hold within the present Section. Notice also that, under $H5)_\infty$ and $H6)^T_{p_0}$, the derivative $\nabla \frac{p_t}{p_\infty}$ is defined everywhere and according to the proof of Lemma 2.3 (see in particular Remark A.4 at the end of Section A.1) the function $\nabla \frac{p_t}{p_\infty}$ in Lemma 2.3 may be taken to be the true gradient. We will throughout in the sequel make this choice , in particular in Equation (2.9) which states that the U-entropy dissipation is equal to the U-Fisher information.

Under $H4)$, if moreover a and b are bounded with a uniformly elliptic, then $H6)^T_{p_0}$ holds for any compactly supported probability density p_0, by [10] Chapter 9. We refer to [16] for conditions ensuring that $H6)^T_{p_0}$ holds under hypoellipticity.

To compute the dissipation of the U-Fischer information, in all the sequel we make the following regularity assumption on U:

$H7)$ The convex function $U : [0, \infty) \to \mathbb{R}$ is of class C^4 on $(0, +\infty)$, continuous on $[0, +\infty)$ and satisfies $U(1) = U'(1) = 0$.

The assumption that $U'(1) = 0$ is inspired in the analysis on admissible entropies developed in Arnold et al. [1] and is granted without modifying the functions $p \mapsto H_U(p|p_\infty)$ and $p \mapsto I_U(p|p_\infty)$ by replacing $U(r)$ by $U(r) - U'(1)(r - 1)$ if needed. Notice that if $H7)$ holds, $U(r)$ attains the minimum 0 at $r = 1$ and therefore $U \geq 0$ by convexity. Following [3] p202 (see also [1, 7]), we introduce an additional assumption on U:

$H7')$ $\forall r \in (0,\infty)$, $(U^{(3)}(r))^2 \leq \frac{1}{2}U''(r)U^{(4)}(r)$,

which is satisfied for instance by $U(r) = r\ln r - (r-1)$ and by $U(r) = (r-1)^2$. Let us recall consequences of $H7)'$ pointed out in [1] (see Remark 2.3 therein) which will be used in proving the following results.

Remark 3.1 *Condition $H7')$ implies that $\left(\frac{1}{U''}\right)'' \leq 0$ at points where $U'' \neq 0$. Since $U'' \geq 0$, and excluding the uninteresting case where U'' identically vanishes, the previous implies that $\frac{1}{U''}$ is finite in $[0,\infty)$, and therefore that U is strictly convex. We then deduce from $H7')$ that $U^{(4)} \geq 0$ in $(0,\infty)$. By concavity and positivity of $\frac{1}{U''}$ this function is moreover non decreasing, and we deduce that $U^{(3)} \leq 0$ in $(0,\infty)$.*

We do not assume that the entropy function U is C^4 on the closed interval $[0,+\infty)$, since we want to deal with $U(r) = r\ln(r) - (r-1)$. That is why we introduce some regularization U_δ indexed by a positive parameter δ : we chose U_δ such that $U_\delta(r) = U(r+\delta)$ for $r \geq 0$ and U_δ is extended to a C^4 function on $\mathbb{R}$. In the next proposition as well as in the remaining of the paper, we will omit the argument (t, Y_t) in order to obtain more compact formulae.

Proposition 3.2 *Under $H4)$, $H5)_{p_\infty}$, $H6)_{p_0}^T$ and $H7)$, one has on the time-interval $[0,T]$*

$$d\left[U_\delta''(\rho)\nabla^*\rho a\nabla\rho\right] = tr(\Lambda_\delta\Gamma)dt + U_\delta''(\rho)\bar{\theta}dt + d\hat{M}^{(\delta)} \text{ with } tr(\Lambda_\delta\Gamma) \geq 0 \text{ under } H7)'$$

and where $\hat{M}_t^{(\delta)} = \int_0^t \partial_k\left[U_\delta''(\rho)\nabla^*\rho a\nabla\rho\right]\sigma_{kr}d\bar{W}_s^r$ is a $\mathcal{G}_t - \mathbb{P}_\infty^T-$local martingale,

$$\bar{\theta} = 2\Big\{\partial_{l'}\rho\partial_l\rho\left[\frac{1}{4}(\partial_k\sigma_{lj}a_{km}\partial_m\sigma_{l'j} - \sigma_{ki}\partial_k\sigma_{lj}\sigma_{mj}\partial_m\sigma_{l'i}) + \frac{1}{2}\bar{b}_m\partial_m a_{ll'} + \frac{1}{2}\sigma_{l'i}a_{mk}\partial_{mk}\sigma_{li} - a_{ml'}\partial_m\bar{b}_l\right]$$
$$+ \left[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}\right]\partial_{l'}\rho\partial_m\sigma_{li}\partial_{kl}\rho\Big\}.$$

and Λ_δ and Γ are the square matrices defined by

$$\Lambda_\delta := \begin{bmatrix} U_\delta''(\rho) & U_\delta^{(3)}(\rho) \\ U_\delta^{(3)}(\rho) & \frac{1}{2}U_\delta^{(4)}(\rho) \end{bmatrix} \qquad \Gamma := \begin{bmatrix} \Gamma_{11} & (\sigma_{\bullet i}\cdot\nabla\rho)\nabla^*\rho\ a\nabla(\sigma_{\bullet i}\cdot\nabla\rho) \\ (\sigma_{\bullet i}\cdot\nabla\rho)\nabla^*\rho\ a\nabla(\sigma_{\bullet i}\cdot\nabla\rho) & |\nabla^*\rho a\nabla\rho|^2 \end{bmatrix}$$

with $\Gamma_{11} = \sum_{i,j=1}^d \left(\sigma_{kj}\sigma_{li}\partial_{kl}\rho + \frac{1}{2}(\sigma_{kj}\partial_k\sigma_{li} + \sigma_{ki}\partial_k\sigma_{lj})\partial_l\rho\right)^2$

The computation of $d\left[U_\delta''(\rho)\nabla^*\rho a\nabla\rho\right]$ is postponed to Appendix B. Let us nevertheless discuss the sign of the term $tr(\Lambda_\delta\Gamma)$ which is inspired from [3] p202 and also from the term $tr(\mathbf{XY})$ in [2] pp 163-164 (see Appendix C for a detailed comparison with the computations in that paper). Since, by Cauchy Schwarz inequality,

$$((\sigma_{\bullet i}\cdot\nabla\rho)\nabla^*\rho\ a\nabla(\sigma_{\bullet i}\cdot\nabla\rho))^2 = \left((\sigma^*\nabla\rho)_i(\sigma^*\nabla\rho)_j\left(\sigma_{kj}\sigma_{li}\partial_{kl}\rho + \frac{1}{2}(\sigma_{kj}\partial_k\sigma_{li} + \sigma_{ki}\partial_k\sigma_{lj})\partial_l\rho\right)\right)^2$$
$$\leq \Gamma_{11}\sum_{i,j=1}^d (\sigma^*\nabla\rho)_i^2(\sigma^*\nabla\rho)_j^2 = \Gamma_{11}\left|\nabla^*\rho a\nabla\rho\right|^2,$$

the determinant of the matrix Γ is nonnegative and this matrix is positive semidefinite. Under $H7')$, Λ_δ is also positive semidefinite and $tr(\Lambda_\delta\Gamma) \geq 0$.

Remark 3.3 *In a previous version of this paper, the coefficient Γ_{11} was chosen equal to*

$$\sum_{i,j=1}^{d}(\sigma_{kj}\sigma_{li}\partial_{kl}\rho+\sigma_{kj}\partial_k\sigma_{li}\partial_l\rho)^2=\sum_{i,j=1}^{d}(\sigma^*\nabla(\sigma^*\nabla\rho)_i)_j^2=\nabla^*((\sigma\nabla\rho)_i a\nabla((\sigma\nabla\rho)_i).$$

We thank Anton Arnold for pointing out to us that the positive semidefiniteness of the matrix Γ is preserved under the new choice of this coefficient. Notice that, by symmetry of $\sigma_{kj}\sigma_{li}\partial_{kl}\rho$ in i and j,

$$\begin{aligned}\sum_{i,j=1}^{d}(\sigma_{kj}\sigma_{li}\partial_{kl}\rho+\sigma_{kj}\partial_k\sigma_{li}\partial_l\rho)^2-\Gamma_{11}&=\frac{1}{4}\sum_{i,j=1}^{d}((\sigma_{kj}\partial_k\sigma_{li}-\sigma_{ki}\partial_k\sigma_{lj})\partial_l\rho)^2\\&=\frac{1}{2}(\partial_k\sigma_{lj}a_{km}\partial_m\sigma_{l'j}-\sigma_{ki}\partial_k\sigma_{lj}\sigma_{mj}\partial_m\sigma_{l'i})\partial_l\rho\partial_{l'}\rho\end{aligned}$$

is a nonnegative quadratic form applied to $\nabla\rho$ which implies that the Bakry Emery criterion below improves upon the one of the previous version.

We introduce one last assumption on the density flow $\rho_t=\frac{p_{T-t}}{p_\infty}$:

$H6')_{p_0}^T$ For each $T'\in(0,T)$ the following integrals are finite:

- $\int_0^{T'}\left|U^{(3)}(\rho)\vee-1\right|^2|\nabla^*\rho a\nabla\rho|^3p_\infty(x)dxdt$
- $\int_0^{T'}(U''(\rho)\wedge 1)^2\nabla^*(\nabla^*\rho a\nabla\rho)a\nabla(\nabla^*\rho a\nabla\rho)p_\infty(x)dxdt$
- $\int_0^{T'}(U''(\rho)\wedge 1)\left[\,|(\sigma_{l'i}a_{m\bullet}-\sigma_{\bullet i}a_{ml'})\partial_m\sigma_{li}|+|\partial_k\left([\sigma_{l'i}a_{mk}-\sigma_{ki}a_{ml'}]\,\partial_m\sigma_{li}\right)|\,\right]|\partial_{l'}\rho||\partial_l\rho|p_\infty(x)dxdt$
- $\int_0^{T'}(U''(\rho)\wedge 1)\left[\,|(\sigma_{l'i}a_{mk}-\sigma_{ki}a_{ml'})\partial_m\sigma_{li}(\partial_l\rho\partial_k\ln p_\infty+\partial_{lk}\rho)|\,\right]|\partial_{l'}\rho|p_\infty(x)dxdt$

We also denote by $H6)_{p_0}^\infty$ (resp. $H6')_{p_0}^\infty$) the assumption satisfied when $H6)_{p_0}^T$ (resp. $H6')_{p_0}^T$) holds for each $T>0$.

Theorem 3.4 *Let Θ denote the $d\times d$ symmetric matrix defined by*

$$\begin{aligned}\Theta_{ll'}=&-\frac{1}{2}b_m\partial_m a_{ll'}+\frac{1}{2}(a_{kl'}\partial_k b_l+a_{kl}\partial_k b_{l'})-\frac{1}{4}a_{mk}\partial_{mk}a_{ll'}-\frac{1}{2}(a_{kl'}\partial_{kj}a_{lj}+a_{kl}\partial_{kj}a_{l'j})\\&-a_{kl}a_{jl'}\partial_{kj}\ln(p_\infty)-\frac{1}{2}(a_{kl}\partial_k a_{l'j}+a_{kl'}\partial_k a_{lj})\partial_j\ln(p_\infty)-\frac{1}{4}(a_{mk}\partial_m\sigma_{li}\partial_k\sigma_{l'i}+\sigma_{ki}\partial_k\sigma_{lj}\sigma_{mj}\partial_m\sigma_{l'i})\\&+\frac{1}{2}\sigma_{ki}(\partial_m\sigma_{li}a_{ml'}+\partial_m\sigma_{l'i}a_{ml})\partial_k\ln(p_\infty)+\frac{1}{2}\partial_k[\sigma_{ki}(\partial_m\sigma_{li}a_{ml'}+\partial_m\sigma_{l'i}a_{ml})]\end{aligned}$$

and assume that $\Theta(x)$ is $p_\infty(x)dx-$a.e. positive semidefinite. Then, under $H4)$, $H5)_{p_\infty}$, $H6)_{p_0}^T$ $H6')_{p_0}^T$, $H7)$ and $H7')$, for a.e. $t\in[0,T]$ one has

$$\frac{d}{dt}\int_{\rho_t>0}U''(\rho_t)[\nabla^*\rho_t a\nabla\rho_t]p_\infty dx\geq 2\int_{\rho_t>0}U''(\rho_t)\nabla^*\rho_t\Theta\nabla\rho_t p_\infty dx. \tag{3.2}$$

If, moreover $I_U(p_0|p_\infty)<+\infty$, $H6)_{p_0}^\infty$ and $H6')_{p_0}^\infty$ hold and the matrix Θ satisfies the non-intrinsic Bakry-Emery criterion

NIBEC) $\exists\lambda>0,\ \forall x\in\mathbb{R}^d,\ \Theta(x)\geq\lambda a(x),$

then $\forall t \geq 0$, $I_U(p_t|p_\infty) \leq e^{-2\lambda t} I_U(p_0|p_\infty)$ *and the non-increasing function* $t \mapsto H_U(p_t|p_\infty)$ *also converges at exponential rate* 2λ *to its limit as* $t \to \infty$.

