# Towards a Quantitative Averaging Principle for Stochastic Differential Equations

Bob Pepin *

March 12, 2018

**Abstract**

This work explores the use of a forward-backward martingale method together with a decoupling argument and entropic estimates between the conditional and averaged measures to prove a strong averaging principle for stochastic differential equations with order of convergence 1/2. We obtain explicit expressions for all the constants involved. At the price of some extra assumptions on the time marginals and an exponential bound in time, we loosen the usual boundedness and Lipschitz assumptions. We conclude with an application of our result to Temperature-Accelerated Molecular Dynamics.

## 1 Introduction and notation

### 1.1 Motivation and main result

We are interested in stochastic differential equations of the form

$$\begin{aligned} dX_t &= \varepsilon^{-1} b_X(X_t, Y_t)dt + \varepsilon^{-1/2}\sigma_X(X_t, Y_t)dB_t^X, \quad X_0 = x_0, \\ dY_t &= b_Y(X_t, Y_t)dt + \sigma_Y(Y_t)dB_t^Y, \quad Y_0 = y_0 \end{aligned}$$

for some $\varepsilon > 0$ and $x_0 \in \mathbb{R}^n$, $y_0 \in \mathbb{R}^m$. The precise assumptions on the coefficients are stated in Assumption 1 and they essentially amount to $b_X$ being one-sided Lipschitz outside a compact set, $b_Y$ being differentiable with bouded derivative, $\sigma_X$ being bounded and the process being elliptic.

It is well known (see for example [FW12]) that when all the coefficients and their first derivatives are bounded, $Y$ (which depends on $\varepsilon$) can be approximated by a process $\bar{Y}$ on $\mathbb{R}^m$ in the sense that for all $T > 0$ fixed

$$\mathbb{P}\left(\sup_{0\le t\le T} |Y_t - \bar{Y}_t| > \varepsilon\right) \to 0 \text{ as } \varepsilon \to 0.$$

*bob.pepin@uni.lu

The process $\bar{Y}$ solves the SDE

$$d\bar{Y}_t = \bar{b}_Y(\bar{Y}_t)dt + \sigma_Y(\bar{Y}_t)dB_t^Y, \quad \bar{Y}_0 = y_0$$

with

$$\bar{b}(y) = \int_{\mathbb{R}^n} b_Y(x,y)\mu^y(dx).$$

Here $(\mu^y)_{y\in\mathbb{R}^m}$ is a family of measures on $\mathbb{R}^n$ such that for each $y$, $\mu^y$ is the unique stationary measure of $X^y$ with

$$dX_t^y = b_X(X_t^y, y)dt + \sigma_X(X_t^y, y)dB_t^X.$$

The work [Liu10] replaces the boundedness assumption on $b_X$ and $\sigma_X$ by a dissipativity condition and shows the following rate of convergence of the time marginals:

$$\sup_{0\le t\le T} \mathbb{E}|Y_t - \bar{Y}_t| \le C\varepsilon^{1/2}$$

for some constant $C$ independent of $\varepsilon$.

In [LLO16] the author relax the growth conditions on the coefficients of the SDE and show that when $(X_t, Y_t)$ is a reversible diffusion process with stationary measure $\mu = e^{-V(x,y)}dxdy$ such that for each $y$, a Poincaré inequality holds for $e^{-V(x,y)}dx$, then there exists a constant $C$ independent of $\varepsilon$ such that

$$\mathbb{E} \sup_{0\le t\le T} |Y_t - \bar{Y}_t| \le C\varepsilon^{1/2}.$$

The present work extends the approach from [LLO16] to the non-stationary case and drops the boundedness assumption on $b_Y$, $\sigma_Y$ commonly found in the averaging literature. The general setting and notation will be outlined in Section 1.2. Section 2 presents a forward-backward martingale argument under the assumption of a Poincaré inequality for the regular conditional probability density $\rho_t^y$ of $X_t$ given $Y_t = y$. By dropping the stationarity assumption, we have to deal with the fact that $\rho_t^y$ is no longer equal to $\mu^y$ defined above. This is done in Section 3 by developing the relative entropy between $\rho_t^y$ and $\mu^y$ along the trajectories of $Y$. Dropping the boundedness assumption on $b_Y$ forces us to consider the mutual interaction between $X_t$ and $Y_t$. In Section 4 we address this problem when the timescales of $X$ and $Y$ are sufficiently separated. The main theorem is proven in Section 5. Section 6 applies the theorem to a particular class of SDEs to obtain sufficient conditions such that for any $T > 0$ and $\varepsilon$ sufficiently small

$$\mathbb{E} \sup_{0\le t\le T} \left|Y_t - \bar{Y}_t\right| \le C\varepsilon^{1/2}$$

where $C$ will be explicitly given in terms of the coefficients of the SDE and the Poincaré constant for $\rho_t^y$.

### 1.2 Setting and notation

The results in sections 2 to 5 will be stated in the setting of an SDE on $\mathcal{X}\times\mathcal{Y}=\mathbb{R}^n\times\mathbb{R}^m$ of the form

$$\begin{aligned} dX_t &= b_X(X_t,Y_t)dt+\sigma_X(X_t,Y_t)dB_t^X,\ X_0=x\\ dY_t &= b_Y(X_t,Y_t)dt+\sigma_Y(X_t,Y_t)dB_t^Y,\ Y_0=y \end{aligned}$$

where $x\in\mathcal{X}=\mathbb{R}^n$, $y\in\mathcal{Y}=\mathbb{R}^m$, $B^X,B^Y$ are independent standard Brownian motions on $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and $b_X=(b_X^i)_{1\le i\le n}$, $b_Y=(b_Y^i)_{1\le i\le m}$, $\sigma_X$ and $\sigma_Y$ are continuous mappings from $\mathcal{X}\times\mathcal{Y}$ to $\mathcal{X}$, $\mathcal{Y}$, $\mathbb{R}^{n\times n}$ and $\mathbb{R}^{m\times m}$ respectively.

The matrices $A_X=\left(a_X^{ij}(x,y)\right)_{i,j\le n}$ and $A_Y=\left(a_Y^{ij}(x,y)\right)_{i,j\le m}$ are defined by

$$A_X(x,y)=\tfrac{1}{2}\sigma_X(x,y)\sigma_X(x,y)^T,\qquad A_Y(x,y)=\tfrac{1}{2}\sigma_Y(x,y)\sigma_Y(x,y)^T$$

and the infinitesimal generator $L$ of $(X,Y)$ has a decomposition $L=L^X+L^Y$ such that

$$\begin{aligned} L^Xf &= \sum_{i=1}^n b_X^i\partial_{x_i}f+\sum_{i,j=1}^n a_X^{ij}\partial^2_{x_ix_j}f,\\ L^Yf &= \sum_{i=1}^m b_Y^i\partial_{x_i}f+\sum_{i,j=1}^m a_Y^{ij}\partial^2_{x_ix_j}f,\\ Lf &= (L^X+L^Y)f. \end{aligned}$$

We will also make use of the square field operators $\Gamma$ and $\Gamma^X$, defined by

$$\begin{aligned} \Gamma(f,g) &= \tfrac{1}{2}(L(fg)-gLf-fLg)=\sum_{i,j=1}^n a_X^{ij}\partial_{x_i}f\partial_{x_j}g+\sum_{i,j=1}^m a_Y^{ij}\partial_{y_i}f\partial_{y_j}g,\\ \Gamma^X(f,g) &= \tfrac{1}{2}(L^X(fg)-gL^Xf-fL^Xg)=\sum_{i,j=1}^n a_X^{ij}\partial_{x_i}f\partial_{x_j}g. \end{aligned}$$

We denote $\rho_t(dx,dy)$ the marginal distribution of $(X,Y)$ at time $t$, i.e. for $\varphi\in C_c^\infty$

$$\mathbb{E}[\varphi(X_t,Y_t)]=\int_{\mathcal{X}\times\mathcal{Y}}\varphi(x,y)\rho_t(dx,dy)$$

and we let $\rho_t^y(dx)$ be the regular conditional probability density of $P(X_t\in dx|Y_t=y)$.

If a measure $\mu(dx,dy)$ is absolutely continuous with respect to Lebesgue measure we will make a slight abuse of notation and denote $\mu(x,y)$ its density.

We will also make use of a family of auxiliary processes $(X^y)_{y\in\mathcal{Y}}$ defined by

$$dX_t^y=b_X(X_t,y)dt+\sigma_X(X_t,y)dB_t^X,\ X_0^y=x$$

which we assume to be uniformly ergodic and we denote $\mu^y$ the unique stationary invariant measure of $X^y$.

We will furthermore use another auxiliary process $\tilde{X}$ solution to

$$d\tilde{X}_t = b_X(\tilde{X}_t, Y_t)dt + \sigma_X(\tilde{X}_t, Y_t)d\tilde{B}_t^X,\ \tilde{X}_0 = x$$

where $\tilde{B}^X$ is an $n$-dimensional Brownian motion independent of $B^X$ and $B^Y$ and we denote $\tilde{\rho}_t^y$ the regular conditional probability density of $P(\tilde{X}_t \in dx | Y_t = y)$.