Remark 3.5

- *The matrix* Θ *and therefore our Bakry-Emery criterion are non-intrinsic in the sense that they cannot in general be written in terms of the diffusion matrix* a *only without making explicit use of* σ. *This is because we have got rid of the nonnegative term* $tr(\Lambda_\delta \Gamma)$ *which appears in the first equation in Proposition 3.2 and involves the non-intrisic term* Γ_{11}.
- *In case* $a = 2\nu I_d$ *and* $b = -(\nabla V + F)$ *with* F *such that* $\nabla.(e^{-V/\nu} F) = 0$, *then* $p_\infty \propto e^{-V/\nu}$, $\bar{b} = -b + 2\nu \nabla \ln p_\infty = -\nabla V + F$ *and* $\Theta = \nu(2\nabla^2 V - \nabla F - \nabla F^*)$. *For the canonical choice* $\sigma = \sqrt{2\nu} I_d$, *condition NIBEC) therefore writes* $\exists \lambda > 0$, $\forall x \in \mathbb{R}^d$, $\nabla^2 V(x) - \frac{\nabla F + \nabla F^*}{2}(x) \geq \lambda I_d$ *which is exactly condition (A2) in the introduction of [2], page 158.*

The proof of (3.2) is postponed to Appendix B.2. Let us deduce the last assertion of Theorem 3.4. Reverting time in (3.2) and using *NIBEC)*, one obtains that for $r \geq 0$,

$$\frac{d}{dr} I_U(p_r|p_\infty) \leq -2\lambda I_U(p_r|p_\infty).$$

Hence $\forall r \geq t \geq 0$, $I_U(p_r|p_\infty) \leq e^{-2\lambda(r-t)} I_U(p_t|p_\infty)$. Since by Theorem 2.4, one has $\frac{d}{dr} H_U(p_r|p_\infty) = -I_U(p_r|p_\infty)$, one deduces that

$$0 \leq H_U(p_t|p_\infty) - \lim_{r\to\infty} H_U(p_r|p_\infty) = \int_t^\infty I_U(p_r|p_\infty) dt \leq \frac{I_U(p_t|p_\infty)}{2\lambda} \leq \frac{e^{-2\lambda t} I_U(p_0|p_\infty)}{2\lambda}. \quad (3.3)$$

We deduce

Theorem 3.6 *Assume* $H4)$, $H5)_{p_\infty}$, $H6)_{p_0}^\infty$ $H6')_{p_0}^\infty$, $H7)$ *and* $H7')$, *that the matrix* $\Theta(x)$ *is* $p_\infty(x)dx$ *– a.e. positive semidefinite, that the diffusion matrix* a *is locally uniformly strictly positive definite and that* $H_U(p_s|p_\infty) < +\infty$ *for some* $s \geq 0$. *Then* $H_U(p_t|p_\infty)$ *converges to* 0 *as* $t \to \infty$. *Moreover, under* $NIBEC)$, *for* $t > s$, *one has the convex Sobolev inequality*

$$H_U(p_t|p_\infty) \leq \frac{1}{2\lambda} I_U(p_t|p_\infty), \quad (3.4)$$

$$\text{and } \forall t \geq s,\ H_U(p_t|p_\infty) \leq e^{-2\lambda(t-s)} H_U(p_s|p_\infty). \quad (3.5)$$

Proof . Reverting time in (3.2), we obtain that $t \mapsto I_U(p_t|p_\infty)$ is non-increasing. When $H_U(p_s|p_\infty)$ is finite for some $s \geq 0$, writing (2.9) on the interval $[s, T]$ in place of $[0, T]$ with arbitrarily large T, we deduce that $I_U(p_t|p_\infty)$ is finite on $(s, +\infty)$ and tends to 0 as $t \to \infty$. When a is locally uniformly strictly positive definite, the beginning of the proof of Theorem 2.5 (before Part(a)) [2], ensures that p_t tends to p_∞ in $L^1(\mathbb{R}^d)$. As a consequence, in the notations of Proposition 1.1, $\mathbb{E}\left|\frac{dP_t}{dQ_t}(X_t^Q) - 1\right|$ tends to 0 as $t \to \infty$ and therefore the a.s. limit $\mathbb{E}\left(\frac{dP_t}{dQ_t}(X_t^Q)\middle| \cap_{s\geq 0} \mathcal{F}_s\right)$ of $\frac{dP_t}{dQ_t}(X_t^Q)$ is equal to 1. By Corollary 1.2, one concludes that $H_U(p_t|p_\infty)$ tends to $U(1) = 0$. Under $NIBEC)$, for $t > s$, $I_U(p_t|p_\infty) < +\infty$ and reasoning like in the derivation of (3.3), one obtains (3.4). This implies that

$$\frac{d}{dt} H_U(p_t|p_\infty) = -I_U(p_t|p_\infty) \leq -2\lambda H_U(p_t|p_\infty)$$

from which the last assertion follows readily. ∎

Remark 3.7 *In view of Corollary 1.2, Proposition 1.3 and Remark 1.4, the local uniform strict positive definiteness assumption on the diffusion matrix a may be replaced by some hypoellipticity assumption to ensure that $H_U(p_t|p_\infty)$ tends to 0 as $t \to \infty$ at exponential rate 2λ as soon as $H_U(p_s|p_\infty)$ is finite for some $s \geq 0$. By the last step of the proof of Theorem 3.6, this implies* (3.4) *and* (3.5) *under $NIBEC$).*

4 Examples

Consider the reversible diffusion process in $\mathbb{R}^2$ with coefficients given for each $(x_1, x_2) \in \mathbb{R}^2$ by

$$a(x_1, x_2) = I_2, \quad \text{and} \quad b(x_1, x_2) = -\nabla V(x_1, x_2),$$

where V is the globally C^2 convex potential

$$V(x_1, x_2) := |x_1|^2 + |x_1 - x_2|^{2+\alpha} + |x_2|^{2+\alpha}$$

for some $\alpha \in (0, 1)$. The invariant measure is $p_\infty \propto e^{-2V}$, and we have

$$\begin{aligned}\partial_1 V =& 2x_1 + (2+\alpha) sign(x_1 - x_2)|x_1 - x_2|^{1+\alpha}\\ \partial_2 V =& (2+\alpha) sign(x_2)|x_2|^{1+\alpha} + (2+\alpha) sign(x_2 - x_1)|x_2 - x_1|^{1+\alpha}\end{aligned}$$

and

$$\nabla^2 V = \begin{pmatrix} 2 & 0 \\ 0 & (2+\alpha)(1+\alpha)|x_2|^\alpha \end{pmatrix} + (2+\alpha)(1+\alpha)|x_1 - x_2|^\alpha \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

The drift $b = -\nabla V$ is locally Lipschitz continuous. Moreover, $(x_1, x_2)\nabla V(x_1, x_2) \geq 0$ and $\sqrt{\partial_{ik} V \partial_{ik} V(x_1, x_2)} \leq C\sqrt{1 + |x_2|^{2\alpha} + |x_1 - x_2|^{2\alpha}}$ so that $\limsup_{|(x_1,x_2)|\to+\infty} \frac{\sqrt{\partial_{ik} V \partial_{ik} V(x_1,x_2)}}{V(x_1,x_2)} = 0$. Last $\Delta V(x_1, x_2) \leq C(1 + |x_2|^\alpha + |x_1 - x_2|^\alpha)$ whereas

$$\begin{aligned}|\nabla V|^2(x_1, x_2) \geq& (2|x_2| + (2+\alpha)|x_1 - x_2|^{1+\alpha})^2 \mathbf{1}_{sign(x_2) \neq sign(x_2 - x_1)}\\ &+ (2+\alpha)^2(|x_2|^{1+\alpha} + |x_1 - x_2|^{1+\alpha})^2 \mathbf{1}_{sign(x_2) = sign(x_2 - x_1)}\end{aligned}$$

since $sign(x_2) \neq sign(x_2 - x_1)$ iff $x_1 \geq x_2 \geq 0$ or $x_1 \leq x_2 \leq 0$. Therefore $\limsup_{|(x_1,x_2)|\to+\infty} \frac{\Delta V}{|\nabla V|^2}(x_1, x_2) = 0$ and, by Remark 2.5 b), $H1)''$ is satisfied.

The classic Bakry-Emery criterion fails since $\nabla^2 V(0, 0)$ is singular but a logarithmic Sobolev inequality can be obtained by the perturbative argument of Holley and Stroock [12]. The potential V is also a particular case of the examples considered by Arnold, Carlen and Ju in the Section 3 of [2]. We notice that in order to check that p_∞ satisfies the convex Sobolev inequality (3.4), they first modify the Fokker-Planck equation by adding a non-symmetric drift term F as described in Remark 3.5 ii) above. Exponential convergence to 0 of $H_U(p_t|p_\infty)$ for the solution p_t of the original Fokker-Planck equation is only deduced in a second step.

It is nevertheless of interest to see how our non-intrisic Bakry Emery criterion allows us to prove directly that p_∞ satisfies the convex Sobolev inequality (3.4) and that $H_U(p_t|p_\infty)$ converges exponentially to 0. In contrast to [2] we modify the stochastic differential equation associated with the diffusion processes, by changing the square root σ of the diffusion matrix, but not the law of its solution or the associated Foker-Planck equation. We consider

$$\sigma = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$$

for a function $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ of class C^2 to be chosen later. We obtain after some computations

$$\Theta = \nabla^2 V - \frac{1}{4}|\nabla\phi|^2 I_2 - \frac{1}{4}\begin{pmatrix} (\partial_2\phi)^2 & -\partial_1\phi\partial_2\phi \\ -\partial_1\phi\partial_2\phi & (\partial_1\phi)^2 \end{pmatrix} + \begin{pmatrix} \partial_{12}\phi & \frac{\partial_{22}\phi-\partial_{11}\phi}{2} \\ \frac{\partial_{22}\phi-\partial_{11}\phi}{2} & -\partial_{12}\phi \end{pmatrix}$$
$$+ \begin{pmatrix} -2\partial_1\phi\partial_2 V & \partial_1\phi\partial_1 V - \partial_2\phi\partial_2 V \\ \partial_1\phi\partial_1 V - \partial_2\phi\partial_2 V & 2\partial_2\phi\partial_1 V \end{pmatrix}$$

We now consider a parameter $\varepsilon > 0$ which will be chosen small and a C^2 function $\varphi : \mathbb{R} \to \mathbb{R}$ such that $\varphi(s) = s$ if $|s| \le 1$ and $\varphi(s) = 0$ if $|s| \ge 2$. Then, we define

$$\phi(x_1, x_2) = -\varepsilon\varphi_\varepsilon(x_1)\varphi_\varepsilon(x_2), \quad (x_1, x_2) \in \mathbb{R}^2$$

where $\varphi_\varepsilon(s) = \varepsilon\varphi(s/\varepsilon)$. Notice that

$$\varphi_\varepsilon = O(\varepsilon), \quad \varphi''_\varepsilon = O(1/\varepsilon), \quad \text{and } \varphi'_\varepsilon = \begin{cases} 1 & \text{if } |s| \le \varepsilon, \\ O(1) & \text{if } \varepsilon < |s| < 2\varepsilon, \\ 0 & \text{if } |s| \ge 2\varepsilon. \end{cases}$$

Then, defining $B_\varepsilon := \{(x_1, x_2) \in \mathbb{R}^2 \, s.t. \, |x_1| \vee |x_2| \le \varepsilon\}$ and $C_\varepsilon := B_{2\varepsilon} \backslash B_\varepsilon$, we have

$$\partial_1\phi(x_1, x_2), \partial_2\phi(x_1, x_2) = \begin{cases} O(\varepsilon^2) & \text{if } (x_1, x_2) \in B_{2\varepsilon}, \\ 0 & \text{if } (x_1, x_2) \in B^c_{2\varepsilon}, \end{cases}$$

$$\partial_{12}\phi(x_1, x_2) = \begin{cases} -\varepsilon & \text{if } (x_1, x_2) \in B_\varepsilon, \\ O(\varepsilon) & \text{if } (x_1, x_2) \in C_\varepsilon, \\ 0 & \text{if } (x_1, x_2) \in B^c_{2\varepsilon}, \end{cases}$$

$$\frac{1}{2}(\partial_{11}\phi(x_1, x_2) - \partial_{22}\phi(x_1, x_2)) = \begin{cases} 0 & \text{if } (x_1, x_2) \in B_\varepsilon, \\ O(\varepsilon) & \text{if } (x_1, x_2) \in C_\varepsilon, \\ 0 & \text{if } (x_1, x_2) \in B^c_{2\varepsilon}, \end{cases}$$

and $\partial_1 V = O(\varepsilon), \partial_2 V = O(\varepsilon^{1+\alpha})$ on $B_{2\varepsilon}$. It follows that

$$\Theta = \nabla^2 V + \begin{pmatrix} -\varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix} + O(\varepsilon^3) \ge \begin{pmatrix} 2-\varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix} + O(\varepsilon^3) \quad \text{on } B_\varepsilon.$$

Next, the smallest eigenvalue of $\nabla^2 V(x_1, x_2)$, is given by

$$\gamma_- := 1 + \kappa_1 + \kappa_2/2 - \sqrt{1 + \kappa_1^2 - \kappa_2 + \kappa_2^2/4} \ge 1 + \kappa_2/2 - \sqrt{(\kappa_2/2 - 1)^2} = \kappa_2 \wedge 2 \tag{4.1}$$

with $\kappa_1 = \kappa_1(x_1, x_2) := (2+\alpha)(1+\alpha)|x_1 - x_2|^\alpha$ and $\kappa_2 = \kappa_2(x_1, x_2) := (2+\alpha)(1+\alpha)|x_2|^\alpha$. Since $\gamma_- = \kappa_1 + \kappa_2 + O(\kappa_1^2 + \kappa_2^2)$ as $\kappa_1^2 + \kappa_2^2 \to 0$ and $|x_2|^\alpha + |x_1 - x_2|^\alpha \ge (|x_2| + |x_1 - x_2|)^\alpha \ge |x_1|^\alpha$, one deduces that on C_ε,

$$\Theta = \nabla^2 V + O(\varepsilon) \ge (2+\alpha)(1+\alpha)\varepsilon^\alpha I_2 + o(\varepsilon^\alpha).$$

Last, by (4.1), $\inf_{(x_1,x_2)\in B^c_{2\varepsilon}} \gamma_- \ge ((2+\alpha)(1+\alpha)(2\varepsilon)^\alpha) \wedge 2 > 0$. One concludes that for ε small enough *NIBEC)* holds.

We next study a related second example of application of our criterion, where $\nabla^2 V$ is singular on a ball with positive radius. Once again, the perturbative argument of Holley Stroock [12] also ensures that a logarithmic Sobolev inequality holds for this choice of potential.