For the section on decoupling and the main theorem we need in addition to a separation of timescales the following regularity conditions on the coefficients of $(X, Y)$:

**Assumption 1.** *Regularity of the coefficients:*

- *$b_X$ verifies a one-sided Lipschitz condition with constant $\kappa_X$ and perturbation $\alpha$:*

  $(x_1 - x_2)^T(b_X(x_1, y) - b_X(x_2, y)) \leq -\kappa_X |x_1 - x_2|^2 + \alpha$ *for all* $x_1, x_2 \in \mathcal{X}, y \in \mathcal{Y}$

- *$b_Y$ has a bounded first derivative in $x$:*

$$\kappa_Y{}^2 := \frac{1}{m} \sum_{i=1}^{m} \sup_{x,y} |\nabla_x b_Y^i(x, y)|^2 < \infty$$

- *$A_X$ is nondegenerate uniformly with respect to $(x, y)$, i.e. there exist two constants $0 < \lambda_X \leq \Lambda_X < \infty$ such that the following matrix inequalities hold (in the sense of nonnegative definiteness):*

$$\lambda_X Id \leq A_X(x, y) \leq \Lambda_X Id$$

- *$\sigma_Y$ is invertible and $A_Y$ is uniformly elliptic with respect to $(x, y)$, i.e. there exists a constant $\lambda_Y > 0$ such that the following matrix inequality holds (in the sense of nonnegative definiteness):*

$$\lambda_Y Id \leq A_Y(x, y)$$

**Assumption 2.** *Regularity of the time marginals:*

- *There exists $M_0$ such that for $|x|^2 + |y|^2 > M_0$, $r > 0$, $\alpha > 0$*

$$\nabla_x \log \rho_t(x, y)^T x + \nabla_y \log \rho_t(x, y)^T y \leq -r(|x|^2 + |y|^2)^{\alpha/2}.$$

- *The regular conditional probability densities $\tilde{\rho}_t^y$ of $P(\tilde{X}_t \in dx | Y_t = y)$ satisfy Poincaré inequalities with constants $c_P(y)$ independent of $\varepsilon$:*

$$\int (f - \tilde{\rho}_t^y(f))^2 d\tilde{\rho}_t^y \leq c_P(y) \int |\sigma_X \nabla_x f|^2 d\tilde{\rho}_t^y.$$

In order to characterise the separation of timescales, we introduce a parameter $\gamma$ defined by

$$\gamma = \frac{\kappa_X{}^2 \lambda_Y}{\Lambda_X \kappa_Y{}^2}.$$

## 2 Approximation by conditional expectations

We will start with a Lemma for a form of the Lyons-Meyer-Zheng forward-backward martingale decomposition.

**Lemma 3** (Forward-backward martingale decomposition)**.** *For a diffusion process $\xi_t$ with generator $L_t$ and square field operator $\Gamma_t$ we have for $f(s,\cdot) \in \mathcal{D}(L_s + \tilde{L}_{T-s})$ and $1 \le p \le 2$*

$$\mathbb{E} \sup_{0\le t\le T} \left| \int_0^t -(L_s + \tilde{L}_{T-s}) f(s,\xi_s) ds \right|^p \le 3^{p-1}(2C_p + 1) \left( \mathbb{E} \int_0^T 2\Gamma_t(f)(\xi_t) dt \right)^{p/2}$$

*where $\tilde{L}_s$ is the generator of the time-reversed process $\tilde{\xi}_t = \xi_{T-t}$ and $C_p$ is the constant in the upper bound of the Burkholder-Davis-Gundy inequality for $L^p$.*

*Proof.* First, suppose that $f(t,x)$ is once differentiable in $t$ and twice differentiable in $x$ so that we can apply the Itô formula.

We express $f(t,\xi_t) - f(0,\xi_0)$ in two different ways, using the fact that $\xi_t = \tilde{\xi}_{T-t}$:

$$\begin{aligned} f(t,\xi_t) - f(0,\xi_0) &= \int_0^t (\partial_s + L_s) f(s,\xi_s) ds + M_t && (1) \\ f(0,\xi_0) - f(t,\xi_t) &= (f(0,\tilde{\xi}_T) - f(T,\tilde{\xi}_0)) - (f(t,\tilde{\xi}_{T-t}) - f(T,\tilde{\xi}_0)) \\ &= \int_{T-t}^T (-\partial_s + \tilde{L}_s) f(T-s,\tilde{\xi}_s) ds + \tilde{M}_T - \tilde{M}_{T-t} \\ &= \int_0^t (-\partial_s + \tilde{L}_{T-s}) f(s,\tilde{\xi}_{T-s}) ds + \tilde{M}_T - \tilde{M}_{T-t} \\ &= \int_0^t (-\partial_s + \tilde{L}_{T-s}) f(s,\xi_s) ds + \tilde{M}_T - \tilde{M}_{T-t} && (2) \end{aligned}$$

where $M$ and $\tilde{M}$ are martingales with

$$\begin{aligned} \langle M \rangle_T &= \int_0^T 2\Gamma_s(f)(s,\xi_s) ds, \\ \langle \tilde{M} \rangle_T &= \int_0^T 2\Gamma_{T-s}(f)(T-s,\tilde{\xi}_s) ds = \int_0^T 2\Gamma_s(f)(s,\xi_s) ds = \langle M \rangle_T. \end{aligned}$$

Summing (1) and (2), we get

$$\int_0^t -(L_s + \tilde{L}_{T-s}) f(s,\xi_s) = M_t + \tilde{M}_T - \tilde{M}_{T-t}.$$

We have by the Burkholder-Davis-Gundy $L^p$-inequality that

$$\mathbb{E} \sup_{0\le t\le T} |M_t|^p \le C_p \mathbb{E}[\langle M \rangle_T^{p/2}]$$

$$\mathbb{E} \sup_{0\le t\le T} |\tilde{M}_{T-t}|^p = \mathbb{E} \sup_{0\le t\le T} |\tilde{M}_t|^p \le C_p \mathbb{E}[\langle \tilde{M} \rangle_T^{p/2}] = C_p \mathbb{E}[\langle M \rangle_T^{p/2}]$$

so that

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t -(L_s+\tilde{L}_{T-s}f)(s,\xi_s)\right|^p ds\right] &= \mathbb{E}\sup_{0\le t\le T}\left|M_t+\tilde{M}_T-\tilde{M}_{T-t}\right|^p\\ &\le 3^{p-1}\left(\mathbb{E}\sup_{0\le t\le T}|M_t|^p+\mathbb{E}|\tilde{M}_T|^p+\mathbb{E}\sup_{0\le t\le T}|\tilde{M}_{T-t}|^p\right)\\ &\le 3^{p-1}(2C_p+1)(\mathbb{E}\langle M\rangle_T)^{p/2}\\ &\le 3^{p-1}(2C_p+1)\left(\mathbb{E}\int_0^T 2\Gamma_t(f)(t,\xi_t)dt\right)^{p/2}\end{aligned}$$

For a general $f(t,x)$, $C^2$ in $x$ and locally integrable in $t$, we approximate first in space by stopping $\xi_t$ and then in time by mollifying $f(\cdot,x)$.

For $R>0,\varepsilon>0$ and a function $f(t,x)$ we will use the notation

$$\begin{aligned}(f)^R(t,x) &= f(t,x\frac{|x|\wedge R}{|x|}),\\ (f)_\varepsilon(t,x) &= \int_{-\infty}^{+\infty} f(s,x)\phi_\varepsilon(t-s)ds\end{aligned}$$

where $\phi_\varepsilon$ is a mollifier. In particular, $(f)^R(t,\cdot)$ is bounded and $(f)_\varepsilon(\cdot,x)$ is differentiable.

Let $K_t=L_t+\tilde{L}_{T-t}$. $K_t$ is a second order partial differential operator and so can be written as

$$K_tf(t,x)=\sum b^i(t,x)\partial_{x_i}f(t,x)+\sum a^{ij}(t,x)\partial^2_{x_ix_j}f(t,x)$$

for some functions $b^i$ and $a^{ij}$.

Define the stopping times $\tau_R=\inf\{t>0:|\xi_t|\ge R\}$. Then

$$\begin{aligned}\mathbb{E}\sup_{0\le t\le T}\left|\int_0^{t\wedge\tau_R} K_s(f)_\varepsilon(s,\xi_s)ds\right|^p &= \mathbb{E}\sup_{0\le t\le\tau_R\wedge T}\left|\int_0^t (K_s(f)_\varepsilon)^R(s,\xi_s)ds\right|^p\\ &\le \mathbb{E}\sup_{0\le t\le T}\left|\int_0^t (K_s(f)_\varepsilon)^R(s,\xi_s)ds\right|^p\\ &\le 3^{p-1}(2C_p+1)\left(\mathbb{E}\int_0^T 2(\Gamma_t((f)_\varepsilon)^R(t,\xi_t)dt\right)^{p/2}. \quad (1)\end{aligned}$$

By differentiating inside the integral for $(f)_\varepsilon$ we get

$$\begin{aligned}\int_0^{t\wedge\tau_R} K_s(f-(f)_\varepsilon)(s,\xi_s)ds \le& \sup_{0\le t\le T,|x|\le R}|b^i(t,x)|\int_0^T\sup_{|x|\le R}|(\partial_{x_i}f-(\partial_{x_i}f)_\varepsilon)(s,x)|ds\\ &+\sup_{0\le t\le T,|x|\le R}|a^{ij}(t,x)|\int_0^T\sup_{|x|\le R}|(\partial^2_{x_ix_j}f-(\partial^2_{x_ix_j}f)_\varepsilon)(s,x)|ds.\end{aligned}$$

As $\varepsilon \to 0$, $(g)_\varepsilon \to g$ in $L^1([0,T], L^\infty(B_R))$ and the integrals on the right hand side go to 0. We now let first $\varepsilon \to 0$ with dominated convergence and then $R \to \infty$ with monotone convergence to get

$$\mathbb{E} \sup_{0\le t\le T} \Big| \int_0^t K_s f(s,\xi_s) ds \Big|^p = \lim_{R\to\infty} \lim_{\varepsilon\to 0} \mathbb{E} \sup_{0\le t\le T} \Big| \int_0^{t\wedge\tau_R} K_s (f)_\varepsilon (s,\xi_s) ds \Big|^p$$