Let v be a convex C^2 function which vanishes on $[-\frac{1}{4},\frac{1}{4}]$ and such that $v''=2$ on $(-\infty,\frac{1}{2}]\cup[\frac{1}{2},+\infty)$. We set $v_\varepsilon(s)=\varepsilon^2 v\left(\frac{s}{\varepsilon}\right)$ and $V_\varepsilon(x_1,x_2)=x_1^2+v_\varepsilon(x_2)+v_\varepsilon(x_1-x_2)$. For $\varepsilon<\frac{1}{3}$, let φ_ε be a C^2 function such that

$$\varphi_\varepsilon(s)=\begin{cases} s \text{ when } |s|\le\varepsilon \\ 0 \text{ when } |s|\ge 1\end{cases}$$

and such that $\frac{-2\varepsilon}{1-\varepsilon}\le\varphi_\varepsilon'\le 1$, $|\varphi_\varepsilon|\le 2\varepsilon$ and $|\varphi_\varepsilon''|\le C$ where C is a constant not depending on ε. We set $\phi(x_1,x_2)=-\varphi_\varepsilon(x_1)\varphi_\varepsilon(x_2)$ so that $-1\le\partial_{12}\phi(x_1,x_2)\le\frac{2\varepsilon}{1-\varepsilon}$ with the first inequality being an equality on B_ε. One has $|\partial_{22}\phi-\partial_{11}\phi|\le 4C\varepsilon$ and $|\nabla\phi|=O(\varepsilon)$. As a consequence, $\Theta=\hat\Theta+O(\varepsilon)$ where

$$\hat\Theta=\begin{pmatrix} 2+v_\varepsilon''(x_1-x_2)+\partial_{12}\phi(x_1,x_2) & -v_\varepsilon''(x_1-x_2) \\ -v_\varepsilon''(x_1-x_2) & v_\varepsilon''(x_2)+v_\varepsilon''(x_1-x_2)-\partial_{12}\phi(x_1,x_2)\end{pmatrix}.$$

On B_ε, $\partial_{12}\phi(x_1,x_2)=-1$ and $\hat\Theta\ge I_2$. When $|x_2|\ge\frac{\varepsilon}{2}$, then $v_\varepsilon''(x_2)=2$ so that $\hat\Theta\ge(2-1)\wedge\left(2-\frac{2\varepsilon}{1-\varepsilon}\right)I_2$. When $|x_2|\le\frac{\varepsilon}{2}$ and $|x_1|>\varepsilon$ then $|x_1-x_2|\ge\frac{\varepsilon}{2}$ so that $v_\varepsilon''(x_1-x_2)=2$ and

$$\hat\Theta\ge\begin{pmatrix} 4+\partial_{12}\phi & -2 \\ -2 & 2-\partial_{12}\phi\end{pmatrix}\ge\left(3-\sqrt{5+2\partial_{12}\phi-(\partial_{12}\phi)^2}\right)I_2\ge\left(3-\sqrt{5+\frac{4\varepsilon}{1-\varepsilon}}\right)I_2.$$

We conclude that

$$\forall\lambda\in(0,3-\sqrt{5}),\ \text{for } \varepsilon>0 \text{ and small enough }, \forall x\in\mathbb{R}^d,\ \Theta(x)\ge\lambda I_2.$$

A Proofs of the main results of Section 2

A.1 Proof of Lemma 2.3

The proof of part a) relies on the following technical result:

Lemma A.1 *Assume that $H1)$, $H2)_P$ and $H3)_P$ hold.*

i) For each $i=1\dots,d$ and a.e. $t\in(0,T]$, the distribution $[a_{ij}(t,\cdot)\partial_j p_t]:=\partial_j(a_{ij}(t,\cdot)p_t)-p_t\partial_j a_{ij}(t,\cdot)$ is a function in $L^1_{loc}(dx)$ and, as a Radon measure in $[0,T]\times\mathbb{R}^d$, one has $[a_{ij}(t,\cdot)\partial_j p_t](x)dx\ dt\ll p_t(x)dx\ dt$. A version of the Radon-Nikodyn density (measurable in (t,x)) is given by $[a_{ij}(t,\cdot)\partial_j p_t](x)/p_t(x)$. Moreover, there exists a measurable function $(t,x)\mapsto K^p(t,x)\in\mathbb{R}^d$ such that for each $i=1\dots,d$

$$[a_{ij}(t,\cdot)\partial_j p_t](x)/p_t(x)=a_{i\bullet}(t,x)K^p(t,x),\ p_t(x)dx\ dt \text{ a.e.}$$

where $a_{i\bullet}$ denotes the row vector $(a_{i1},\dots,a_{id})$.

ii) If moreover $H2)_Q$, $H3)_Q$ and $P_0\ll Q_0$ hold, one has $[a_{ij}(t,\cdot)\partial_j p_t](x)dx\ dt\ll q_t(x)dx\ dt$ and $[a_{ij}(t,\cdot)\partial_j p_t](x)/q_t(x)$ is a (measurable in (t,x)) version of the Radon-Nikodyn derivative. Furthermore, it holds $p_{T-t}(x)dx\ dt$ (but not necessarily $q_{T-t}(x)dx\ dt$) a.e. that

$$\begin{aligned}\bar b_i^P(t,x)-\bar b_i^Q(t,x)=&[\bar a_{ij}(t,\cdot)\partial_j p_{T-t}](x)/p_{T-t}(x)-[\bar a_{ij}(t,\cdot)\partial_j q_{T-t}](x)/q_{T-t}(x)\\ =&\bar a_{i\bullet}(t,x)(K^p(T-t,x)-K^q(T-t,x)),\end{aligned}$$

and $q_{T-t}(x)dx\ dt$ (and thus $p_{T-t}(x)dx\ dt$) a.e. that

$$\frac{p_{T-t}(x)}{q_{T-t}(x)}(\bar b_i^P(t,x)-\bar b_i^Q(t,x))=\frac{p_{T-t}(x)}{q_{T-t}(x)}\bar a_{i\bullet}(t,x)(K^p(T-t,x)-K^q(T-t,x)).$$

Proof . The Lipschitz character of a (following from $H1$)) ensures that a has a.e. defined spatial derivatives of order 1 in $L^\infty_{loc}([0,T]\times\mathbb{R}^d)$ and thus that the distribution $a_{ij}(t,\cdot)\partial_j p_t$ as defined is a function in $L^1_{loc}([0,T]\times\mathbb{R}^d)$ under $H3)_P$. This implies, by Lemma A.2 in [17] (see also Lemma A.2 in [11]), that $a_{ij}(t,x)\partial_j p_t(x)$ vanishes a.e. on $\{x : p_t(x)=0\}$. This fact easily yields the remaining assertions, except for the existence of the functions K^p or K^q, which we establish in what follows.

We will on one hand use the fact, asserted in the proof of Lemma A.2 in [17], that for a.e. $t>0$ and each bounded open set O, $a_{ij}(t,x)\partial_j p_t(x)$ is the $\sigma(L^1(O),L^\infty(O))$-weak limit of some subsequence of $a_{ij}(t,x)\partial_j[\rho_n * p_t](x)$, for compactly supported regularizing kernels $\rho_n(x)=n^d\rho(nx)$. It is indeed shown in Lemma A.1 in [11] that for a suitable bounded sequence $\alpha_n>0$, $\alpha_n^{-1}|x|\ |\nabla\rho_n(x)|$ is again a regularizing kernel. The local Lipschitz character of a then yields the domination $\forall x\in O$, $|a_{ij}(t,x)\partial_j[\rho_n * p_t](x)|\le|\rho_n * \partial_j(a_{ij}(t,\cdot)p_t)(x)|+C\alpha_n^{-1}\int|x-y|\ |\nabla\rho_n(x-y)|p_t(y)dy$, the right-hand-side being, by the previous, an $L^1(O)$-converging sequence. Weak compactness is then provided by the Dunford-Pettis criterion, and the limit is identified integrating by parts against smooth test functions compactly supported in O. On the other hand, we will use the fact that diagonalizing the symmetric positive semidefinite matrix $(a_{ij}(t,x))=[u_1(t,x),\dots,u_d(t,x)]\Lambda(t,x)[u_1(t,x),\dots,u_d(t,x)]^*$ provides orthonormal vectors $(u_i(t,x))_{i=1}^d$ and the corresponding eigenvalues and diagonal components $(\lambda_i(t,x))_{i=1}^d$ of $\Lambda(t,x)$, that are measurable as functions of (t,x).

We take O as before and $a_{ij}(t,x)\partial_j[\rho_n * p_t](x)$ to be the subsequence described above. Defining the vectorial functions $w^{(n)}:=[u_1,\dots,u_d]^*\nabla[\rho_n * p_t]$ and $v_k=sign(u_k^*[a\nabla p])u_k,\ k=1,\dots,d$, we have

$$\int_{O\cap\{\lambda_k=0\}}|v_k^*[a\nabla p_t]|=\lim_{n\to\infty}\int_{O\cap\{\lambda_k=0\}}v_k^*[a\nabla[\rho_n * p_t]]=\lim_{n\to\infty}\int_{O\cap\{\lambda_k=0\}}\lambda_k w_k^{(n)}sign(u_k^*[a\nabla p_t])=0,$$

since $a\nabla[\rho_n * p_t]=\sum_{j=1}^d\lambda_j w_j^{(n)}u_j$ by the spectral decomposition of a. Consequently, for each t and a.e. $x\in\mathbb{R}^d$, the vector $[a(t,x)\nabla p_t(x)]$ belongs to the linear space $\left\langle(u_i(t,x))_{i=1,\dots,d;\lambda_i(t,x)\neq0}\right\rangle$. Denote now by $w=(w_j)_{j=1}^d:=(u_j^*a\nabla p_t)_{j=1}^d$ the coordinates of $a\nabla p_t$ w.r.t. the orthogonal basis $(u_j(t,x))_{j=1,\dots,d}$, so that w is a measurable function of (t,x). If we moreover denote by $\overline{\Lambda}$ the diagonal matrix with diagonal coefficients $\lambda_j^{-1}\mathbf{1}_{\lambda_j\neq0}, j=1,\dots,d$, and set $v:=[u_1,\dots,u_d]\overline{\Lambda}w$, then

$$av=[u_1,\dots,u_d]\Lambda[u_1,\dots,u_d]^*[u_1,\dots,u_d]\overline{\Lambda}w=[u_1,\dots,u_d]\Lambda\overline{\Lambda}w=[u_1,\dots,u_d]w$$

since $w=(w_j\mathbf{1}_{\lambda_j\neq0})_{j=1}^d$. That is, $(t,x)\mapsto v(t,x)\in\mathbb{R}^d$ is a measurable function such that for almost every $t\in[0,T]$ and each i, $a_{i\bullet}(t,x)v(t,x)=[a_{ij}\partial_j p_t(x)]$, dx a.e. Finally, $K^p(t,x):=v(t,x)/p_t(x)\mathbf{1}_{p_t(x)>0}$ has the required properties. ∎

We can now take $\nabla\ln\frac{p_t}{q_t}(x)$ to be an arbitrary representant of the equivalence class of the function $K^p(t,x)-K^q(t,x)$ under the relation $f(t,x)-g(t,x)\in Ker(a(t,x)),\ p_t(x)dx\ dt\ a.e.$ The identity in Lemma 2.3 a) is then satisfied by construction.

The proof of part b) of Lemma 2.3 relies first on a martingale representation property ensured by the extremality assumption:

Lemma A.2 *Assume that $H1), H2)_Q$ and $H3)_Q$ hold. For each $i=1,\dots,d$,*

$$M_t^i:=Y_t^i-Y_0^i-\int_0^t\bar{b}_i^Q(s,Y_s)ds,\ t\in[0,T]$$

is a continuous local martingale w/r to $\mathbb{Q}^T$ and $(\mathcal{G}_t)$, and $\langle M^i, M^j\rangle_t = \int_0^t \bar{a}^{ij}(s, Y_s)ds$ for all $i, j = 1, \dots, d$. Moreover, if $\mathbb{Q}^T$ is an extremal solution to the martingale problem $(MP)_Q$, then for any martingale $(N_t)_{t\in[0,T]}$ w/r to $\mathbb{Q}^T$ and $(\mathcal{G}_t)$ such that $N_0 = 0$, there exist predictable processes $(h_t^j)_{t\in[0,T],j=1,\dots d}$ with $\sum_{i,j=1}^d \int_0^T h_s^i \bar{a}_{ij}(s, Y_s)h_s^j ds < \infty$, $\mathbb{Q}^T$ a.s., and such that $(\int_0^t h_s \cdot dM_s = \sum_{j=1}^d \int_0^t h_s^j dM_s^j)_{t\in[0,T]}$ is a modification of $(N_t)_{t\in[0,T]}$. In particular, $(N_t)_{t\in[0,T]}$ has a continuous modification.

Proof . Standard localization arguments show that M_t^f in $(MP)_Q$ is a continuous local martingale for any $f \in C^2$ (see e.g. Proposition 2.2 in Ch. VII of [19], its proof for deterministic initial condition being also valid in the general case). Moreover, since $M_t^i = M_t^f$ for $f(x) = x^i$, by Proposition 2.4 , Ch. VII of [19] we get $\langle M^i, M^j\rangle_t = \int_0^t \bar{a}^{ij}(s, Y_s)ds$. The measure $\mathbb{Q}^T$ is therefore a solution to the Problem (12.9) in Jacod [14] in the filtered space $\big(C([0,T], \mathbb{R}^d), (\mathcal{G}_t)_{t\in[0,T]}\big)$, with data given by $\mathcal{G}_0$ and $(Y_t)_{t\in[0,T]}$, and characteristics respectively corresponding to: Q_T as the initial law, the $d-$dimensional process $(B^i = \int_0^\cdot \bar{b}^i(s, Y_s)ds)_{i=1}^d$, the matrix process $(C^{ij} = \int_0^\cdot \bar{a}^{ij}(s, Y_s)ds)_{i,j=1}^d$ and the trivial random measure process on $\mathbb{R}^d$ given by $\mu_t \equiv 0$. The extremality assumption together with Theorem 12.21 in [14] and the continuity of the canonical process under $\mathbb{Q}^T$ imply that any $L^2(\mathbb{Q}^T)-$bounded $(\mathcal{G}_t)-$martingale is the the $L^2(\mathbb{Q}^T)$ limit of linear combinations of stochastic integrals with respect to $M_t^i, i = 1, \dots, d$ (see also Proposition 12.10 in [14]). The statement follows by localization arguments.

∎

Part b) of Lemma 2.3 is contained in parts i) and ii) of the next result. For completeness, its part iii) additionally gives the exponential form of $(D_t)_{t\in[0,T]}$ defined in (2.3). By convenience, the convention $\inf \emptyset = +\infty$ is adopted and the filtration $(\mathcal{G}_t)_{t\in[0,T]}$ is extended to the whole interval $[0, \infty)$ by putting $\mathcal{G}_t = \mathcal{G}_T$ for all $t \in [T, \infty)$.

Lemma A.3 *Assume that $H1)$, $H2)_Q$, $H3)_Q$ and $H3)_P$ hold together. Suppose moreover that $P_0 \ll Q_0$ and that $\mathbb{Q}^T$ is an extremal solution to the martingale problem $(MP)_Q$. Recall that $(t, x) \mapsto \nabla[\frac{p_t}{q_t}](x)$ is $q_t(x)dx\ dt$ a.e. defined in $[0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ by $\nabla\left[\frac{p_t}{q_t}\right](x) := \frac{p_t}{q_t}(x)\nabla\left[\ln \frac{p_t}{q_t}\right](x)$.*

i) With R the $(\mathcal{G}_t)$-stopping time $R := \inf\{s \in [0, T] : D_s = 0\}$, we have $\mathbb{Q}^T-$a.s. that

$$\forall t \in [0, T],\ \int_0^t \left(\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s, Y_s)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R}\ ds < \infty,\ \textit{and}$$
$$\forall t \in [0, R),\ \int_0^t \left(\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s, Y_s)\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)ds < \infty \ \textit{on}\ \{R > 0\}.$$

ii) The process $(D_t)_{t\in[0,T]}$ has a continuous version, denoted in the same way, such that

$$\begin{aligned}\mathbb{Q}^T\ a.s,\ \forall t \in [0, T],\ D_t =& \frac{p_T}{q_T}(Y_0) + \int_0^t \nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R} \cdot dM_s \\ =& \frac{p_T}{q_T}(Y_0) + \int_0^t \nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{\{\frac{p_{T-s}}{q_{T-s}}(Y_s)>0\}} \cdot dM_s \\ \textit{and}\ \langle D\rangle_t =& \int_0^t \left(\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s, Y_s)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\mathbf{1}_{s<R}\ ds.\end{aligned}$$

iii) Finally, if we define the $(\mathcal{G}_t)$-stopping times $\tau^o := 0 \cdot \mathbf{1}_{D_0=0} + \infty \cdot \mathbf{1}_{D_0>0}$ and

$$\tau := \inf\left\{t \in [0,T] : \int_0^t \left(\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)ds = \infty\right\},$$

then $\mathbb{Q}^T-$a.s. $R = \tau \wedge \tau^o$, and $\forall t \in [0,T]$,

$$D_t = \mathbf{1}_{\{t<\tau\}}\frac{dp_T}{dq_T}(Y_0)\times$$
$$\exp\left\{\int_0^t \nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\cdot dM_s - \frac{1}{2}\int_0^t \left(\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\ln\frac{p_{T-s}}{q_{T-s}}\right](Y_s)ds\right\}. \tag{A.1}$$

In particular, on $\{R>0\}$ the second integral in i) is a.s. divergent at $t=R$.