For the right hand side of (1), note that

$$(\Gamma_t(f) - \Gamma_t((f)_\varepsilon))^R = \Gamma_t(f-(f)_\varepsilon, f+(f)_\varepsilon)^R = (a^{ij})^R (\partial_{x_i} f - (\partial_{x_i} f)_\varepsilon)^R (\partial_{x_j} f + (\partial_{x_j} f)_\varepsilon)^R$$

so that

$$\int_0^T |(\Gamma_t(f) - \Gamma_t((f)_\varepsilon))^R| \le \sup_{0\le t\le T, |x|\le R} a^{ij}(t,x)(\partial_{x_j} f + (\partial_{x_j} f)_\varepsilon) \int_0^T \sup_{|x|\le R} |\partial_{x_i} f - (\partial_{x_i} f)_\varepsilon| dt.$$

Now the convergence follows again by first letting $\varepsilon \to 0$ with dominated convergence and then $R \to \infty$ with monotone convergence. $\square$

**Lemma 4.** *Let $L$ and $\hat{L}$ be generators of diffusion processes with common invariant measure $\mu$ and square field operators $\Gamma$ and $\hat{\Gamma}$ respectively. Let $f, g$ be a pair of functions such that*

$$Lf = \hat{L}g \text{ and } \int \hat{\Gamma}(f) d\mu \le \int \Gamma(f) d\mu.$$

*Then*

$$\int \Gamma(f) d\mu \le \int \hat{\Gamma}(g) d\mu.$$

*Proof.*

$$\begin{aligned}
\int \Gamma(f) d\mu = \int f L f d\mu = \int f \hat{L} g d\mu &= \int \hat{\Gamma}(f,g) d\mu \\
&\le \Big( \int \hat{\Gamma}(f) d\mu \Big)^{1/2} \Big( \int \hat{\Gamma}(g) d\mu \Big)^{1/2} \\
&\le \Big( \int \Gamma(f) d\mu \Big)^{1/2} \Big( \int \hat{\Gamma}(g) d\mu \Big)^{1/2}.
\end{aligned}$$

The result follows by dividing both sides by $\left(\int \Gamma(f) d\mu\right)^{1/2}$. $\square$

**Lemma 5.** *Consider a generator $L$ with invariant measure $\mu$ and associated square field operator $\Gamma$. Assume that the following Poincaré inequality holds:*

$$\int (\varphi - \mu(\varphi))^2 d\mu \le c_P \int \Gamma(\varphi) d\mu.$$

*Then for any sufficiently nice $f$*

$$\int \Gamma(f) d\mu \le c_P \int (-Lf)^2 d\mu \le {c_P}^2 \int \Gamma(-Lf) d\mu$$

*Proof.* Since both $\Gamma$ and $L$ are differential operators, we can assume that $\mu(f) = 0$. Now,

$$\left(\int \Gamma(f)d\mu\right)^2 = \left(-\int fLfd\mu\right)^2 \leq \int f^2 d\mu \int (-Lf)^2 d\mu \leq c_P \int \Gamma(f)d\mu \int (-Lf)^2 d\mu$$

and the first inequality follows after dividing both sides by $\int \Gamma(f)d\mu$. For the second inequality, we apply the Poincaré inequality again with $\varphi = (-Lf)$. $\square$

**Proposition 6.** *In the general setting of section 1.2 with Assumption 2 let $\nu_t^\eta(dx)$ be the regular conditional probability density of $\mathbb{P}(X_t \in dx|\phi(Y_t) = \eta)$ for a measurable function $\phi\colon \mathcal{Y} \to \mathbb{R}^l$. If $\nu_t^\eta$ satisfies a Poincaré inequality with constant $c_P(\eta)$ independent of $t$ with respect to $\Gamma^X$ then for any function $f_t(x, y)$ with at most polynomial growth in $x$ and $y$ such that $f_t(\cdot) \in C^2(\mathcal{X} \times \mathcal{Y})$, $\int_{\mathcal{X}} f_t(x, y)\nu_t^{\phi(y)}(dx) = 0$ and $1 \leq p \leq 2$*

$$\mathbb{E} \sup_{0\leq t\leq T}\left|\int_0^t f_s(X_s, Y_s)ds\right|^p \;\leq\; 3^{p-1}2^{-p/2}(2C_p + 1)\left(\mathbb{E}\int_0^T c_P(\phi(Y_t))f_t^2(X_t, Y_t)dt\right)^{p/2}$$

*where $C_p$ is the constant in the upper bound of the Burkholder-Davis-Gundy inequality for $L^p$.*

*Proof of Proposition 6.* The generator of the time-reversed process $(X, Y)_{T-t}$ is [HP86]

$$\begin{aligned}\tilde{L}_t\varphi = &-\sum_{i=1}^n b_X^i\partial_{x_i}\varphi - \sum_{i=1}^m b_Y^i\partial_{y_i}\varphi + \sum_{i,j=1}^n a_X^{ij}\partial^2_{x_ix_j}\varphi + \sum_{i,j=1}^m a_Y^{ij}\partial^2_{y_iy_j}\varphi \\ &+ \frac{1}{p_{T-t}}\sum_{i,j=1}^n \partial_{x_j}(2a_X^{ij}p_{T-t})\partial_{x_i}\varphi + \frac{1}{p_{T-t}}\sum_{i,j=1}^m \partial_{y_j}(2a_Y^{ij}p_{T-t})\partial_{y_i}\varphi\end{aligned}$$

so that the symmetrized generator is

$$\begin{aligned}K_t\varphi &:= \frac{(L + \tilde{L}_{T-t})\varphi}{2} \\ &= \frac{1}{p_t}\sum_{i,j=1}^n \partial_{x_j}(a_X^{ij}p_t)\partial_{x_i}\varphi + \sum_{i,j=1}^n a_X^{ij}\partial^2_{x_ix_j}\varphi + \frac{1}{p_t}\sum_{i,j=1}^m \partial_{y_j}(a_Y^{ij}p_t)\partial_{y_i}\varphi + \sum_{i,j=1}^m a_Y^{ij}\partial^2_{y_iy_j}\varphi \\ &= \sum_{i,j=1}^n \frac{1}{p_t}\partial_{x_i}(p_ta_X^{ij}\partial_{x_j}\varphi) + \sum_{i,j=1}^m \frac{1}{p_t}\partial_{y_i}(p_ta_Y^{ij}\partial_{y_j}\varphi).\end{aligned}$$

For fixed $\tau \geq 0$, we see from the expression for $K$ that $p_\tau(dx, dy)$ is an invariant measure for $K_\tau$ (use integration by parts).

By the properties of conditional expectation $\int f_\tau dp_\tau = 0$. From Assumption 2 and Theorem 1 in [PV01] it follows that for each $\tau$ there exists a unique solution $F_\tau \in C^2(\mathcal{X} \times \mathcal{Y})$ to the Poisson Problem $K_\tau F_\tau = f_\tau$.

We can now apply the forward-backward martingale decomposition via Lemma 3 to obtain

$$
\begin{aligned}
\mathbb{E}\sup_{0\le t\le T}\Big|\int_0^t f_s(X_s,Y_s)ds\Big|^p &= \mathbb{E}\sup_{0\le t\le T}\Big|\int_0^t K_sF_s(X_s,Y_s)ds\Big|^p \\
&= 2^{-p}\mathbb{E}\sup_{0\le t\le T}\Big|\int_0^t (L+\tilde{L}_{T-s})F_s(X_s,Y_s)ds\Big|^p \\
&\le 2^{-p}3^{p-1}(2C_p+1)\bigg(\mathbb{E}\int_0^T 2\Gamma(F_s)(X_s,Y_s)ds\bigg)^{p/2}.
\end{aligned}
$$

Now, we want to pass from $\Gamma$ to $\Gamma^X$ in order to use our Poincaré inequality for $\nu_t^\eta$.

For $\varphi\in C^2(\mathcal{X})$ and $y\in\mathcal{Y},\tau\ge 0$ fixed let $\hat{K}^{\tau,y}\varphi$ be the the reversible generator associated to $\Gamma^X(\varphi)(\cdot,y)$ and $\nu_\tau^{\phi(y)}$.

Since $\nu_\tau^{\phi(y)}$ satisfies a Poincaré inequality and $\int f_\tau(x,y)\nu_\tau^{\phi(y)}(dx)=0$ by assumption,

$$
\hat{K}_\tau\hat{F}^{\tau,y}(x)=f_\tau(x,y)
$$

has a unique solution $\hat{F}^{\tau,y}(x)$.