Proof . By Lemma A.2, the $\mathbb{Q}^T$-martingale $(D_t)_{t\in[0,T]}$ admits the continuous version $D_0 + \sum_{j=1}^d \int_0^t h_s^j dM_s^j$ still denoted by D_t for simplicity. The martingale representation property and standard properties of stochastic integrals moreover imply that D_t is determined by the processes $\langle D, M^i\rangle = \int_0^\cdot \bar{a}_{ij}(t,Y_t)h_t^j dt$, $i = 1,\dots,d$. Consequently, h_t can be replaced (leaving D_t unchanged) by any predictable process k_t such that for each i, $\int_0^\cdot \sum_{j=1}^d h_t^j \bar{a}_{ij}(t,Y_t)dt = \int_0^\cdot \bar{a}_{ij}(t,Y_t)k_t^j dt$ $\mathbb{Q}^T$ a.s. (the fact that $\int_0^T k_s^i \bar{a}_{ij}(s,Y_s)k_s^j ds = \int_0^T \sum_{i,j=1}^d h_s^j \bar{a}_{ij}(s,Y_s)h_s^i ds < \infty$ $\mathbb{Q}^T$ a.s. follows then immediately). Furthermore, since $D_t = D_{t\wedge R}$ by standard properties of nonnegative continuous martingales, we may and shall assume that $\mathbb{Q}^T$ a.s. $h_t = h_t\mathbf{1}_{t<R} = h_t\mathbf{1}_{D_t>0}$ for all $t \in [0,T]$. Let us also notice that, by Fubini's Theorem, it $\mathbb{Q}^T-$a.s. holds that $D_s = \frac{p_{T-s}}{q_{T-s}}(Y_s)$ (and then $\mathbf{1}_{\{R>s\}} = \mathbf{1}_{\{\frac{p_{T-s}}{q_{T-s}}(Y_s)>0\}}$) for a.e. $s \in [0,T]$.

Now, by our assumptions and Theorem 2.2 a), $\mathbb{P}^T \ll \mathbb{Q}^T$ are probability measures respectively solving the martingale problems $(MP)_P$ and $(MP)_Q$. The processes $\int_0^\cdot \bar{b}_i^P(t,Y_t)dt$ and $\int_0^\cdot \bar{b}_i^Q(t,Y_t)dt + \int_0^\cdot (D_t)^{-1}h_t^j d\langle M^i, M^j\rangle_t$ then are $\mathbb{P}^T-$ indistinguishable (see e.g. Proposition 12.18 v) in [14]). Using these facts, the expression for $\langle M^i, M^j\rangle$ in Lemma A.2 and part ii) of Lemma A.1 we deduce first that, $\mathbb{P}^T-$a.s.,

$$\bar{b}_i^P(t,Y_t) - \bar{b}_i^Q(t,Y_t) = \bar{a}_{ij}(t,Y_t)\left(h_t^j\frac{q_{T-t}}{p_{T-t}}(Y_t)\right) = \bar{a}_{i\bullet}(t,Y_t)(K^p(T-t,Y_t) - K^q(T-t,Y_t)) \tag{A.2}$$

for a.e. $t \in [0,T]$ and each i. By part ii) of Lemma A.1 we then also get

$$\int_0^\cdot \bar{a}_{ij}(t,Y_t)h_t^j dt = \int_0^\cdot \bar{a}_{i\bullet}(t,Y_t)(K^p(T-t,Y_t) - K^q(T-t,Y_t))\frac{p_{T-t}(Y_t)}{q_{T-t}(Y_t)}dt,\ i = 1,\dots,d,$$

$\mathbb{P}^T-$a.s., and then $\mathbb{Q}^T-$a.s. because of our assumption on h. From these identities and our previous discussion we deduce the that we can choose $h_t = \nabla\frac{p_{T-t}}{q_{T-t}}(Y_t)\mathbf{1}_{\{\frac{p_{T-t}}{q_{T-t}}(Y_t)>0\}} = \nabla\frac{p_{T-t}}{q_{T-t}}(Y_t)\mathbf{1}_{\{R>t\}}$. This proves part ii). The first property of the process $\nabla\frac{p_{T-t}}{q_{T-t}}(Y_t)$ in i) is thus consequence of the general properties of h in the representation formula for D_t. The second assertion in i) easily follows from the first one, taking into account the definitions of $\nabla\frac{p_{T-t}}{q_{T-t}}(Y_t)$ and $\nabla\ln\frac{p_{T-t}}{q_{T-t}}(Y_t)$ and the properties of D_t.

To establish iii), we again use the extremality of $\mathbb{Q}^T$ in order to apply Theorem 12.48 in [14]. Thanks to part ii) of Lemma A.1 and equation (A.2), the objects z, K, B and T_n in (12.32),

(12.35) and (12.42) of [14] alluded in that theorem, respectively correspond in our setting to $\frac{p_T}{q_T}(Y_0)$, $\nabla \ln \left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)$, the process $A_t := \int_0^t \left(\nabla \left[\ln \frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s, Y_s) \nabla \left[\ln \frac{p_{T-s}}{q_{T-s}}\right](Y_s) ds$ and the stopping time $\tau_n := \inf\{t \in [0,T] : A_t \geq n\}$. This and Lemma 12.36 d) in [14], yield the fact that D_t equals (A.1), $\mathbb{Q}^T$ a.s in the set $\cup_{n\in\mathbb{N}}\{t \in [0,T] : t \leq \tau_n\}$. It is also established therein that $\tau_n \nearrow \tau$ $\mathbb{Q}^T$ a.s., impling that the latter random set equals the interval $[0,\tau) \cap [0,T]$. Moreover, on this interval, the integrals which appear in the exponential factor in (A.1) are finite. Therefore, either $\tau^o = 0$ and then $R = 0$, or $\tau^o = \infty$ and then $R \geq \tau$.

By Theorem 12.48 in [14] as well, we have $D_t = \liminf_{n\to\infty} D_{\tau_n}$ for t in $[\tau, T]$, $\mathbb{Q}^T$ a.s. Thus, $t \mapsto D_t$ is constant in $[\tau, T]$, $\mathbb{Q}^T$ a.s.. By Theorem 12.39 in [14] we have $\mathbb{P}^T(\tau < \infty) = 0$. Since $\mathbb{P}^T(\tau < \infty) = \mathbb{P}^T(\tau \leq T) = \mathbb{E}^T(1_{\{\tau\leq T\}} D_T)$, the a.s. constancy of $t \mapsto D_t$ on $[\tau, T]$ ensures that $\mathbb{Q}^T$ a.s. $D_t = 0$ for all $t \in [\tau, T]$, when the latter interval is non empty. As a consequence $\mathbb{Q}^T$ a.s., $R \leq \tau$ so that $R = \tau \wedge \tau^o$. This completes the proof. ∎

Remark A.4 *We notice from the proof of Lemma A.3 that the function $\nabla \frac{p_t}{q_t}(x)$ therein can be replaced by any representant of the equivalence class of the function $\frac{p_t}{q_t}(x)\left(K^p(t,x) - K^q(t,x)\right)$ under the relation $f(t,x) - g(t,x) \in Ker(a(t,x))$, $q_t(x)dx\ dt$ a.e.. If p_t and q_t are of class C^1 and q_t does not vanish on $\mathbb{R}^d$, the true gradient $\nabla \frac{p_t}{q_t}$ is equal to $\frac{p_t}{q_t}\left(\mathbf{1}_{p_t>0}\frac{\nabla p_t}{p_t} - \frac{\nabla q_t}{q_t}\right)$ (as the gradient of non-negative function vanishes at zeros of that function) and is such a representant.*

A.2 Proof of Theorem 2.4

Since by Lemma 2.3, $(D_t)_{t\in[0,T]}$ is a continuous non-negative $\mathbb{Q}^T$-martingale and U'_- is locally bounded on $(0,+\infty)$, $t \mapsto \int_0^t \left[U'_-(D_s)\right]^2 d\langle D\rangle_s$ is finite and continuous on $[0,T]$ when $R > T$ and finite and continuous on $[0,R)$ otherwise. In the latter case, $\int_0^R \left[U'_-(D_s)\right]^2 d\langle D\rangle_s$ makes sense but is possibly infinite. Define for any positive integer n the stopping time

$$R_n := \inf\left\{t \in [0, T\wedge R] : D_t \leq \frac{1}{n} \text{ or } \int_0^t \left[U'_-(D_s)\right]^2 d\langle D\rangle_s \geq n\right\}.$$

For all $t \in [0,T]$, $\int_0^{t\wedge R_n} \left[U'_-(D_s)\right]^2 d\langle D\rangle_s \leq n$ and $\mathbb{E}\left(\int_0^{t\wedge R_n} U'_-(D_s)dD_s\right) = 0$. Moreover $R_n \nearrow R$ as $n \to \infty$.

Let $t \in [0,T]$. By Tanaka's formula,

$$U(D_{t\wedge R_n}) = U(D_0) + \int_0^{t\wedge R_n} U'_-(D_s)dD_s + \frac{1}{2}\int_{(0,+\infty)} L^r_{t\wedge R_n}(D)U''(dr). \tag{A.3}$$

The assumption that $H_U(P_0|Q_0) < \infty$ and Remark 2.1 a) imply that $(U(D_s))_{s\in[0,T]}$ is a uniformly integrable $\mathbb{Q}^T$-submartingale. Since the $\mathbb{Q}^T$-expectation of the stochastic integral is zero, one deduces

$$\mathbb{E}^T\left(U(D_{t\wedge R_n})\right) = \mathbb{E}^T(U(D_0)) + \frac{1}{2}\mathbb{E}^T\left(\int_{(0,+\infty)} L^r_{t\wedge R_n}(D)U''(dr)\right).$$

When $n \to \infty$, since U is continuous on $(0,+\infty)$ by convexity, $U(D^T_{t\wedge R_n})$ converges to $U(D_{t\wedge R}) + \Delta U(0)1_{\{0<R\leq t\}} = U(D_t) + \Delta U(0)1_{\{0<R\leq t\}}$ and by uniform integrability, $\mathbb{E}^T(U(D_{t\wedge R_n}))$ converges to $\mathbb{E}^T(U(D_t)) + \Delta U(0)\mathbb{Q}^T(0 < R \leq t)$. Dealing with the expectation of the integral in

the right-hand-side by monotone convergence, one obtains

$$\mathbb{E}^T(U(D_t)) = \mathbb{E}^T(U(D_0)) - \Delta U(0)\mathbb{Q}^T(0 < R \le t) + \frac{1}{2}\mathbb{E}^T\left(\int_{(0,+\infty)} L^r_{t\wedge R}(D)U''(dr)\right).$$

Since according to Lemma 2.3 b), D is equal to zero on $[R,T]$, one can replace $t \wedge R$ by t in the last expectation. Replacing t by $T-t$ in this equation, one gets (2.8). Moreover $\mathbb{Q}^T$ a.s., $\int_{(0,+\infty)} L^r_t(D)U''(dr)$ is the finite limit of the integral with respect to $U''(dr)$ in the right-hand-side of (A.3) as $n \to \infty$. Since the left-hand-side converges to $U(D_t) + \Delta U(0)1_{\{0<R\le t\}}$ we deduce that the stochastic integral in the right-hand-side also has a finite limit. Hence $\int_0^{t\wedge R}[U'(D_s)]^2\, d\langle D\rangle_s < +\infty$, $\int_0^{t\wedge R} U'(D_s)dD_s$ makes sense and (2.5) holds. When U is continuous on $[0,+\infty)$ and C^2 on $(0,+\infty)$, (2.6) follows by the occupation times formula and Lemma 2.3 b) and (2.8) written for $t=0$ combined with the same arguments imply that

$$\begin{aligned} H_U(P_0|Q_0) =& H_U(P_T|Q_T) \\ &+ \frac{1}{2}\mathbb{E}^T\left(\int_0^T U''(D_s)1_{\{s<R\}}\left(\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)ds\right). \end{aligned}$$

Since Y_s admits the density q_{T-s} and for almost all $s \in [0,T)$, $D_s = \frac{p_{T-s}}{q_{T-s}}(Y_s)$ and $\{R > s\} = \{\frac{p_{T-s}}{q_{T-s}}(Y_s) > 0\}$, by changing variables $s \mapsto T-s$ we have established (2.9).