If we set $\hat{K}_\tau\varphi(x,y)=(\hat{K}^{\tau,y}\varphi(\cdot,y))(x)$ and $\hat{F}_\tau(x,y)=\hat{F}^{\tau,y}(x)$ then

$$
\begin{aligned}
\int_{\mathcal{X}\times\mathcal{Y}}\hat{K}_\tau\varphi(x,y)p_t(dx,dy) = \int_{\mathcal{Y}}\int_{\mathcal{X}}(\hat{K}^{\tau,y}\varphi(\cdot,y))(x)\nu_t^{\phi(y)}p_t(\mathcal{X},dy)=0 \text{ and}\\
\hat{K}_\tau\hat{F}_\tau(x,y)=f_\tau(x,y)=K_\tau F_\tau(x,y).
\end{aligned}
$$

By Lemma 4 we get that

$$
\int_{\mathcal{X}\times\mathcal{Y}}\Gamma(F_t)dp_t\le\int_{\mathcal{X}\times\mathcal{Y}}\Gamma^X(\hat{F}_t)dp_t.
$$

Since $\hat{K}\hat{F}_t=f_t$ and $\hat{K}_t$ is the generator associated with $\Gamma^X$ and $\nu_t^{\phi(y)}$, we can use the Poincaré inequality on $\nu_t^{\phi(y)}$ in Lemma 5 to estimate the right hand side by

$$
\begin{aligned}
\int_{\mathcal{X}\times\mathcal{Y}}\Gamma^X(\hat{F}_t)(x,y)p_t(dx,dy) &= \int_{\mathcal{Y}}\int_{\mathcal{X}}\Gamma^X(\hat{F}_t)(x,y)\nu_t^{\phi(y)}(dx)p_t(\mathcal{X},dy)\\
&\le\int_{\mathcal{Y}}c_P(\phi(y))\int_{\mathcal{X}}{f_t}^2(x,y)\nu_t^{\phi(y)}(dx)p_t(\mathcal{X},dy)\\
&=\int_{\mathcal{X}\times\mathcal{Y}}c_P(\phi(y)){f_t}^2(x,y)p_t(dx,dy)
\end{aligned}
$$

which completes the proof.

□

# 3 Distance between conditional and averaged measures

We will first show a general result on the relative entropy between $\rho_t^{Y_t}$ and $\mu^{Y_t}$ by studying the relative entropy along the trajectories of $Y_t$. We are still in the setting of section 1.2.

**Proposition 7.** *Let $f_t(x,y) = \frac{d\rho_t^y}{d\mu^y}(x)$. If $\mu^y$ satisfies a Logarithmic Sobolev inequality with constant $c_L$ uniformly in $y$ with respect to $\Gamma^X$ then for $r \in \mathbb{R}$*

$$\mathbb{E}\,\mathrm{H}(\rho_t^{Y_t}|\mu^{Y_t})e^{rt} \leq \mathbb{E}\,\mathrm{H}(\rho_0^{Y_0}|\mu^{Y_0}) - \left(\frac{2}{c_L} - r\right)\int_0^t \mathbb{E}\,\mathrm{H}(\rho_s^{Y_s}|\mu^{Y_s})e^{rs}ds + \int_0^t \mathbb{E}[L^Y \log f_s(X_s,Y_s)]e^{rs}ds.$$

*Proof.* We have

$$H(\rho_t^y|\mu^y) = \int_{\mathcal{X}} f_t \log f_t \mu^y(dx) = \mathbb{E}[\log f_t(X_t,Y_t)|Y_t = y]$$

so that the quantity we want to estimate is

$$\mathbb{E}H(\rho_t^{Y_t}|\mu^{Y_t}) = \mathbb{E}[\log f_t(X_t,Y_t)].$$

Now by Itô's formula

$$\begin{aligned} de^{rt}\log f_t(X_t,Y_t) &= \left((\partial_t + L)\log f_t(X_t,Y_t) + r\log f_t(X_t,Y_t)\right)e^{rt}dt + dM_t \\ &= \big((\partial_t \log\rho_t^y(x))(X_t,Y_t) + L^X \log f_t(X_t,Y_t) + L^Y \log f_t(X_t,Y_t) + r\log f_t(X_t,Y_t)\big)e^{rt}dt \\ &\quad + dM_t \end{aligned}$$

where $M_t$ is a local martingale.

Since $\rho_t^y dx$ is a probability measure, we have

$$\begin{aligned} \mathbb{E}[\partial_t \log\rho_t^y(x)(X_t,Y_t)|Y_t = y] &= \int_{\mathcal{X}} (\partial_t \log\rho_t^y(x))\rho_t^y(x)dx \\ &= \int_{\mathcal{X}} \partial_t \rho_t^y(x)dx \\ &= \partial_t \int_{\mathcal{X}} \rho_t^y(x)dx = 0. \end{aligned}$$

By the definition of $\mu^y$ as an invariant measure for $X^y$ we have for all $\varphi$ in the domain of $L^X$

$$\int_{\mathcal{X}} L^X \varphi(x,y)d\mu^y = 0. \tag{2}$$

From the Logarithmic Sobolev inequality for $\mu^y$ we get

$$\mathrm{H}(\rho_t^y|\mu^y) \leq \tfrac{1}{2}c_L\,\mathrm{I}(\rho_t^y|\mu^y) = \tfrac{1}{2}c_L \int_{\mathcal{X}} \frac{\Gamma^X(f_t)(x,y)}{f_t(x)}\mu^y(x)dx.$$

Together with the formula $L^X(g\circ f) = g'(f)L^X f + g''(f)\Gamma^X(f)$ this implies

$$\begin{aligned}
\mathbb{E}[L^X \log f_t(X_t,Y_t)|Y_t=y] &= \int_{\mathcal{X}} L^X(\log f_t)(x,y)\rho_t^y(x)dx \\
&= \int_{\mathcal{X}} L^X f_t(x,y)\mu^y(x)dx - \int_{\mathcal{X}} \frac{\Gamma^X(f_t)(x,y)}{f_t(x)}\mu^y(x)dx \\
&= -\,\mathrm{I}(\rho_t^y|\mu^y) \\
&\le -\frac{2}{c_L}\,\mathrm{H}(\rho_t^y|\mu^y).
\end{aligned}$$

By the tower property for conditional expectation and the preceding results, $\mathbb{E}[(\partial_t \log\rho_t^y(x))(X_t,Y_t)] = 0$ and $\mathbb{E}[L^X \log f_t(X_t,Y_t)] \le -\frac{2}{c_L}\mathbb{E}\,\mathrm{H}(\rho_t^{Y_t}|\mu^{Y_t})$ so that

$$\begin{aligned}
\mathbb{E}\,\mathrm{H}(\rho_t^{Y_t}|\mu^{Y_t})e^{rt} &= E[\log f_t(X_t,Y_t)e^{rt}] \\
&\le \mathbb{E}\,\mathrm{H}(\rho_0^{Y_0}|\mu^{Y_0}) - \left(\frac{2}{c_L}-r\right)\int_0^t \mathbb{E}\,\mathrm{H}(\rho_s^{Y_s}|\mu^{Y_s})e^{rs}ds + \int_0^t \mathbb{E}[L^Y \log f_s(X_s,Y_s)]e^{rs}ds.
\end{aligned}$$

□

We now proceed to estimate the term $\mathbb{E}[L^Y \log f_s(X_t,Y_t)]$ in a restricted setting where the coefficients of $L^Y$ are independent of $x$ and $\mu^y$ has a density $\mu^y(x) = Z(y)^{-1}e^{-V(x,y)}$ where $V$ has bounded first and second derivatives in $y$.

**Lemma 8.** *If the coefficients $b_Y^i$ and $a_Y^{ij}$ of $L^Y$ only depend on $y$ then for $f_t(x,y) = \frac{d\rho_t^y}{d\mu^y}(x)$*

$$\int_{\mathcal{X}} L^Y \log f_t d\rho_t^y \le -\int_{\mathcal{X}} L^Y \log\mu^y d\rho_t^y$$

*Proof.* Let $g_t(x,y) = \rho_t^y(x)$. Provided that all the integrals exist, we have

$$\begin{aligned}
\int_{\mathcal{X}} L^Y(\log g_t)(x,y)\rho_t^y(dx) &= \int_{\mathcal{X}} L^Y(\log g_t(x,\cdot))(y)\rho_t^y(dx) \\
&= \int_{\mathcal{X}} L^Y(g_t(x,\cdot))(y)dx - \int_{\mathcal{X}} \frac{\Gamma^Y(g_t(x,\cdot))(y)}{g_t(x,y)}\rho_t^y(dx) \\
&\le \int_{\mathcal{X}} L^Y(g_t(x,\cdot))(y)dx \\
&= L^Y\left(\int_{\mathcal{X}} g_t(x,\cdot)dx\right)(y) = 0
\end{aligned}$$

since $g_t(x,y)dx$ is a probability measure. Now the result follows since

$$L^Y \log f_t = L^Y \log g_t - L^Y \log\mu^y.$$

□

**Lemma 9.** *Consider a probability measure $\mu(dx,dy)$ with density $\mu(x,y)$ on $\mathcal{X}\times\mathcal{Y}$ and let $Z(y)=\int_{\mathcal{X}}\mu(x,y)dx$, $\mu^y(dx)=\mu(dx,y)/Z(y)$. We have the identities*

$$\begin{aligned}
\partial_{y_i}\log Z(y) &= \int_{\mathcal{X}}\partial_{y_i}\log\mu(x,y)\mu^y(dx),\\
\partial^2_{y_iy_j}\log Z(y) &= \int_{\mathcal{X}}\partial^2_{y_iy_j}\log\mu(x,y)\mu^y(dx)+\mathrm{Cov}_{\mu^y}(\partial_{y_i}\log\mu,\partial_{y_j}\log\mu).
\end{aligned}$$

*Proof.* By differentiating under the integral

$$\partial_{y_i}\log Z(y)=\frac{\partial_{y_i}Z(y)}{Z(y)}=\int_{\mathcal{X}}\partial_{y_i}\mu(x,y)\frac{dx}{Z(y)}=\int_{\mathcal{X}}\frac{\partial_{y_i}\mu(x,y)}{\mu(x,y)}\frac{\mu(x,y)dx}{Z(y)}=\int_{\mathcal{X}}\partial_{y_i}\log\mu(x,y)\mu^y(dx)$$

and

$$\begin{aligned}
&\partial^2_{y_iy_j}\log Z(y)\\
&\quad=\partial_{y_i}\int_{\mathcal{X}}\partial_{y_j}\log\mu(x,y)\mu^y(dx)\\
&\quad=\int_{\mathcal{X}}\partial_{y_i}\partial_{y_j}\log\mu(x,y)\mu^y(dx)+\int_{\mathcal{X}}\partial_{y_j}\log\mu(x,y)\frac{\partial_{y_i}\mu(x,y)}{Z(y)}dx-\int_{\mathcal{X}}\partial_{y_j}\log\mu(x,y)\mu(x,y)\frac{\partial_{y_i}Z(y)}{Z(y)^2}dx\\
&\quad=\int_{\mathcal{X}}\partial^2_{y_iy_j}\log\mu(x,y)\mu^y(dx)\\
&\qquad+\int_{\mathcal{X}}\partial_{y_j}\log\mu(x,y)\partial_{y_i}\log\mu(x,y)\mu^y(dx)-\partial_{y_i}\log Z(y)\int_{\mathcal{X}}\partial_{y_j}\log\mu(x,y)\mu^y(dx)\\
&\quad=\int_{\mathcal{X}}\partial^2_{y_iy_j}\log\mu(x,y)\mu^y(dx)+\mathrm{Cov}_{\mu^y}(\partial_{y_i}\log\mu,\partial_{y_j}\log\mu).
\end{aligned}$$

□

**Lemma 10.** *For any Lipschitz function $f$*

$$\left|\int fd\mu^y-\int fd\rho_t^y\right|^2\le\|f\|^2_{\mathrm{Lip}}\Lambda_Xc_LH(\rho_t^y|\mu^y)$$

*uniformly in $y\in\mathcal{Y}$.*

*Proof.* By the Logarithmic Sobolev inequality of $\mu^y$ with respect to $\Gamma^X$ and the uniform boundedness of $A$ we have

$$\mathrm{Ent}_{\mu^y}(f^2)\le 2c_L\int\Gamma^X(f)d\mu^y=2c_L\int(\nabla_xf)^TA(\cdot,y)(\nabla_xf)d\mu^y\le 2c_L\Lambda_X\int|\nabla_xf|^2d\mu^y$$

which says that $\mu^y$ satisfies a Logarithmic Sobolev inequality with respect to the usual square field operator $|\nabla_x|^2$ with constant $c_L\Lambda_X$. By the Otto-Villani theorem, this implies a $T_2$ inequality with the same constant:

$$W_2(\rho_t^y,\mu^y)^2\le c_L\Lambda_XH(\rho_t^y|\mu^y).$$

By the Kantorovich duality formulation of $W_1$ and monotonicity of Kantorovich norms it follows from the preceding $T_2$ inequality that

$$\Big|\sup_{\|f\|_{\mathrm{Lip}}\le 1}\int f d(\rho_t^y-\mu^y)\Big|^2 = W_1(\rho_t^y,\mu^y)^2 \le W_2(\rho_t^y,\mu^y)^2 \le c_L\Lambda_X H(\rho_t^y|\mu^y)$$

from which the result follows. □

**Proposition 11.** *If $b_Y$, $\sigma_Y$ depend only on $y$ and $\mu^y(dx)=Z(y)^{-1}e^{-V(x,y)}dx$ such that $\|\partial_{y_i}V(\cdot,y)\|_{\mathrm{Lip}}<\infty$, $\|\partial^2_{y_iy_j}V(\cdot,y)\|_{\mathrm{Lip}}<\infty$ for all $y$ then*

$$\mathbb{E}L^Y f_t(X_t,Y_t)\le \frac{\Lambda_X c_L}{2}\mathbb{E}\left(\sum_{i=1}^m\|\partial_{y_i}V(\cdot,Y_t)\|^2_{\mathrm{Lip}}+\sum_{i,j=1}^m\|\partial^2_{y_iy_j}V(\cdot,Y_t)\|^2_{\mathrm{Lip}}\right)H(\rho_t^{Y_t}|\mu^{Y_t})+\mathbb{E}\Phi(Y_s)$$

*where*

$$\Phi(y)=\tfrac12\sum_{i=1}^m b_Y^i(y)^2+\tfrac12\sum_{i,j=1}^m a_Y^{ij}(y)^2+\sum_{i,j=1}^m a_Y^{ij}(y)\,\mathrm{Cov}_{\mu^y}(\partial_{y_i}V,\partial_{y_j}V).$$

*Proof.* Using Lemmas Lemma 8, 9 and 10 together with the inequality $2ab\le a^2+b^2$ we get

$$\begin{aligned}
&\int_{\mathcal{X}} L^Y\log f_t d\rho_t^y\\
&\quad= -\int_{\mathcal{X}} L^Y\log\mu^y d\rho_t^y\\
&\quad= L^Y\log Z(y)-\int_{\mathcal{X}} L^Y\log\mu d\rho_t^y\\
&\quad= b_Y^i(y)\int_{\mathcal{X}}\partial_{y_i}\log\mu\, d(\mu^y-\rho_t^y)+a_Y^{ij}(y)\int_{\mathcal{X}}\partial^2_{y_iy_j}\log\mu\, d(\mu^y-\rho_t^y)+a_Y^{ij}(y)\,\mathrm{Cov}_{\mu^y}(\partial_{y_i}\log\mu,\partial_{y_j}\log\mu)\\
&\quad\le \tfrac12 b_Y^i(y)^2+\tfrac12\|\partial_{y_i}\log\mu\|^2_{\mathrm{Lip}}\Lambda_X c_L H(\rho_t^y|\mu^y)+\tfrac12 a_Y^{ij}(y)^2+\tfrac12\|\partial^2_{y_iy_j}\log\mu\|^2_{\mathrm{Lip}}\Lambda_X c_L H(\rho_t^y|\mu^y)\\
&\qquad+a_Y^{ij}(y)\,\mathrm{Cov}_{\mu^y}(\partial_{y_i}\log\mu,\partial_{y_j}\log\mu).
\end{aligned}$$

The result now follows from the tower property of conditional expectation. □

# 4 Decoupling

We are still in the general setting of section 1.2. We also require that $\sigma_Y(x,y)=\sigma_Y(y)$ only depends on $y$ and that Assumption 1 is in force. The key requirement for the results in this section is a sufficient separation of timescales expressed by assumptions on $\gamma$.

The goal in this subsection is to estimate expressions of the type $\mathbb{E}F(X,Y)$ by $\mathbb{E}F(\tilde{X},Y)$ for any functional $F$ on $\mathcal{W}_{\mathcal{X}}\times\mathcal{W}_{\mathcal{Y}}$.

Denoting $\mathbb{P}$ the Wiener measure on $C([0,T],\mathcal{X}\times\mathcal{Y})$, define a new probability measure $\mathbb{Q}=\mathcal{E}(M)\mathbb{P}$ with

$$dM_t=\Big(\sigma_Y(Y_t)^{-1}(b_Y(\tilde{X}_t,Y_t)-b_Y(X_t,Y_t))\Big)^T dB_t^Y.$$

Corollary 16 will show in particular that under our assumption on $\gamma$ $\mathcal{E}(M)$ is a true martingale so that $\mathbb{Q}$ is indeed a probability measure.

Under this conditions, there is a $\mathbb{Q}$-Brownian motion $\tilde{B}^Y$ such that

$$dY_t=b_Y(\tilde{X}_t,Y_t)dt+\sigma_Y(Y_t)d\tilde{B}_t^Y$$

with

$$d\tilde{B}_t^Y=dB_t^Y-\sigma_Y(Y_t)^{-1}(b_Y(\tilde{X}_t,Y_t)-b_Y(X_t,Y_t))dt.$$

The following Proposition 13 states the key property of $\mathbb{Q}$ which we are going to use.

**Lemma 12.** *Under $\mathbb{Q}$, $B^X$, $\tilde{B}^X$ and $\tilde{B}^Y$ are independent Brownian motions.*

*Proof.* Girsanov's theorem states that if $L$ is a continuous $\mathbb{P}$-local martingale, then $L-\langle L,M\rangle$ is a continuous $\mathbb{Q}$-local martingale. Thus $\tilde{B}^Y=B^Y-\langle B^Y,M\rangle$ is a continuous $\mathbb{Q}$-local martingale by definition, and $B^X,\tilde{B}^X$ are continuous $\mathbb{Q}$-local martingales since $\langle B^X,M\rangle=0$ and $\langle\tilde{B}^X,M\rangle=0$. Since the quadratic variation process is invariant under a change of measure we can conclude using Lévy's characterisation theorem. □

**Proposition 13.** *The laws of $(X,Y,\tilde{X})$ under $\mathbb{P}$ and of $(\tilde{X},Y,X)$ under $\mathbb{Q}$ are equal.*

*Proof.* $(X,Y)$ solves the martingale problem for $L$ under $\mathbb{P}$, and $(\tilde{X},Y)$ solves the martingale problem for $L$ under $\mathbb{Q}$. Since $b_X$ and $b_Y$ are locally Lipschitz, the martingale problem has a unique solution. □

Note in particular that under $\mathbb{Q}$ $B_t^X$ and $Y$ are independent.

The rest of this section is dedicated to show that we can estimate expectations under $\mathbb{P}$ by expectations under $\mathbb{Q}$ when we have a sufficient separation of timescales.