A.3 Proof of Corollary 2.7

We notice first that

$$\forall \delta \in (0,1),\ \mathbb{E}^T\int_0^T \mathbf{1}_{|D_s-1|<\delta}\left(\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^* \bar{a}(s,Y_s)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)ds < \infty. \tag{A.4}$$

Indeed, for $\delta \in (0,1)$, we can easily construct a C^2 convex function $\hat{U}$ on $\mathbb{R}$ such that $\forall r \in \mathbb{R}$, $0 \le \hat{U}(r) \le |r-1|$ and $\forall r \in [1-\delta, 1+\delta]$, $\hat{U}''(r) \ge \alpha$ for some $\alpha > 0$, so that the integral in (A.4) is bounded thanks to (2.9) by $\frac{1}{\alpha}H_{\hat{U}}(P_0|Q_0) \le \frac{1}{\alpha}\|P_0 - Q_0\|_{\mathrm{TV}}$. For $r \in \mathbb{R}$, since

$$L^r_t(D) = 2\left((D_t - r)^+ - (D_0 - r)^+ - \int_0^t \mathbf{1}_{D_s>r}dD_s\right),$$

by Doob's inequality we obtain $|\mathbb{E}^T(L^r_t(D) - L^r_t(D))| \le 4|r-1| + 2\left(\mathbb{E}^T\int_0^t \mathbf{1}_{\{1\wedge r<D_s\le r\vee 1\}}d\langle D\rangle_s\right)^{1/2}$. Hence Lemma 2.3 b) and (A.4) imply that $r \mapsto \mathbb{E}^T(L^r_t(D))$ is continuous (and finite) at $r=1$. With the occupation times formula, one deduces that

$$\begin{aligned} 2\mathbb{E}^T(L^1_t(D)) &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{1-\varepsilon}^{1+\varepsilon}\mathbb{E}^T(L^r_t(D))dr \\ &= \lim_{\varepsilon\to 0}\mathbb{E}^T\frac{1}{\varepsilon}\int_0^t \mathbf{1}_{\{|D_s-1|<\varepsilon\}}\left(\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)\right)^*\bar{a}(s,Y_s)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](Y_s)ds \\ &= \lim_{\varepsilon\to 0}\int_0^t\frac{1}{\varepsilon}\int_{\{|\frac{p_{T-s}}{q_{T-s}}(x)-1|<\varepsilon\}}\nabla^*\left[\frac{p_{T-s}}{q_{T-s}}\right](x)\bar{a}(s,x)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](x)q_{T-s}(x)dxds. \end{aligned}$$

Define now the function $\varphi_\varepsilon(r) := \mathbf{1}_{[-\varepsilon,\varepsilon]}(r)r\varepsilon^{-1} + \mathbf{1}_{(\varepsilon,\infty)}(r) - \mathbf{1}_{(-\infty,-\varepsilon)}(r)$. Since the function $\varepsilon \mapsto \int_0^t\int_{\{|\frac{p_{T-s}}{q_{T-s}}(x)-1|\le\varepsilon\}} q_{T-s}(x)dxds$ is increasing and right continuous, we can chose $\varepsilon_k \searrow 0$ a

sequence with $\int_0^t \int_{\{|\frac{p_{T-s}}{q_{T-s}}(x)-1|=\varepsilon_k\}} q_{T-s}(x)dxds = 0$ so that

$$\begin{aligned}
2\mathbb{E}^T(L_t^1(D)) &= \lim_{k\to\infty} \int_0^t \int_{\mathbb{R}^d} \nabla^* \left[\varphi_{\varepsilon_k}\left(\frac{p_{T-s}}{q_{T-s}}-1\right)\right](x)\bar{a}(s,x)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](x)q_{T-s}(x)dxds \\
&= -\lim_{k\to\infty} \int_0^t \int_{\mathbb{R}^d} \varphi_{\varepsilon_k}\left(\frac{p_{T-s}}{q_{T-s}}-1\right)(x)\nabla\cdot\left[\bar{a}(s,x)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](x)q_{T-s}(x)\right]dxds \\
&= -\int_0^t \int_{\mathbb{R}^d} \widetilde{sign}\left(\frac{p_{T-s}}{q_{T-s}}-1\right)(x)\nabla\cdot\left[\bar{a}(s,x)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](x)q_{T-s}(x)\right]dxds
\end{aligned}$$

where the last equality is a consequence of the integrability assumption made on $\nabla\cdot\left[\bar{a}(s,x)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](x)q_{T-s}(x)\right]$. To justify the integration by parts at the second equality, we introduce functions $\phi_n \in C_0^\infty(\mathbb{R}^d)$ such that $\mathbf{1}_{B(0,r_n)} \le \phi_n \le \mathbf{1}_{B(0,2r_n)}$ and $0 \le |\nabla\phi_n| \le 2/r_n$, and functions $\varphi_{\varepsilon_k,m} : \mathbb{R} \to \mathbb{R}$ of class C^1 such that $\varphi_{\varepsilon_k,m} \to \varphi_{\varepsilon_k}$, $|\varphi_{\varepsilon_k,m}| \le |\varphi_{\varepsilon_k}|$ on $\mathbb{R}$ and $\varphi'_{\varepsilon_k,m} \to \varphi'_{\varepsilon_k}$, $|\varphi'_{\varepsilon_k,m}| \le |\varphi'_{\varepsilon_k}|$ on $\mathbb{R}\backslash\{-\varepsilon_k,+\varepsilon_k\}$ as $m\to\infty$, and using the assumptions, (A.4) and the choice of ε_k, we take the limits $n\to\infty$ then $m\to\infty$ by dominated convergence in the equality

$$\begin{aligned}
&\int_{\mathbb{R}^d} \varphi'_{\varepsilon_k,m}\left(\frac{p_{T-s}}{q_{T-s}}-1\right)(x)\nabla^*\left[\frac{p_{T-s}}{q_{T-s}}\right](x)a(T-s,x)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](x)q_{T-s}(x)\phi_n(x)dx \\
&= -\int_{\mathbb{R}^d} \varphi_{\varepsilon_k,m}\left(\frac{p_{T-s}}{q_{T-s}}-1\right)(x)\nabla\cdot\left(a(T-s,x)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](x)q_{T-s}(x)\right)\phi_n(x)dx \\
&\quad -\int_{\mathbb{R}^d} \varphi_{\varepsilon_k,m}\left(\frac{p_{T-s}}{q_{T-s}}-1\right)(x)\nabla^*\phi_n(x)a(T-s,x)\nabla\left[\frac{p_{T-s}}{q_{T-s}}\right](x)q_{T-s}(x)dx.
\end{aligned}$$

A.4 Proof of Proposition 2.9

By Theorem 1.3.8 [15], if $(X_t)_{t\ge0}$ is Feller the tail sigma field is trivial if $\|P_t - Q_t\|_{TV} \to 0$ as $t\to\infty$ for all pair of initial laws P_t and Q_t. Since $\|P_t - Q_t\|_{TV} \le \|P_t - p_\infty dx\|_{TV} + \|p_\infty dx - Q_t\|_{TV}$ and, by Theorem 2.1.3 p.162 [4], the local uniform ellipticity assumption ensures that P_t admits a density with respect to the Lebesgue measure for any positive t, it is enough to show that $(X_t)_{t\ge0}$ is Feller and that $\|P_t - p_\infty dx\|_{TV} \to 0$ as $t\to\infty$ when P_0 admits a density p_0 with respect to the Lebesgue measure.

We first check the convergence to the invariant density in total variation norm. For $k\in\mathbb{N}^*$ consider the probability density

$$p_0^k(x) = (p_0(x)\wedge kp_\infty(x)) + p_\infty(x)\int_{p_0>kp_\infty}(p_0(y)-kp_\infty(y))dy.$$

Since p_∞ is positive, on one hand we have $\lim_{k\to\infty}\|p_0 - p_0^k\|_1 = 0$ and $p_0^k \le (k+1)p_\infty$. On the other hand, the total variation distance between the marginal laws at time t of the solutions to (0.1) started from the initial densities p_0 and p_0^k is not larger than $\|p_0 - p_0^k\|_1$. Therefore we can moreover restrict ourselves to the case when $\frac{p_0}{p_\infty}$ is bounded. Then,

$$\int_{\mathbb{R}^d}\left(\frac{p_0}{p_\infty}(x)-1\right)^2 p_\infty(x)dx \le \left(\int_{\mathbb{R}^d}\left(\frac{p_0}{p_\infty}(x)-1\right)^4 p_\infty(x)dx\right)^{1/2} < +\infty.$$

We set $Q_0 = p_\infty dx$. By Remarks 2.5 a) and 2.5 c), conditions $H1)$, $H2)_Q$, $H3)_Q$ and $H3)_P$ hold and for each $T>0$, $\mathbb{Q}^T$ is an extremal solution of the martingale problem (MP). Applying

Theorem 2.4 respectively with $U(r) = (r-1)^4$ and $U(r) = (r-1)^2$, one obtains that $t \mapsto \int_{\mathbb{R}^d} \left(\frac{p_t}{p_\infty}(x) - 1\right)^2 p_\infty(x)dx$ is non-increasing and that

$$\sup_{t\geq 0} \int_{\mathbb{R}^d} \left(\frac{p_t}{p_\infty}(x) - 1\right)^4 p_\infty(x)dx + \int_0^\infty \int_{\{\frac{p_t}{p_\infty}(x)>0\}} \left(\nabla^* \left[\frac{p_t}{p_\infty}\right] a \nabla \left[p_0 \frac{p_t}{p_\infty}\right]\right)(x) p_\infty(x)dxdt < +\infty. \tag{A.5}$$

Since a is locally uniformly elliptic, the proof of Lemma A.1 ensures that dt a.e., the gradient ∇p_t (resp. ∇p_∞) of p_t (resp. p_∞) in the sense of distributions is a locally integrable function on $\mathbb{R}^d$ that vanishes a.e. on $\{x : p_t(x) = 0\}$. Moreover, one can choose therein $K^p(t,x) = \mathbf{1}_{\{p_t(x)>0\}} \frac{\nabla p_t}{p_t}(x)$ and $K^q(t,x) = \frac{\nabla p_\infty}{p_\infty}(x)$. Then, in (A.5), $\nabla \left[\frac{p_t}{p_\infty}\right] = \frac{\nabla p_t}{p_\infty} - \frac{p_t \nabla p_\infty}{p_\infty^2}$ is a.e. equal to 0 when $\frac{p_t}{p_\infty}$ is equal to 0 so that the restriction of the spatial integration to $\{\frac{p_t}{p_\infty}(x) > 0\}$ can be removed. Since p_∞ is assumed to be locally Lipschitz continuous and bounded away from 0, the function $\frac{1}{p_\infty}$ is locally bounded with a locally bounded gradient in the sense of distributions equal to $-\frac{\nabla p_\infty}{p_\infty^2}$. We deduce that the gradient $\nabla \frac{p_t}{p_\infty}$ of $\frac{p_t}{p_\infty}$ in the sense of distributions is equal to $\frac{\nabla p_t}{p_\infty}(x) - \frac{p_t \nabla p_\infty}{p_\infty^2}$ and therefore to $\nabla \left[\frac{p_t}{p_\infty}\right]$.

From the finiteness of the time-integral in (A.5), we deduce the existence of a sequence $(t_n)_n$ tending to $+\infty$ such that $\lim_{n\to\infty} \int_{\mathbb{R}^d} \left(\nabla^* \frac{p_{t_n}}{p_\infty} a \nabla \frac{p_{t_n}}{p_\infty}\right)(x) p_\infty(x)dx = 0$. For $A > 0$, writing the integral on $\mathbb{R}^d$ as the sum of the integrals on the ball $B(0,A)$ and its complementary $B(0,A)^c$, one has

$$\begin{aligned}
&\int_{\mathbb{R}^d} \left(\frac{p_{t_n}}{p_\infty}(x) - 1\right)^2 p_\infty(x)dx \\
&\leq \int_{B(0,A)} \left(\frac{p_{t_n}}{p_\infty}(x) - \frac{\int_{B(0,A)} p_{t_n}(y)dy}{\int_{B(0,A)} p_\infty(y)dy}\right)^2 p_\infty(x)dx + \frac{\left(\int_{B(0,A)} (p_{t_n} - p_\infty)(y)dy\right)^2}{\int_{B(0,A)} p_\infty(y)dy} \\
&+ \left(\int_{B(0,A)^c} \left(\frac{p_{t_n}}{p_\infty}(x) - 1\right)^4 p_\infty(x)dx \int_{B(0,A)^c} p_\infty(x)dx\right)^{1/2} \\
&\leq \int_{B(0,A)} \left(\frac{p_{t_n}}{p_\infty}(x) - \frac{\int_{B(0,A)} \frac{p_{t_n}}{p_\infty}(y)dy}{\int_{B(0,A)} dy}\right)^2 p_\infty(x)dx + \frac{\left(\int_{B(0,A)^c} (\frac{p_{t_n}}{p_\infty}(y) - 1) p_\infty(y)dy\right)^2}{\int_{B(0,A)} p_\infty(y)dy} \\
&+ \left(\int_{\mathbb{R}^d} \left(\frac{p_0}{p_\infty}(x) - 1\right)^4 p_\infty(x)dx \int_{B(0,A)^c} p_\infty(x)dx\right)^{1/2}.
\end{aligned}$$

Since $\left(\int_{B(0,A)^c} (\frac{p_{t_n}}{p_\infty}(y) - 1) p_\infty(y)dy\right)^2 \leq \int_{\mathbb{R}^d} (\frac{p_0}{p_\infty}(y) - 1)^2 p_\infty(y)dy \int_{B(0,A)^c} p_\infty(y)dy$, the sum of the last two terms on the right-hand-side tends to 0 uniformly in n as $A \to \infty$. Using (2.10) and denoting by $C_A < +\infty$ the constant of the Poincaré-Wirtinger inequality satisfied by the Lebesgue measure on the ball $B(0,A)$, one checks that the first term is smaller than

$$C_A \frac{\sup_{B(0,A)} p_\infty}{\varepsilon_A \inf_{B(0,A)} p_\infty} \int_{\mathbb{R}^d} \left(\nabla^* \frac{p_{t_n}}{p_\infty} a \nabla \frac{p_{t_n}}{p_\infty}\right)(x) p_\infty(x)dx$$

which tends to 0 as $n \to \infty$. Hence $\lim_{n\to\infty} \int_{\mathbb{R}^d} \left(\frac{p_{t_n}}{p_\infty}(x) - 1\right)^2 p_\infty(x)dx = 0$. Since $\|p_t - p_\infty\|_1^2 \leq \int_{\mathbb{R}^d} \left(\frac{p_t}{p_\infty}(x) - 1\right)^2 p_\infty(x)dx$ where the right-hand-side is non-increasing with t, we conclude that $\lim_{t\to\infty} \|p_t - p_\infty\|_1 = 0$.

It remains us to show that $(X_t)_{t\geq 0}$ is Feller. Using Itô's calculus and Gronwall's Lemma we check under the assumptions on the coefficients that the solution X_t^x of (0.1) starting from $x \in \mathbb{R}^d$ satisfies $\mathbb{E}\left((1+|X_t^x|^2)^{-1}\right) \leq C(1+|x|^2)^{-1}$ for some $C > 0$. Then, for any continuous function $f : \mathbb{R}^d \to \mathbb{R}$ going to 0 at infinity the inequality

$$|\mathbb{E}(f(X_t^x))| \leq \sup_{|y|\leq A} |f(y)| C \frac{(1+A^2)}{(1+|x|^2)} + \sup_{|y|>A} |f(y)|$$

for all $A > 0$ (following from the previous estimate and Markov's inequality) implies that $\mathbb{E}(f(X_t^x)) \to 0$ when $x \to \infty$.