**Lemma 14.** *For any $p>1,q>1$ and $\mathcal{F}_t$-measurable variable $X$*

$$\Big(\mathbb{E}X\Big)^p\le\Big(\mathbb{E}_{\mathbb{Q}}X^p\Big)\Big(\mathbb{E}e^{\lambda(p,q)\langle M\rangle_t}\Big)^{\frac{p-1}{q}}\ \textit{with}\ \lambda(p,q)=\frac{q}{2(p-1)^2}\left(p+\frac{1}{q-1}\right)$$

*Proof.* We have

$$\begin{aligned}\mathbb{E}X&=\mathbb{E}[X\mathcal{E}(M)^{1/p}\mathcal{E}(M)^{-1/p}]\le(\mathbb{E}X^p\mathcal{E}(M))^{1/p}(\mathbb{E}\mathcal{E}(M)^{-p'/p})^{1/p'}\\&=(\mathbb{E}_{\mathbb{Q}}X^p)^{1/p}(\mathbb{E}\mathcal{E}(M)^{-p'/p})^{1/p'}\ \text{with}\ \tfrac{1}{p}+\tfrac{1}{p'}=1\end{aligned}$$

Furthermore, using that for any $\alpha \in \mathbb{R}$ we have $\mathcal{E}(M)^{-\alpha} = \mathcal{E}^{\alpha}(-M)e^{\alpha(1+\alpha)/2}$, we get

$$\begin{aligned}
\mathbb{E}[\mathcal{E}(M)^{-p'/p}] = \mathbb{E}\left[\left(\mathcal{E}(M)^{-q'p'/p}\right)^{1/q'}\right] &= \mathbb{E}\left[\left(\mathcal{E}^{q'p'/p}(-M)\right)^{1/q'}\left(e^{\frac{q'p'}{2p}(\frac{q'p'}{p}+1)\langle M\rangle}\right)^{1/q'}\right] \\
&\leq \left(\mathbb{E}\mathcal{E}^{q'p'/p}(-M)\right)^{1/q'}\left(\mathbb{E}e^{\frac{qp'}{2p}(\frac{q'p'}{p}+1)\langle M\rangle}\right)^{1/q} \text{ with } \frac{1}{q}+\frac{1}{q'}=1 \\
&\leq \left(\mathbb{E}e^{\frac{q}{2(p-1)^2}\left(p+\frac{1}{q-1}\right)\langle M\rangle_t}\right)^{1/q}
\end{aligned}$$

The first expectation in the second line is $\leq 1$ since $\mathcal{E}^{q'p'/p}(-M)$ is a positive local martingale and therefore a supermartingale. Expressing $q'$ and $p'$ in terms of $p$ and $q$ in the second expectation, we pass to the last line and conclude. □

**Lemma 15.** *Under assumption 1 for*

$$\beta \leq \frac{\gamma}{4}$$

*we have*

$$\mathbb{E}\exp\left(\beta\langle M\rangle_t\right) \leq \exp\left(\frac{2\beta\kappa_X(\alpha+n\bar{\lambda}_X)t}{\Lambda_X\gamma}\right)$$

*Proof.* From the definition of $M_t$ we have

$$d\langle M\rangle_t = \left|\sigma_Y{}^{-1}\left(b(X_t,Y_t)-b(\tilde{X}_t,Y_t)\right)\right|^2 dt \leq \frac{1}{\lambda_Y}\left|b(X_t,Y_t)-b(\tilde{X}_t,Y_t)\right|^2 dt \leq \frac{\kappa_Y}{\lambda_Y}\left|X_t-\tilde{X}_t\right|^2 dt.$$

We also have

$$\begin{aligned}
d|X_t-\tilde{X}_t|^2 &= 2\left(X_t-\tilde{X}_t\right)^T\left(b_X(X_t,Y_t)-b_X(\tilde{X}_t,Y_t)\right)dt \\
&\quad + 2\left(X_t-\tilde{X}_t\right)^T\left(\sigma_X(X_t,Y_t)dB_t^X-\sigma_X(\tilde{X}_t,Y_t)d\tilde{B}_t^X\right) \\
&\quad + 2\operatorname{Tr}(A_X(X_t,Y_t))dt + 2\operatorname{Tr}(A_X(\tilde{X}_t,Y_t))dt \\
&\overset{(\mathrm{m})}{\leq} -2\kappa_X|X_t-\tilde{X}_t|^2dt + 2(\alpha+n\bar{\lambda}_X)dt
\end{aligned}$$

where $\overset{(\mathrm{m})}{\leq}$ means inequality modulo local martingales, and

$$\begin{aligned}
d\langle|X_t-\tilde{X}_t|^2\rangle &= 4(X_t-\tilde{X}_t)^T\left(A_X(X_t,Y_t)+A_X(\tilde{X}_t,Y_t)\right)(X_t-\tilde{X}_t) \\
&\leq 8\Lambda_X|X_t-\tilde{X}_t|^2
\end{aligned}$$

so that

$$\begin{aligned}
de^{\frac{r}{2}|X_t-\tilde{X}_t|^2}e^{\beta\langle M\rangle_t} &= \left(\frac{r}{2}d|X_t-\tilde{X}_t|^2 + \beta d\langle M\rangle_t + \frac{r^2}{8}d\langle|X_t-\tilde{X}_t|^2\rangle\right)e^{\frac{r}{2}|X_t-\tilde{X}_t|^2}e^{\beta\langle M\rangle_t} \\
&\overset{(\mathrm{m})}{\leq} \left(\left(r^2\Lambda_X - r\kappa_X + \frac{\beta\kappa_Y^2}{\lambda_Y^2}\right)|X_t-\tilde{X}_t|^2 + r(\alpha+n\bar{\lambda}_X)\right)e^{\frac{r}{2}|X_t-\tilde{X}_t|^2}e^{\beta\langle M\rangle_t}dt \\
&= \left(\Lambda_X(r-r_-)(r-r_+)|X_t-\tilde{X}_t|^2 + r(\alpha+n\bar{\lambda}_X)\right)e^{\frac{r}{2}|X_t-\tilde{X}_t|^2}e^{\beta\langle M\rangle_t}dt
\end{aligned}$$

with

$$r_{\pm} = \frac{\kappa_X}{2\Lambda_X}\left(1 \pm \sqrt{1-4\beta/\gamma}\right).$$

According to our assumptions, $1-4\beta/\gamma > 0$ and we have, choosing $r = r_-$

$$de^{\frac{r_-}{2}|X_t-\tilde{X}_t|^2}e^{\beta\langle M\rangle_t} \overset{\text{(m)}}{\leq} r_-(\alpha+n\bar{\lambda}_X)e^{\frac{r_-}{2}|X_t-\tilde{X}_t|^2}e^{\beta\langle M\rangle_t}dt$$

so that

$$e^{\frac{r_-}{2}|X_t-\tilde{X}_t|^2}e^{\beta\langle M\rangle_t} \overset{\text{(m)}}{\leq} e^{r_-(\alpha+n\bar{\lambda}_X)t}$$

and

$$\mathbb{E}e^{\beta\langle M\rangle_t} \leq \mathbb{E}e^{\frac{r_-}{2}|X_t-\tilde{X}_t|^2}e^{\beta\langle M\rangle_t} \leq e^{r_-(\alpha+n\bar{\lambda}_X)t}.$$

Since $1-\sqrt{1-x} \leq x$ for $0 \leq x \leq 1$ we have furthermore

$$r_- \leq \frac{\kappa_X}{2\Lambda_X}\frac{4\beta}{\gamma}$$

so that

$$\mathbb{E}e^{\beta\langle M\rangle_t} \leq \exp\left(\frac{2\beta\kappa_X(\alpha+n\bar{\lambda}_X)t}{\Lambda_X\gamma}\right).$$

□

**Corollary 16.** *If* $\gamma > 2$ *then*

$$\mathcal{E}(M)_t \textit{ is a true martingale.}$$

*Proof.* Since $\frac{1}{2} < \frac{\gamma}{4}$ by our assumption we get from the previous Proposition that

$$\mathbb{E}\left[e^{\frac{1}{2}\langle M\rangle_t}\right] < \infty$$

and Novikov's criterion leads directly to the conclusion. □

**Proposition 17.** *Under assumption 1 for any* $\mathcal{F}_t$*-measurable random variable* $Z$ *and* $1+\frac{2}{\gamma}+2\sqrt{\frac{2}{\gamma}} \leq p \leq 2$

$$(\mathbb{E}Z)^p \leq \mathbb{E}_{\mathbb{Q}}\left[Z^p\right]\exp\left(\frac{p\,\kappa_X(\alpha+n\bar{\lambda}_X)\,t}{\left(p-1-\sqrt{2/\gamma}\right)\Lambda_X\,\gamma}\right)$$

*Proof.* We would like to apply Lemmas 14 and 15, so we need to find conditions that ensure the existence of a $q$ such that $\lambda(p,q) \leq \frac{\gamma}{4}$.

After some straightforward computations we get the identities

$$\begin{aligned}\lambda(p,q)-\frac{\gamma}{4} &= \frac{p(q-q_-)(q-q_+)}{2(p-1)^2(q-1)},\\ q_\pm &= \frac{\gamma(p-1)}{4p}\left(p-1+\frac{2}{\gamma}\pm\sqrt{(p-p_-)(p-p_+)}\right),\\ p_\pm &= 1+\tfrac{2}{\gamma}\pm 2\sqrt{\tfrac{2}{\gamma}}.\end{aligned}$$

Our assumption on $p$ implies that $1+\frac{2}{\gamma}+2\sqrt{\frac{2}{\gamma}}\le 2 \iff \gamma \ge \frac{1}{(\sqrt{3}-\sqrt{2})^2} > 2$ so that $p-p_- > p-1+\frac{2}{\gamma} > 0$ and by our assumption on $p$, $p-p_+>0$ as well so that $q_\pm$ is real and $\lambda(p,q_+)=\frac{\gamma}{4}$.