A.5 Sufficient conditions for superquadratic potentials to satisfy $H1)''$

Lemma A.5 *Let $b(x) = -\nabla V(x)$ for a nonnegative C^2 potential V in $\mathbb{R}^d$ satisfying (2.7), and σ be any globally Lipschitz continuous choice of the square root of the identity I_d. Then, condition H1)" holds for the diffusion process $dX_t = \sigma(X_t)dW_t - \nabla V(X_t)dt$.*

Proof . Computing $d|X_t|^2$, one sees that the first condition in (2.7) prevents explosion for the SDE which has locally Lipschitz coefficients. Since for $c > 0$,

$$de^{cV(X_t)} = e^{cV(X_t)} \left(c\nabla^* V(X_t)\sigma(X_t)dW_t + \frac{c}{2}[\Delta V + (c-2)|\nabla V|^2](X_t)dt \right),$$

the second condition ensures that for c small enough, $\mathbb{E}(e^{cV(X_t)}) \leq e^{K(c)t}\mathbb{E}(e^{cV(X_0)})$ for some finite constant $K(c)$ only depending on V and c. The third assumption ensures the existence of a finite constant $\tilde{K}(\frac{c}{T})$ only depending on $\frac{c}{T}$ and V such that

$$\mathbb{E}\left(\exp(4\int_0^T \sqrt{\partial_{ik}V\partial_{ik}V(X_t)}dt)\right) \leq \tilde{K}(\frac{c}{T})\mathbb{E}\left(\exp(\frac{c}{T}\int_0^T V(X_t)dt)\right).$$

By Jensen's inequality, one deduces that

$$\mathbb{E}\left(\exp(4\int_0^T \sqrt{\partial_{ik}V\partial_{ik}V(X_t)}dt)\right) \leq \frac{\tilde{K}(\frac{c}{T})}{T}\int_0^T \mathbb{E}(e^{cV(X_t)})dt \leq \tilde{K}(\frac{c}{T})e^{K(c)T}\mathbb{E}(e^{cV(X_0)}).$$

∎

B Proofs of the main results of Section 3

B.1 Proof of Proposition 3.2

We will make use of the stochastic flow defined by the two-parameter process $\xi_t(x)$ satisfying

$$d\xi_t^i(x) = \sigma_{ik}(\xi_t(x))d\bar{W}_t^k + \bar{b}_i(\xi_t(x))dt, \quad (t,x) \in [0,T) \times \mathbb{R}^d, \ i = 1, \dots d, \tag{B.1}$$

and $\xi_0(x) = x$, noting that $\xi_t(Y_0) = Y_t$. We shall also deal with the family of continuous $\mathcal{G}_t - \mathbb{P}_\infty^T -$ local martingales $(D_t(x) : t \in [0,T])_{x\in\mathbb{R}^d}$ defined by

$$dD_t(x) = [\sigma_{ik}\partial_i\rho](t,\xi_t(x))d\bar{W}_t^k \quad , \quad D_0(x) = \frac{p_T}{p_\infty}(x) = \rho_0(x). \tag{B.2}$$

According to Lemma 2.3, $D_t(Y_0)$ is equal to the process D_t defined in (2.2). Writing $\nabla\rho_t(\xi_t(x)) = (\nabla_x^*\xi_t(x))^{-1}\nabla_x[\rho_t(\xi_t(x))]$ we remark that, thanks to the Itô product rule, $d\nabla\rho_t(\xi_t(x))$ can be obtained with by computing $d(\nabla_x\xi_t(x))^{-1}$ and $d\nabla_x[\rho_t(\xi_t(x))]$, which is part of the content of the two next Lemmas:

Lemma B.1 *The process $(t,x)\mapsto\xi_t(x)$ has a $\mathbb{P}_\infty^T$ a.s. continuous version such that the mapping $x\mapsto\xi_t(x)$ is a global diffeomorphism of class $C^{1,\alpha}$ for some $\alpha\in(0,1)$ and every $t\in[0,T]$. Moreover, we have*

$$d\partial_j\xi_t^i(x) = \partial_p\sigma_{ik}(t,\xi_t(x))\partial_j\xi_t^p(x)d\bar{W}_t^k + \partial_p\bar{b}_i(t,\xi_t(x))\partial_j\xi_t^p(x)dt, \quad (t,x)\in[0,T)\times\mathbb{R}^d \tag{B.3}$$

with $\partial_j\xi_0^i(x)=\delta_{ij}$. Finally, writing $\nabla\xi_t(x)=(\partial_j\xi_t^i(x))_{ij}$, it holds that

$$\begin{aligned} d(\nabla\xi_t(x))_{kl}^{-1} = &-(\nabla\xi_t(x))_{ki}^{-1}[\partial_l\sigma_{ir}](\xi_t(x))d\bar{W}_t^r - \nabla\xi_t(x))_{ki}^{-1}[\partial_l\bar{b}_i](\xi_t(x))dt \\ &+(\nabla\xi_t(x))_{ki}^{-1}[\partial_m\sigma_{ir}\partial_l\sigma_{mr}](\xi_t(x))dt, \qquad (t,x)\in[0,T)\times\mathbb{R}^d. \end{aligned} \tag{B.4}$$

Proof . Under assumptions $H4)$ and $H5)_{p_\infty}$, classic results by Kunita [15] (see Theorem 4.7.2) imply the asserted regularity properties of the stochastic flow, as well as the $\mathbb{P}_\infty^T$ a.s. existence of the inverse matrix $(\nabla\xi_t(x))^{-1}$ for all $(t,x)\in[0,T]\times\mathbb{R}^d$. Since the smooth map $A\mapsto A^{-1}$, defined on non singular $d\times d$ matrices, has first and second derivatives respectively given by the linear and bilinear operators $F\mapsto -A^{-1}FA^{-1}$ and $(F,K)\mapsto A^{-1}FA^{-1}KA^{-1}+A^{-1}KA^{-1}FA^{-1}$ (where F,K are generic square-matrices), we deduce that for $A=(A_{ij})_{i,j=1\dots d}$,

$$\frac{\partial(A^{-1})_{kl}}{\partial A_{ij}} = -A_{ki}^{-1}A_{jl}^{-1}, \quad \text{and} \quad \frac{\partial^2(A^{-1})_{kl}}{\partial A_{ij}\partial A_{mn}} = A_{ki}^{-1}A_{jm}^{-1}A_{nl}^{-1} + A_{km}^{-1}A_{ni}^{-1}A_{jl}^{-1}$$

for all $k,l,i,j,m,n\in\{1,\dots,d\}$. Equation B.4 follows by applying Itô's formula to each of the functions $A\mapsto(A^{-1})_{kl}$ and the semimartingales $(\partial_j\xi_t^i(x))$, $i,j=1\dots d$.

∎

Lemma B.2 *The process $D_t(x)$ has a modification still denoted $D_t(x)$ such that $\mathbb{P}_\infty^T$ a.s. the function $(t,x)\mapsto D_t(x)$ is continuous and $x\mapsto D_t(x)$ is of class C^1 for each t. This modification is indistinguishable from $(\rho_t(\xi_t(x)):(t,x)\in[0,T)\times\mathbb{R}^d)$ and we have*

$$d\partial_k D_t(x) = \partial_m\left[\sigma_{ir}\partial_i\rho\right](t,\xi_t(x))\partial_k\xi_t^m(x)d\bar{W}_t^r = d\left[\partial_m\rho(t,\xi_t(x))\partial_k\xi_t^m(x)\right] \tag{B.5}$$

for all $(t,x)\in[0,T)\times\mathbb{R}^d$.

Proof . Thanks to the regularity of $x\mapsto\xi_t(x)$ established in Lemma B.1 and assumptions $H5)_{p_\infty}$ and $H6)_{p_0}^T$, the statements follow from Theorem 3.3.3 of Kunita [15] (see also Exercise 3.1.5 therein). ∎

We can now proceed to prove Proposition 3.2. Evaluating expressions (B.4) and (B.5) in $x=Y_0$, we obtain using Itô's product rule that

$$\begin{aligned} d\partial_l\rho_t(Y_t) &= \left[\sigma_{kr}\partial_{lk}\rho\right](t,Y_t)d\bar{W}_t^r - \left[\sigma_{kr}\partial_{kj}\rho\partial_l\sigma_{jr} + \partial_k\rho\partial_l\bar{b}_k\right](t,Y_t)dt \\ &= \left[\sigma_{kr}\partial_{lk}\rho\right](t,Y_t)d\bar{W}_t^r - \left[\frac{1}{2}\partial_{kj}\rho\partial_l a_{kj} + \partial_k\rho\partial_l\bar{b}_k\right](t,Y_t)dt. \end{aligned} \tag{B.6}$$

For the remaining of the proof, the argument (t, Y_t) will be omitted for notational simplicity. By Itô 's formula we get $d\sigma_{li} = [\sigma_{mr}\partial_m\sigma_{li}]\, d\bar{W}^r_t + \left[\bar{b}_m\partial_m\sigma_{li} + \frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}\right] dt$. We then have

$$\begin{aligned} d\left[\sigma_{li}\partial_l\rho\right] &= \sigma_{li}d\partial_l\rho + \partial_l\rho d\sigma_{li} + d\langle\partial_l\rho, \sigma_{li}\rangle \\ &= \partial_k\left[\partial_l\rho\sigma_{li}\right]\sigma_{kr}d\bar{W}^r + \partial_l\rho\left[\bar{b}_m\partial_m\sigma_{li} + \frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}\right] - \sigma_{li}\left[\sigma_{kr}\partial_{kj}\rho\partial_l\sigma_{jr} + \partial_k\rho\partial_l\bar{b}_k\right] \\ &\quad + \ a_{mk}\partial_{lk}\rho\partial_m\sigma_{li} \end{aligned}$$

where we used in the stochastic integral the fact that $\partial_l\rho\sigma_{mr}\partial_m\sigma_{li} + \sigma_{li}\sigma_{kr}\partial_{lk}\rho = \partial_l\rho\sigma_{kr}\partial_k\sigma_{li} + \sigma_{li}\sigma_{kr}\partial_{lk}\rho = \partial_k\left[\partial_l\rho\sigma_{li}\right]\sigma_{kr}$. It follows that

$$\begin{aligned} d\left[\nabla^*\rho a\nabla\rho\right] &= d\left[\sigma_{li}\partial_l\rho\ \sigma_{l'i}\partial_{l'}\rho\right] \\ &= 2\ \sigma_{l'i}\ \partial_{l'}\rho\partial_k\left[\sigma_{li}\partial_l\rho\right]\sigma_{kr}d\bar{W}^r + \ 2\Big\{\left[\sigma_{l'i}\partial_{l'}\rho a_{mk}\partial_m\sigma_{li}\partial_{lk}\rho\right] \\ &\quad + \sigma_{l'i}\partial_{l'}\rho\partial_l\rho\left[\bar{b}_m\partial_m\sigma_{li} + \frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}\right] - a_{ll'}\partial_{l'}\rho\left[\sigma_{kr}\partial_{kj}\rho\partial_l\sigma_{jr} + \partial_k\rho\partial_l\bar{b}_k\right]\Big\}dt \\ &\quad + a_{kk'}\partial_k\left[\partial_l\rho\sigma_{li}\right]\partial_{k'}\left[\partial_{l'}\rho\sigma_{l'i}\right]dt \end{aligned}$$

On the other hand, using (B.2) at $x = Y_0$ we have $dU''_\delta(\rho) = U^{(3)}_\delta(\rho)\sigma_{nr}\partial_n\rho\, d\bar{W}^r + \frac{1}{2}U^{(4)}_\delta(\rho)a_{nj}\partial_n\rho\partial_j\rho\, dt$ which combined with the previous expression yields

$$\begin{aligned} d\left[U''_\delta(\rho)\nabla^*\rho a\nabla\rho\right] =& 2U''_\delta(\rho)\Big\{\left[\sigma_{l'i}\partial_{l'}\rho a_{mk}\partial_m\sigma_{li}\partial_{lk}\rho\right] + \sigma_{l'i}\partial_{l'}\rho\partial_l\rho\left[\bar{b}_m\partial_m\sigma_{li} + \frac{1}{2}a_{mk}\partial_{mk}\sigma_{li}\right] \\ &\qquad - a_{ll'}\partial_{l'}\rho\left[\sigma_{kr}\partial_{kj}\rho\partial_l\sigma_{jr} + \partial_k\rho\partial_l\bar{b}_k\right]\Big\}dt \ + \ d\hat{M}^{(\delta)} \\ &+ \ U''_\delta(\rho)a_{kk'}\partial_k\left[\partial_l\rho\sigma_{li}\right]\partial_{k'}\left[\partial_{l'}\rho\sigma_{l'i}\right]dt + \frac{1}{2}U^{(4)}_\delta(\rho)\left|\nabla^*\rho a\nabla\rho\right|^2 dt \\ &+ \ 2U^{(3)}_\delta(\rho)\sigma_{l'i}\partial_{l'}\rho\partial_k\left[\sigma_{li}\partial_l\rho\right]a_{jk}\partial_j\rho dt. \end{aligned} \tag{B.7}$$

Equivalently,

$$\begin{aligned} d\left[U''_\delta(\rho)\nabla^*\rho a\nabla\rho\right] =& 2U''_\delta(\rho)\Big\{\partial_{l'}\rho\partial_l\rho\Big[\frac{1}{4}(\partial_k\sigma_{lj}a_{km}\partial_m\sigma_{l'j} - \sigma_{ki}\partial_k\sigma_{lj}\sigma_{mj}\partial_m\sigma_{l'i}) \\ &\qquad + \frac{1}{2}\bar{b}_m\partial_m a_{ll'} + \frac{1}{2}\sigma_{l'i}a_{mk}\partial_{mk}\sigma_{li} - a_{kl'}\partial_k\bar{b}_l\Big] \\ &\qquad + \left[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}\right]\partial_{l'}\rho\partial_m\sigma_{li}\partial_{kl}\rho\Big\}dt \ + \ d\hat{M}^{(\delta)} + tr[\Lambda_\delta\Gamma]dt. \end{aligned}$$

B.2 Proof of Theorem 3.4

Let us check (3.2). Since U'' is continuous and non increasing in $(0, \infty)$ by Remark 3.1, one has $U''_\delta(r) \nearrow U''(r)$ for each $r > 0$ as $\delta \to 0$. It is therefore enough to obtain (the integrated version of) inequality (3.2) with U''_δ instead of U'', as monotone convergence allows us to pass to the

limit as $\delta \to 0$ on both sides. For $0 \le r \le t < T$ we have by Proposition 3.2 that

$$\begin{aligned}
&[U''_\delta(\rho)\nabla^*\rho a\nabla\rho](t,Y_t) - [U''_\delta(\rho)\nabla^*\rho a\nabla\rho](r,Y_r)\\
&\quad\ge \hat{M}^{(\delta)}_t - \hat{M}^{(\delta)}_r + 2\int_r^t U''_\delta(\rho)\left[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}\right]\partial_{l'}\rho\partial_m\sigma_{li}\partial_{kl}\rho ds\\
&\quad + 2\int_r^t U''_\delta(\rho)\partial_{l'}\rho\partial_l\rho\Big(\frac{1}{4}(\partial_k\sigma_{lj}a_{km}\partial_m\sigma_{l'j} - \sigma_{ki}\partial_k\sigma_{lj}\sigma_{mj}\partial_m\sigma_{l'i})\\
&\qquad\qquad + \frac{1}{2}\left[\bar{b}_m\partial_m a_{ll'} + \sigma_{l'i}a_{mk}\partial_{mk}\sigma_{li}\right] - a_{ml'}\partial_m\bar{b}_l\Big)ds.
\end{aligned} \tag{B.8}$$

Since $\partial_{kl'}\rho U''_\delta(\rho)\left[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}\right] = 0$ and

$$\partial_k(U''_\delta(\rho))\partial_{l'}\rho\left[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}\right] = U^{(3)}_\delta(\rho)\partial_k\rho\partial_{l'}\rho\left[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}\right] = 0,$$

one has

$$\begin{aligned}
U''_\delta(\rho)\left[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}\right]\partial_{l'}\rho\partial_m\sigma_{li}\partial_{kl}\rho &= \frac{1}{p_\infty}\partial_k\left(\partial_l\rho\partial_{l'}\rho U''_\delta(\rho)\left[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}\right]\partial_m\sigma_{li}\,p_\infty\right)\\
&\quad - \frac{\partial_l\rho\partial_{l'}\rho U''_\delta(\rho)}{p_\infty}\partial_k\left([a_{mk}\sigma_{l'i} - \sigma_{ki}a_{ml'}]\partial_m\sigma_{li}p_\infty\right).
\end{aligned} \tag{B.9}$$

Setting

$$\begin{aligned}
\Sigma_{ll'} &\stackrel{\text{def}}{=} \frac{1}{4}(\partial_k\sigma_{lj}a_{km}\partial_m\sigma_{l'j} - \sigma_{ki}\partial_k\sigma_{lj}\sigma_{mj}\partial_m\sigma_{l'i}) + \frac{1}{2}\left[\bar{b}_m\partial_m a_{ll'} + \sigma_{l'i}a_{mk}\partial_{mk}\sigma_{li}\right]\\
&\quad - a_{ml'}\partial_m\bar{b}_l - \frac{1}{p_\infty}\partial_k\left[\left(\frac{1}{2}a_{mk}\partial_m a_{ll'} - \sigma_{ki}a_{ml'}\partial_m\sigma_{li}\right)p_\infty\right]
\end{aligned}$$

we deduce that

$$\begin{aligned}
&[U''_\delta(\rho)\nabla^*\rho a\nabla\rho](t,Y_t) - [U''_\delta(\rho)\nabla^*\rho a\nabla\rho](r,Y_r)\\
&\ge \hat{M}^{(\delta)}_t - \hat{M}^{(\delta)}_r + 2\int_r^t U''_\delta(\rho)\Sigma_{ll'}\partial_{l'}\rho\partial_l\rho\, ds + 2\int_r^t \frac{1}{p_\infty}\partial_k\left(\partial_l\rho\partial_{l'}\rho U''_\delta(\rho)\left[\sigma_{l'i}a_{mk} - \sigma_{ki}a_{ml'}\right]\partial_m\sigma_{li}\,p_\infty\right)ds.
\end{aligned} \tag{B.10}$$

Using (3.1) and the identity $\sigma_{ki}\partial_{k'}\sigma_{li} = \partial_{k'}a_{kl} - \partial_{k'}\sigma_{ki}\sigma_{li}$ one can check that

$$\begin{aligned}
\Theta_{ll'} =& \frac{1}{2}\bar{b}_{k'}\partial_{k'}a_{ll'} + \frac{1}{2}(a_{kl'}\partial_k\bar{b}_l + a_{kl}\partial_k\bar{b}_{l'}) + \frac{1}{4}a_{k'k}\partial_{k'k}a_{ll'} - \frac{1}{4}(a_{k'k}\partial_{k'}\sigma_{li}\partial_k\sigma_{l'i} + \sigma_{ki}\partial_k\sigma_{lj}\sigma_{k'j}\partial_{k'}\sigma_{l'i})\\
&+ \frac{1}{2}\sigma_{ki}(\partial_{k'}\sigma_{li}a_{k'l'} + \partial_{k'}\sigma_{l'i}a_{k'l})\partial_k\ln(p_\infty) - \frac{1}{2}a_{k'k}\partial_{k'}a_{ll'}\partial_k\ln(p_\infty)\\
&+ \frac{1}{2}\partial_k[\sigma_{ki}(\partial_{k'}\sigma_{li}a_{k'l'} + \partial_{k'}\sigma_{l'i}a_{k'l}) - a_{k'k}\partial_{k'}a_{ll'}]\\
=& \frac{\Sigma_{ll'} + \Sigma_{l'l}}{2}
\end{aligned} \tag{B.11}$$

and therefore, the second integral on the right hand side of (B.10) rewrites as $2\int_r^t U''_\delta(\rho)\Theta_{ll'}\partial_{l'}\rho\partial_l\rho\, ds$.

Now, the quadratic variation of $\hat{M}^{(\delta)}$ is bounded above in $[0,T)$ by a constant times

$$\int_0^t \left[|U^{(3)}_\delta(\rho)|^2|\nabla^*\rho a\nabla\rho|^3(Y_s) + \left(U''_\delta(\rho)\right)^2\nabla^*(\nabla^*\rho a\nabla\rho)a\nabla(\nabla^*\rho a\nabla\rho)\right](Y_s)ds.$$

This fact and our assumptions imply that $\hat{M}^\delta$ is a martingale in $[0,T)$ for all $\delta > 0$ sufficiently small. Indeed, we have from Remark 3.1 that $U''_\delta(r) \le U''(\delta) \wedge U''(r)$ and $|U^{(3)}_\delta(r)| \le |U^{(3)}(\delta)| \wedge |U^{(3)}(r)|$ for all $r \ge 0$. Therefore (since $U'' > 0$) we have $U''_\delta(r) \le (U''(r) \wedge 1)\mathbf{1}_{U''(\delta)\le 1} + U''(\delta)(U''(r)/U''(\delta)) \wedge 1)\mathbf{1}_{U''(\delta)>1}$ whence $U''_\delta(r) \le (U''(\delta)+1)(U''(r)\wedge 1)$. As $U^{(3)}$ is non decreasing and non positive, either $|U^{(3)}(\delta)| \neq 0$ for all δ sufficiently small, in which case we similarly get $|U^{(3)}_\delta(r)| \le (|U^{(3)}(\delta)|+1)(|U^{(3)}(r)|\wedge 1)$, or otherwise $U^{(3)}_\delta$ identically vanishes for all δ. Assumption $H6')_{p_\infty}$ and the previous then ensure that $\langle M^{(\delta)}\rangle_t$ has finite expectation for $t \in [0,T)$.

In order to conclude that inequality (3.2) holds for the function U_δ , noting that $\nabla\rho_t$ vanishes on $\{\rho_t = 0\}$, it is enough to show that the last integral in (B.10) has (well defined) null expectation. Using (B.9) and Assumption $H6')_{p_\infty}$ we obtain (with the same control for $U''_\delta(r)$ as before) that

$$\begin{aligned}\mathbb{E}^T_\infty \int_r^t \left| \frac{1}{p_\infty} \partial_k \left(\partial_l\rho \partial_{l'}\rho U''_\delta(\rho) \left[\sigma_{l'i} a_{mk} - \sigma_{ki} a_{ml'}\right] \partial_m \sigma_{li}\, p_\infty \right) \right| (Y_s) ds \\ = \int_r^t \int_{\mathbb{R}^d} \left| \partial_k \left(\partial_l\rho \partial_{l'}\rho U''_\delta(\rho) \left[\sigma_{l'i} a_{mk} - \sigma_{ki} a_{ml'}\right] \partial_m \sigma_{li}\, p_\infty \right) \right| dx ds < \infty\end{aligned} \tag{B.12}$$

which shows that the expectation of the last term in (B.10) is well defined. Moreover, the (everywhere defined) spatial divergence of $g(s,x) := \partial_l\rho_s \partial_{l'}\rho_s U''_\delta(\rho_s) \left[\sigma_{l'i} a_{m\bullet} - \sigma_{\bullet i} a_{ml'}\right] \partial_m \sigma_{li}\, p_\infty$ is $L^1(dx, \mathbb{R}^d)$ for a.e. s. For such s and $\phi_n \in C^\infty_0(\mathbb{R}^d)$ such that $0 \le \phi_n \le 1$, $0 \le |\nabla\phi_n| \le 1$, $\phi_n(x) = 1$ for $x \in B(0,n)$ and $\phi_n(x) = 0$ for $x \in B(0,2n)^c$, we have

$$0 = \int_{\mathbb{R}^d} \nabla.(\phi_n(x) g(s,x)) dx = \int_{\mathbb{R}^d} \phi_n(x) \nabla.g(s,x) dx + \int_{\mathbb{R}^d} \nabla\phi_n(x).g(s,x) dx.$$

Since by Lebesgue's theorem, the second term of the right-hand-side tends to 0 as $n \to \infty$, the limit $\int_{\mathbb{R}^d} \nabla.g(s,x)dx$ of the first term is equal to 0.

C Dissipation of the Fisher information : comparison with the computations and results in [2]

In this section we compare our computations and results with those of [2].

The form of the term $tr(\Lambda_\delta\Gamma)$ in Proposition 3.2 is inspired from the term $tr(\mathbf{XY})$ in [2] pp 163-164 where $\mathbf{X} = 2\Lambda_\delta$. One has

$$\begin{aligned}\Gamma_{12} = \ (\nabla^*\rho\ a)_j\ \partial_j(\sigma_{ki}\partial_k\rho)\sigma_{li}\partial_l\rho =& \frac{1}{2}(\nabla^*\rho\ a)_j \left[\partial_j(\sigma_{ki}\partial_k\rho)\sigma_{li}\partial_l\rho + \partial_j(\sigma_{li}\partial_l\rho)\sigma_{ki}\partial_k\rho\right] \\ =& \frac{1}{2}(\nabla^*\rho\ a)_j \partial_j \left[\partial_l\rho a_{kl}\partial_k\rho\right] = \frac{1}{2}(\nabla^*\rho\ a)\nabla(\nabla^*\rho a \nabla\rho)\end{aligned}$$

which, with $\frac{\partial v}{\partial x} := (\partial_j v_i)_{i,j}$ denoting the Jacobian matrix of vector field v, equals

$$\begin{aligned}\frac{1}{2}(\nabla^*\rho\ a)_j \partial_j \left[\partial_k\rho a_{kl}\partial_l\rho\right] =& \frac{1}{2}(\nabla^*\rho\ a)_j \left(\partial_{kj}\rho\ a_{kl}\ \partial_l\rho + \partial_j\left[a_{kl}\ \partial_l\rho\right]\partial_k\rho\right) \\ =& \frac{1}{2}\nabla^*\rho\ a\frac{\partial(\nabla\rho)}{\partial x} a\nabla\rho + \frac{1}{2}\nabla^*\rho\ a\frac{\partial(a\nabla\rho)}{\partial x}^* \nabla\rho\end{aligned}$$

and corresponds to $4\mathbf{Y}_{12}$ in [2] p. 164 (noting that in their notation, $\mathbf{D}(x) = a(x)/2$). Similarly, $\Gamma_{22} = 4\mathbf{Y}_{22}$. However Γ_{11} cannot in general be identified with $4\mathbf{Y}_{11}$. For instance, in the case

of scalar diffusion $\mathbf{D}(x) = a(x)/2 = D(x)I_d$ for some real valued function D, the term $\Gamma_{11}(x)$ above when written in terms of D reads

$$\frac{1}{2}|\nabla D|^2|\nabla\rho|^2 + \frac{1}{2}(\nabla D.\nabla\rho)^2 + 4D\partial_j D\partial_i\rho\partial_{ij}\rho + 4D^2\sum_{ij}(\partial_{ij}\rho)^2$$

for the choice $\sigma(x) = \sqrt{2D(x)}I_d$, whereas

$$4\mathbf{Y}_{11} = 4\left(D^2\sum_{ij}(\partial_{ij}\rho)^2 + \left(\frac{n}{4}-\frac{1}{2}\right)(\nabla\rho\cdot\nabla D)^2 + 2D\partial_j D\partial_i\rho\partial_{ij}\rho - D(\nabla\rho\cdot\nabla D)\triangle\rho + \frac{1}{2}|\nabla D|^2|\nabla\rho|^2\right).$$

Moreover, our term Γ_{11} is non-intrinsic, in the sense that it cannot in general be written in terms of the diffusion matrix a only (without making explicit use of σ), contrary to the term $\mathbf{Y}_{11}$ in the matrix of [2].

We will next check that the criterion in [2] can also be derived from the computations in Proposition 3.2 in the case a is non singular, which amounts to make an alternative choice in the expression for $d\left[U_\delta''(\rho)\nabla^*\rho a\nabla\rho\right]$ of the quantities in the roles of the coefficient Γ_{11} and of the term $\bar\theta$. This will also allow us to compare and combine both criteria.

Recall first that the matrix $\mathbf{D}(x)$ in [2] equals half of our matrix $a(x)$, and notice that our forward drift term writes in their notation as $b = -\mathbf{D}\nabla\phi - \mathbf{D}F + \nabla.\mathbf{D}$ where $(\nabla.\mathbf{D})_i = \partial_j\mathbf{D}_{ij}$, $e^{-\phi} = p_\infty$ is the invariant density, and F a is vector field satisfying $\nabla.(\mathbf{D}Fe^{-\phi}) = 0$. Thus, $\bar b = a\nabla\ln p_\infty + \nabla.\,a - b = -\mathbf{D}\nabla\phi + \mathbf{D}F + \nabla.\mathbf{D}$.