For our particular values of $p_-$ and $p_+$ we have furthermore $(p-p_-)(p-p_+)\ge (p-p_+)^2$ so that

$$q_+ \ge \frac{\gamma(p-1)(p-1-\sqrt{\frac{2}{\gamma}})}{2p}$$

Now, apply Lemma 14 with $q=q_+$ to obtain

$$\mathbb{E}[Z]^p \le \mathbb{E}_{\mathbb{Q}}\big[Z^p\big]\mathbb{E}\big[e^{\frac{\gamma}{4}\langle M\rangle_t}\big]^{\frac{p-1}{q_+}}.$$

We estimate the second expectation on the right hand side using Proposition 15

$$\begin{aligned}\mathbb{E}\big[e^{\frac{\gamma}{4}\langle M\rangle_t}\big]^{\frac{p-1}{q_+}} &\le \exp\left(\frac{(p-1)}{q_+}\frac{\kappa_X(\alpha+n\bar{\lambda}_X)t}{2\Lambda_X}\right)\\ &\le \exp\left(\frac{p\,\kappa_X(\alpha+n\bar{\lambda}_X)\,t}{\left(p-1-\sqrt{2/\gamma}\right)\Lambda_X\,\gamma}\right)\end{aligned}$$

which leads to our result. □

# 5 Proof of the main theorem

**Lemma 18.** *If $\bar{b}_Y$ is Lipschitz then*

$$\sup_{0\le t\le T}|Y_t-\bar{Y}_t| \le \sup_{0\le t\le T}\left|\int_0^t b_Y(X_s,Y_s)-\bar{b}(Y_s)ds\right|e^{\|\bar{b}\|_{\mathrm{Lip}}T}$$

*Proof.*

$$\begin{aligned}\sup_{0\le t\le T}|Y_t-\bar{Y}_t| &= \sup_{0\le t\le T}\left|\int_0^t b_Y(X_s,Y_s)-\bar{b}_Y(\bar{Y}_s)ds\right| \\ &\le \sup_{0\le t\le T}\left|\int_0^t b_Y(X_s,Y_s)-\bar{b}_Y(Y_s)ds\right| + \|\bar{b}\|_{\mathrm{Lip}}\int_0^T \sup_{0\le s\le t}|Y_s-\bar{Y}_s|ds\end{aligned}$$

and the conclusion follows from Gronwall's inequality. □

**Theorem 19.** *Under Assumption 1 if $\sigma_Y(x,y)=\sigma_Y(y)$, a Poincaré inequality with constant $c_P$ holds for $\tilde{\rho}_t^y$, a Logarithmic Sobolev inequality with constant $c_L$ holds for $\mu^y(dx)=Z(y)^{-1}e^{-V(x,y)}dx$ both with respect to $\Gamma^X$, $X_0\sim\mu^{Y_0}$ and $\bar{b}$ is Lipschitz then for $1\le p\le \frac{2}{1+\frac{2}{\gamma}+2\sqrt{\frac{2}{\gamma}}}$ we have the estimate*

$$\mathbb{E}\left[\sup_{0\le t\le T}|Y_t-\bar{Y}_t|^p\right]^{2/p} \le m{\kappa_Y}^2\Lambda_X\left(27{c_P}^2T+\frac{2{c_L}^2}{4-{c_L}^2\Lambda_X(m{\kappa_Y}^2+3{c_V}^2)}\mathbb{E}\int_0^T\Psi(Y_t)dt\right)$$
$$\exp\left(\frac{2p'\kappa_X(\alpha+n\bar{\lambda}_X)T}{p\gamma\Lambda_X}+2\|\bar{b}\|_{\mathrm{Lip}}T\right)$$

*with*

$$\Psi(y)=\frac{3m{\kappa_Y}^2(\alpha+n\bar{\lambda}_X)}{2\kappa_X}+\tfrac{3}{2}|\bar{b}(y)|^2+\tfrac{1}{2}\sum_{i,j=1}^m a_Y^{ij}(y)^2+\sum_{i,j=1}^m a_Y^{ij}(y)\,\mathrm{Cov}_{\mu^y}(\partial_{y_i}V,\partial_{y_j}V),$$

$$p'=\frac{1}{1-\frac{p}{2}\left(1+\sqrt{\frac{2}{\gamma}}\right)}>\frac{2}{2-p}$$

*and*

$${c_V}^2=\sup_y\left(\sum_{i=1}^m\|\partial_{y_i}V(\cdot,y)\|_{\mathrm{Lip}}^2+\sum_{i,j=1}^m\|\partial_{y_iy_j}^2V(\cdot,y)\|_{\mathrm{Lip}}^2\right).$$

*Proof.* By Lemma 18 we have

$$\mathbb{E}\left[\sup_{0\le t\le T}|Y_t-\bar{Y}_t|^p\right]\le \mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t b_Y(X_s,Y_s)-\bar{b}(Y_s)ds\right|^p\right]e^{p\|\bar{b}\|_{\mathrm{Lip}}T}.$$

Using Proposition 17 we get for $1\le p\le\frac{2}{1+\frac{2}{\gamma}+2\sqrt{\frac{2}{\gamma}}}$ that

$$\begin{aligned}&\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t b_Y(X_s,Y_s)-\bar{b}(Y_s)ds\right|^p\right] \\ &\qquad\le \mathbb{E}_{\mathbb{Q}}\left[\sup_{0\le t\le T}\left|\int_0^t b_Y(X_s,Y_s)-\bar{b}(Y_s)ds\right|^2\right]^{p/2}\exp\left(\frac{p'\kappa_X(\alpha+n\bar{\lambda}_X)T}{\gamma\Lambda_X}\right)\end{aligned}$$

with

$$0 < p' = \frac{1}{1 - \frac{p}{2}\left(1 + \sqrt{\frac{2}{\gamma}}\right)} < \infty.$$

By Proposition 13

$$\mathbb{E}_{\mathbb{Q}}\left[\sup_{0\le t\le T}\left|\int_0^t b_Y(X_s, Y_s) - \bar{b}(Y_s)ds\right|^2\right] = \mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t b_Y(\tilde{X}_s, Y_s) - \bar{b}(Y_s)ds\right|^2\right].$$

Now we decompose

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t b_Y(\tilde{X}_s, Y_s) - \bar{b}(Y_s)ds\right|^2\right] &\le 2\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t b_Y(\tilde{X}_s, Y_s) - \mathbb{E}[b_Y(\tilde{X}_s, Y_s)|(X_s, Y_s)]ds\right|^2\right] \\ &+ 2\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t \mathbb{E}[b_Y(\tilde{X}_s, Y_s)|(X_s, Y_s)] - \bar{b}(Y_s)ds\right|^2\right]. \quad (3)\end{aligned}$$

For the rest of the proof we put ourselves in the setting of section 1.2 where we substitute $\tilde{X}$ for $X$ and $(X, Y)$ for $Y$.

For $1 \le i \le m$ we now apply Proposition 6 with $\phi : (x, y) \mapsto y$, $\nu_t^y = \tilde{\rho}_t^y$ and $f_t(\tilde{x}, x, y) = b_Y^i(\tilde{x}, y) - \mathbb{E}[b_Y^i(\tilde{X}_s, Y_s)|(X_s, Y_s) = (x, y)]$. Since $\tilde{\rho}_t^y$ satisfies a Poincaré inequality by assumption and $\int f_t(\cdot, y)d\tilde{\rho}_t^y = 0$ by the properties of conditional expectation, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t b_Y^i(\tilde{X}_s, Y_s) - \mathbb{E}[b_Y^i(\tilde{X}_s, Y_s)|(X_s, Y_s)]ds\right|^2\right] &\le \frac{27}{2}\int_0^T \mathbb{E}\left[c_P(Y_t) f_t^2(\tilde{X}_t, X_t, Y_t)\right] dt \\ &\le \frac{27}{2}\int_0^T \mathbb{E}\left[c_P(Y_t)^2 \Gamma^X(b_Y^i)(\tilde{X}_t, Y_t)\right] dt \\ &\le \frac{27{c_P}^2\Lambda_X\|\nabla_x b_Y^i\|_\infty^2 T}{2}\end{aligned}$$

where the second inequality follows from the tower property of conditional expectation and applying the Poincaré inequality a second time to $\tilde{\rho}_t^y$ and the last line from $\Gamma^X(b_Y^i) = {\nabla_x b_Y^i}^T A_X \nabla_x b_Y^i \le \Lambda_X|\nabla_x b_Y^i|^2 \le \Lambda_X\|\nabla_x b_Y^i\|_\infty^2$. Summing over the components $b_Y^i$ we get

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t b_Y(\tilde{X}_s, Y_s) - \mathbb{E}[b_Y(\tilde{X}_s, Y_s)|(X_s, Y_s)]ds\right|^2\right] &\le \frac{27{c_P}^2\Lambda_X T}{2}\sum_{i=1}^m \|\nabla_x b_Y^i\|_\infty^2 \\ &= \frac{27{c_P}^2\Lambda_X m{\kappa_Y}^2 T}{2}\end{aligned}$$

We now turn to the second term on the right hand side in the decomposition (3). First, note that

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t \mathbb{E}[b_Y(\tilde{X}_s,Y_s)|(X_s,Y_s)] - \bar{b}(Y_s)ds\right|^2\right] \\
&\le \mathbb{E}\int_0^T \left|\mathbb{E}[b_Y(\tilde{X}_s,Y_s)|(X_s,Y_s)] - \bar{b}(Y_s)ds\right|^2 \\
&= \sum_{i=1}^n \int_0^T \mathbb{E}\left|\int_{\mathcal{X}} b_Y^i(\tilde{x},Y_t)\tilde{\rho}^{X_t,Y_t}(d\tilde{x}) - \int_{\mathcal{X}} b_Y^i(\tilde{x},Y_t)\mu^{Y_t}(d\tilde{x})\right|^2
\end{aligned}$$