The factor of $U_\delta''(\rho)$ in (B.7) takes the intrinsic form

$$\begin{aligned}
&a_{kk'}\left[\partial_{kl}\rho\sigma_{li}\partial_{k'l'}\rho\sigma_{l'i} + \partial_{kl}\rho\sigma_{li}\partial_{l'}\rho\partial_{k'}\sigma_{l'i} + \partial_l\rho\partial_k\sigma_{li}\partial_{k'l'}\rho\sigma_{l'i} + \partial_l\rho\partial_k\sigma_{li}\partial_{l'}\rho\partial_{k'}\sigma_{l'i}\right]\\
&+ 2\sigma_{l'i}\partial_{l'}\rho a_{k'k}\partial_{k'}\sigma_{li}\partial_{lk}\rho + \partial_l\rho\partial_{l'}\rho a_{kk'}\partial_{k'k}\sigma_{li}\sigma_{l'i} - 2a_{ll'}\partial_{l'}\rho\sigma_{kr}\partial_{kk'}\rho\partial_l\sigma_{k'r} + \bar b_m\partial_m a_{ll'}\partial_l\rho\partial_{l'}\rho - 2a_{ll'}\partial_{l'}\rho\partial_k\rho\partial_l\bar b_k\\
&= a_{kk'}\left[\partial_{kl}\rho\partial_{k'l'}\rho a_{ll'} + 2\partial_{kl}\rho\partial_{l'}\rho\partial_{k'}a_{ll'}\right]\\
&+ \frac{1}{2}a_{kk'}\partial_l\rho\partial_{l'}\rho\partial_{kk'}a_{ll'} - a_{ll'}\partial_{l'}\rho\partial_{kk'}\rho\partial_l a_{kk'} + \bar b_m\partial_m a_{ll'}\partial_l\rho\partial_{l'}\rho - 2a_{ll'}\partial_{l'}\rho\partial_k\rho\partial_l\bar b_k,
\end{aligned}$$

where to the second and third terms in the bracket on the l.h.s., brought together, we have added the first term after the bracket, and moreover the fourth term in the bracket on the l.h.s. was added to the the second term outside the bracket. Hence, writing

$$Q_1 := -a_{ll'}\partial_{l'}\rho\partial_{kk'}\rho\partial_l a_{kk'} + \bar b_m\partial_m a_{ll'}\partial_l\rho\partial_{l'}\rho - 2a_{ll'}\partial_{l'}\rho\partial_k\rho\partial_l\bar b_k,$$

$$Q_2 := a_{kk'}\left[\partial_{kl}\rho\partial_{k'l'}\rho a_{ll'} + 2\partial_{kl}\rho\partial_{l'}\rho\partial_{k'}a_{ll'}\right] + \frac{1}{2}a_{kk'}\partial_l\rho\partial_{l'}\rho\partial_{kk'}a_{ll'},$$

and using the last expression above for Γ_{12}, we can write

$$\begin{aligned}
\frac{1}{2}d\left[U_\delta''(\rho)\nabla^*\rho a\nabla\rho\right] &= \frac{1}{2}d\hat M^{(\delta)} + \frac{U_\delta''(\rho)}{2}(Q_1+Q_2)\,dt + \frac{U_\delta^{(4)}(\rho)}{4}\left|\nabla^*\rho a\nabla\rho\right|^2 dt\\
&\quad + \frac{U_\delta^{(3)}(\rho)}{2}\left(\nabla^*\rho\; a\frac{\partial(\nabla\rho)}{\partial x}a\nabla\rho + \nabla^*\rho\; a\frac{\partial(a\nabla\rho)}{\partial x}^*\nabla\rho\right)dt.\\
=&\frac{1}{2}d\hat M^{(\delta)} + \left[\frac{U_\delta^{(4)}(\rho)}{4}\left|\nabla^*\rho a\nabla\rho\right|^2 + \frac{U_\delta^{(3)}(\rho)}{4}\left(\nabla^*\rho\; a\frac{\partial(\nabla\rho)}{\partial x}a\nabla\rho + \nabla^*\rho\; a\frac{\partial(a\nabla\rho)}{\partial x}^*\nabla\rho\right)\right]dt\\
&+\left[\frac{U_\delta''(\rho)}{2}Q_1\right]dt + \left[\frac{U_\delta''(\rho)}{2}Q_2 + \frac{U_\delta^{(3)}(\rho)}{4}\left(\nabla^*\rho\; a\frac{\partial(\nabla\rho)}{\partial x}a\nabla\rho + \nabla^*\rho\; a\frac{\partial(a\nabla\rho)}{\partial x}^*\nabla\rho\right)\right]dt.
\end{aligned}\tag{C.1}$$

The latter identity yields the expression for the dissipation of entropy dissipation computed in [2]. Indeed, denoting respectively by J_1, J_2 and J_3 the expectations of the first, second and third terms in square brackets in the right-hand side, we observe that J_1 is, up to time reversal $t \mapsto T-t$, exactly equal to the term $\tilde{R}_1$ on top of p. 162 in [2]. Starting from the last expression of T_3 p. 160 and the definition (2.23) of $\tilde{R}_2$ and T_4 and replacing $\mathbf{D}F$ by its expression $\bar{b} - \frac{1}{2}(a\nabla \ln(p_\infty) + \nabla.a)$ in our notations, one obtains that $\tilde{R}_2 + T_3 + T_4$ is equal to

$$\int_{\mathbb{R}^d} \frac{U''_\delta(\rho)Q_1}{2} p_\infty - \int_{\mathbb{R}^d} \left[\frac{U''_\delta(\rho)}{4}\left(\partial_i\rho\partial_j a_{ik}\partial_k\rho + 2\partial_{ij}\rho a_{ik}\partial_k\rho\right)\right] \times \partial_l(a_{lj}p_\infty) - \int_{\mathbb{R}^d} \frac{U''_\delta(\rho)}{2}\left[a_{lj}\partial_{ijl}\rho\partial_k\rho a_{ki}\right] p_\infty$$

up to time reversal. The first term corresponds to J_2. Integrating by parts the second term to get rid of the derivative w.r.t. the l-th coordinate in the second factor, one checks that its sum with the last one is equal to J_3. Hence, up to time reversal, we have $J_1+J_2+J_3 = (\tilde{R}_1+T_3)+(\tilde{R}_2+T_4)$ which is the expression for the dissipation of entropy dissipation computed in [2] p. 160.

In order to recover the Bakry Emery criterion in [2], we rewrite $Q_1 + Q_2 = K_1(\rho) + K_2(\rho)$ where

$$K_1(\rho) := \bar{b}_m\partial_m a_{ll'}\partial_l\rho\partial_{l'}\rho - 2a_{ll'}\partial_{l'}\rho\partial_k\rho\partial_l\bar{b}_k + \frac{1}{2}a_{kk'}\partial_l\rho\partial_{l'}\rho\partial_{kk'}a_{ll'}$$

and

$$K_2(\rho) := a_{kk'}\partial_{kl}\rho\, a_{ll'}\partial_{k'l'}\rho + 2a_{kk'}\partial_{kl}\rho\partial_{l'}\rho\partial_{k'}a_{ll'} - a_{k'l'}\partial_{l'}\rho\partial_{kl}\rho\partial_{k'}a_{kl}.$$

When a is non singular, introducing $G_{jk}(\rho) = \partial_{l'}\rho a_{k'l'}\partial_{k'}a_{jk}$ and $H_{lj}(\rho) = \partial_j a_{ll'}\partial_{l'}\rho$ we can write

$$\begin{aligned}
K_2(\rho) =& tr\left[(a\nabla^2\rho)^2 + 2H(\rho)\, a\nabla^2\rho - G(\rho)\nabla^2\rho\right] \\
=& tr\left[(a\nabla^2\rho)^2 + H(\rho)a\nabla^2\rho + aH(\rho)^*\nabla^2\rho - G(\rho)a^{-1}a\nabla^2\rho\right] \\
=& tr\Big[(a\nabla^2\rho)^2 + \frac{1}{2}(H(\rho)a\nabla^2\rho + aH(\rho)^*\nabla^2\rho - G(\rho)a^{-1}a\nabla^2\rho) \\
&+ \frac{1}{2}(a\nabla^2\rho H(\rho) + a\nabla^2\rho aH(\rho)^*a^{-1} - a\nabla^2\rho G(\rho)a^{-1})\Big]
\end{aligned}$$

where we have used the cyclicity of the trace and its invariance by transposition. Following [2], we complete the trace of a squared sum of matrices to get

$$K_2(\rho) = tr\left[a\nabla^2\rho + \frac{1}{2}(H(\rho) + aH(\rho)^*a^{-1} - G(\rho)a^{-1})\right]^2 - \frac{1}{4}tr\left[H(\rho) + aH(\rho)^*a^{-1} - G(\rho)a^{-1}\right]^2.$$

The finite variation part on the right hand side of the first line in (C.1) therefore rewrites

$$\begin{aligned}
&\frac{U''_\delta(\rho)}{2}(K_1(\rho) - \frac{1}{4}tr\left[H(\rho) + aH(\rho)^*a^{-1} - G(\rho)a^{-1}\right]^2) \\
&+ \frac{U''_\delta(\rho)}{2}tr\left[a\nabla^2\rho + \frac{1}{2}(H(\rho) + aH(\rho)^*a^{-1} - G(\rho)a^{-1})\right]^2 dt \\
&+ \frac{U^{(3)}_\delta(\rho)}{2}\left(\nabla^*\rho\; a\frac{\partial(\nabla\rho)}{\partial x}a\nabla\rho + \nabla^*\rho\; a\frac{\partial(a\nabla\rho)}{\partial x}^*\nabla\rho\right) dt \\
&+ \frac{U^{(4)}_\delta(\rho)}{4}\left|\nabla^*\rho a\nabla\rho\right|^2 dt.
\end{aligned} \tag{C.2}$$

The sum of the second, third and fourth lines correspond to the matrix product $\mathbf{XY}$ in [2] and is shown to be nonnegative in p. 164 therein. We can then check that for a smooth function $v : \mathbb{R}^d \to \mathbb{R}$, the term $\frac{1}{2}(K_1(v) - \frac{1}{4}tr\left[H(v) + aH(v)^*a^{-1} - G(v)a^{-1}\right]^2)$ is twice the expression

on the l.h.s. of the inequality (2.13) in p. 158 of [2] (with ∇v corresponding to their vector field "U"). Consequently, their Bakry Emery criterion (2.13) corresponds, in our notation, to imposing the condition

$\exists \lambda > 0$ such that for all smooth function $v : \mathbb{R}^d \to \mathbb{R}$ and all $x \in \mathbb{R}^d$:

$$\frac{1}{2}(K_1(v) - \frac{1}{4} tr \left[H(v) + aH(v)^* a^{-1} - G(v)a^{-1}\right]^2)(x) \geq \lambda \nabla v^* a \nabla v(x),$$

which implies exponential convergence at rate 2λ of the $U-$Fisher information and the $U-$relative entropy.

We may combine this criterion with ours by introducing some C^1 function $\alpha : \mathbb{R}^d \to [0,1]$ and writing the finite variation part on the right hand side of the first line in (C.1) as $(1-\alpha)$ multiplied by the expression (C.2) plus $\frac{1}{2}\alpha$ multiplied by the finite variation part in the right hand side of (B.8). Because of the integration by parts performed in the proof of Theorem 3.4, the mixed criterion involves the derivatives of α. Let

$$\Theta^{\alpha}_{ll'} := \alpha \Theta_{ll'} - \frac{1}{2}\partial_k \alpha \left(\left[\sigma_{l'i} a_{mk} - \sigma_{ki} a_{ml'}\right] \partial_m \sigma_{li} + \left[\sigma_{li} a_{mk} - \sigma_{ki} a_{ml}\right] \partial_m \sigma_{l'i} \right).$$

This ultimate mixed criterion writes

$\exists \lambda > 0$ such that for all smooth function $v : \mathbb{R}^d \to \mathbb{R}$ and all $x \in \mathbb{R}^d$:

$$\nabla v^* \Theta^{\alpha} \nabla v(x) + (1-\alpha(x)) \left(\frac{1}{2}(K_1(v) - \frac{1}{4} tr \left[H(v) + aH(v)^* a^{-1} - G(v)a^{-1}\right]^2)(x) \right) \geq \lambda \nabla v^* a \nabla v(x)$$

and also implies exponential convergence at rate 2λ of the $U-$Fisher information and the $U-$relative entropy.

References

[1] Arnold, A., Markowich, P., Toscani, G., Unterreiter, A., : *On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations*, Comm. Partial Differential Equations 26 (2001), no. 1-2, 43–100.

[2] Arnold, A., Carlen, E. and Ju, Q. : *Large-time behavior of non-symmetric Fokker-Planck type equations*, Communications on Stochastic Analysis, 2(1), 153–175, 2008.

[3] Bakry, D. and Emery, M. : *Diffusions hypercontractives.* Séminaire de probabilités 19 (1985), 177–206.

[4] Bouleau, N. and Hirsch, F. :*Dirichlet Forms and Analysis on Wiener Space.* de Gruyter Studies in Mathematics 14, 1991.

[5] Cattiaux, P. :*A pathwise approach of some classical inequalities.* Potential Anal. 20 (2004), no. 4, 361–394

[6] Cattiaux, P. and Léonard, C.:*Minimization of the Kullback information of diffusion processes*, Ann. Inst. H. Poincaré Probab. Statist. 30 (1994), no. 1, 83–132.

[7] Chafaï, D. : *Entropies, convexity, and functional inequalities: on Φ-entropies and Φ-Sobolev inequalities*, J. Math. Kyoto Univ. 44 (2004), no. 2, 325–363.

[8] Föllmer, H.: *Time reversal on Wiener space* Stochastic processes—mathematics and physics (Bielefeld, 1984), 119–129, Lecture Notes in Math., 1158, Springer, Berlin, 1986.

[9] Föllmer, H. and Wakolbinger, A. : *Time reversal of infinite-dimensional diffusions.* Stochastic Process. Appl. 22 (1986), no. 1, 59–77

[10] Friedman, A .: *Partial differential equations of parabolic type.* Prentice-Hall, Inc., Englewood Cliffs, N.J. 1964

[11] Haussmann, U. G.; Pardoux, É.: *Time reversal of diffusions* Ann. Probab. 14 (1986), no. 4, 1188–1205.

[12] Holley, R. A. and Stroock, D. W. :*Logarithmic Sobolev inequalities and stochastic Ising models.*, J. Statist. Phys. 46 (1987), no. 5-6, 1159–1194.

[13] Hwang, C.-R. ; Hwang-Ma, S.-Y. ; Sheu, S.-J.: *Accelerating diffusions.* Ann. Appl. Probab. 15 (2005), no. 2, 1433–1444

[14] Jacod, J.: *Calcul stochastique et problèmes de martingales.* Lecture Notes in Mathematics, 714. Springer, Berlin, 1979.

[15] Kunita, H. : *Stochastic flows of diffeomorphisms and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, 24. Cambridge University Press, Cambridge, 1990.

[16] Kusuoka, S.; Stroock, D. : *Applications of the Malliavin calculus. II.* J. Fac. Sci. Univ. Tokyo Sect. IA Math. 32 (1985), no. 1, 1–76.

[17] Millet, A.; Nualart, D. and Sanz, M. *Integration by parts and time reversal for diffusion processes*, Ann. Probab. 17 (1989), no. 1, 208–238.

[18] Pardoux, E.: *Grossissement d'une filtration et retournement du temps d'une diffusion*, Séminaire de Probabilités, XX, 1984/85, 48–55, Lecture Notes in Math., 1204, Springer, Berlin, 1986.

[19] Revuz, D.; Yor, M. : *Continuous martingales and Brownian motion.* Third edition. Grundlehren der Mathematischen Wissenschaften, 293. Springer-Verlag, Berlin, 1999.