By Lemma 10 we have

$$\begin{aligned}
&\sum_{i=1}^n \mathbb{E}\left|\int_{\mathcal{X}} b_Y^i(\tilde{x},Y_t)\tilde{\rho}^{X_t,Y_t}(d\tilde{x}) - \int_{\mathcal{X}} b_Y^i(\tilde{x},Y_t)\mu^{Y_t}(d\tilde{x})\right|^2 \\
&\qquad \le \sum_{i=1}^n \|b_Y^i\|_{\mathrm{Lip}}^2 \Lambda_X c_L \mathbb{E}H(\tilde{\rho}_t^{X_t,Y_t}|\mu^{Y_t}) \le m{\kappa_Y}^2 \Lambda_X c_L \mathbb{E}H(\tilde{\rho}_t^{X_t,Y_t}|\mu^{Y_t}).
\end{aligned}$$

Suppose that uniformly in $y$

$$\left(\sum_{i=1}^m \|\partial_{y_i} V(\cdot,y)\|_{\mathrm{Lip}}^2 + \sum_{i,j=1}^m \|\partial_{y_i y_j}^2 V(\cdot,y)\|_{\mathrm{Lip}}^2\right) < {c_V}^2.$$

Now, for some $r\in\mathbb{R}$ to be fixed later, use Propositions 7 and 11 to get

$$\mathbb{E}H(\tilde{\rho}_t^{X_t,Y_t}|\mu^{Y_t})e^{rt} \le -\left(\frac{2}{c_L} - \frac{{c_V}^2\Lambda_X c_L}{2} - r\right)\int_0^t \mathbb{E}H(\rho_s^{X_s,Y_s}|\mu^{Y_s})e^{rs}ds + \int_0^t \mathbb{E}\Phi(X_s,Y_s)e^{rs}ds. \tag{4}$$

We have

$$\mathbb{E}\Phi(X_s,Y_s) = \mathbb{E}\left[\tfrac{1}{2}\sum_{i=1}^m b_Y^i(X_s,Y_s)^2 + \tfrac{1}{2}\sum_{i,j=1}^m a_Y^{ij}(Y_s)^2 + \sum_{i,j=1}^m a_Y^{ij}(Y_s)\operatorname{Cov}_{\mu^{Y_s}}(\partial_{y_i}V,\partial_{y_j}V)\right]$$

and we estimate the first term on the right hand side as follows:

$$\begin{aligned}
\mathbb{E}b_Y^i(X_s,Y_s)^2 &\le 3\mathbb{E}[b_Y^i(X_s,Y_s) - b_Y^i(\tilde{X}_s,Y_s)]^2 \\
&\quad + 3[\mathbb{E}b_Y^i(\tilde{X}_s,Y_s) - \int_{\mathcal{X}} b_Y^i(x,Y_s)\mu^{Y_s}(dx)]^2 + 3\left|\int_{\mathcal{X}} b_Y^i(x,Y_s)\mu^{Y_s}(dx)\right|^2.
\end{aligned}$$

Since $b_Y$ is Lipschitz in the first variable we get for the first term

$$\sum_{i=1}^m \mathbb{E}[b_Y^i(X_s,Y_s) - b_Y^i(\tilde{X}_s,Y_s)]^2 = \mathbb{E}\left|b_Y(X_s,Y_s) - b_Y(\tilde{X}_s,Y_s)\right|^2 \le m{\kappa_Y}^2\mathbb{E}|X_s-\tilde{X}_s|^2 \le \frac{m{\kappa_Y}^2(\alpha+n\bar{\lambda}_X)}{\kappa_X}.$$

Still using the Lipschitzness of $b_Y$, we use Lemma 10 together with the tower property for conditional expectation on the second term to get

$$\sum_{i=1}^{m}[\mathbb{E}b_Y^i(\tilde{X}_s, Y_s) - \int_{\mathcal{X}} b_Y^i(x, Y_s)\mu^{Y_s}(dx)]^2 \leq m{\kappa_Y}^2 c_L \Lambda_X \mathbb{E}\,\mathrm{H}(\tilde{\rho}_s^{X_s,Y_s}|\mu^{Y_s}).$$

This leads us to

$$\mathbb{E}\Phi(X_s, Y_s) \leq \tfrac{3}{2}\left(m{\kappa_Y}^2 c_L \Lambda_X \mathbb{E}\,\mathrm{H}(\tilde{\rho}_s^{X_s,Y_s}|\mu^{Y_s}) + \frac{m{\kappa_Y}^2(\alpha + n\bar{\lambda}_X)}{\kappa_X} + \mathbb{E}|\bar{b}(Y_s)|^2\right) + \tfrac{1}{2}(a_Y^{ij}(Y_s) + a_Y^{ij}(Y_s)\,\mathrm{Cov}_{\mu^{Y_s}}(\partial_{y_i}V, \partial_{y_j}V).$$

Substituting $\Phi$ in (4) we get

$$\begin{aligned}\mathbb{E}H(\tilde{\rho}_t^{X_t,Y_t}|\mu^{Y_t})e^{rt} \leq &-\left(\frac{2}{c_L} - \frac{\Lambda_X c_L(m{\kappa_Y}^2 + 3{c_V}^2)}{2} - r\right)\int_0^t \mathbb{E}H(\rho_s^{X_s,Y_s}|\mu^{Y_s})e^{rs}ds \\ &+\mathbb{E}\int_0^t e^{rs}\frac{3m{\kappa_Y}^2(\alpha + n\bar{\lambda}_X)}{2\kappa_X} + \tfrac{3}{2}|\bar{b}(Y_s)|^2 + \tfrac{1}{2}\sum_{i,j=1}^{m} a_Y^{ij}(Y_s)^2 \\ &+ \sum_{i,j=1}^{m} a_Y^{ij}(Y_s)\,\mathrm{Cov}_{\mu^{Y_s}}(\partial_{y_i}V, \partial_{y_j}V)ds.\end{aligned}$$

Now we choose

$$r = \frac{2}{c_L} - \frac{\Lambda_X c_L(m{\kappa_Y}^2 + 3{c_V}^2)}{2}$$

so that

$$\mathbb{E}H(\tilde{\rho}_t^{X_t,Y_t}|\mu^{Y_t}) \leq \mathbb{E}\int_0^t e^{-r(t-s)}\Psi(Y_s)ds.$$

with

$$\Psi(y) = \frac{3m{\kappa_Y}^2(\alpha + n\bar{\lambda}_X)}{2\kappa_X} + \tfrac{3}{2}|\bar{b}(y)|^2 + \tfrac{1}{2}\sum_{i,j=1}^{m} a_Y^{ij}(y)^2 + \sum_{i,j=1}^{m} a_Y^{ij}(y)\,\mathrm{Cov}_{\mu^y}(\partial_{y_i}V, \partial_{y_j}V).$$

By the preceding inequality and the Young inequality for convolutions on $L^1([0,T])$

$$\begin{aligned}\int_0^T \mathbb{E}H(\tilde{\rho}_t^{X_t,Y_t}|\mu^{Y_t})dt &\leq \mathbb{E}\int_0^T\int_0^t e^{-r(t-s)}\Psi(Y_s)dsdt \\ &\leq \mathbb{E}\int_0^T e^{-rt}dt\int_0^T \Psi(Y_t)dt \\ &= \frac{1}{r}(1 - e^{-rT})\mathbb{E}\int_0^T \Psi(Y_t)dt\end{aligned}$$

so that finally

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t \mathbb{E}[b_Y(\tilde{X}_s,Y_s)|(X_s,Y_s)]-\bar{b}(Y_s)ds\right|^2\right] \le \frac{m{\kappa_Y}^2\Lambda_X c_L}{r}(1-e^{-rT})\mathbb{E}\int_0^T \Psi(Y_t)dt$$
$$= \frac{2{c_L}^2\Lambda_X m{\kappa_Y}^2}{4-{c_L}^2\Lambda_X(m{\kappa_Y}^2+3{c_V}^2)}(1-e^{-rT})\mathbb{E}\int_0^T \Psi(Y_t)dt$$
$$\le \frac{2{c_L}^2\Lambda_X m{\kappa_Y}^2}{4-{c_L}^2\Lambda_X(m{\kappa_Y}^2+3{c_V}^2)}\mathbb{E}\int_0^T \Psi(Y_t)dt.$$

Assembling the previous results, we obtain

$$\mathbb{E}\left[\sup_{0\le t\le T}|Y_t-\bar{Y}_t|^p\right]^{2/p} \le m{\kappa_Y}^2\Lambda_X\left(27{c_P}^2T+\frac{2{c_L}^2}{4-{c_L}^2\Lambda_X(m{\kappa_Y}^2+3{c_V}^2)}\mathbb{E}\int_0^T\Psi(Y_t)dt\right)$$
$$\exp\left(\frac{2p'\kappa_X(\alpha+n\bar{\lambda}_X)T}{p\gamma\Lambda_X}+2\|\bar{b}\|_{\mathrm{Lip}}T\right).$$

□

# 6 Applications

## 6.1 Averaging

For $\varepsilon>0$ fixed consider an SDE of the form

$$dX_t = -\varepsilon^{-1}\nabla_x V(X_t,Y_t)dt+\varepsilon^{-1/2}\sqrt{2\beta_X^{-1}}dB_t^X \tag{5}$$
$$dY_t = b_Y(X_t,Y_t)dt+\sqrt{2\beta_Y^{-1}}dB_t^Y \tag{6}$$

with $Y_0=y_0\in\mathbb{R}^m$ and $X_0\sim\mu^{y_0}=e^{-\beta V(x,y_0)}dx$ and $V(x,y)$ is of the form

$$V(x,y)=\frac{1}{2}(x-g(y))Q(x-g(y))+h(x,y)$$

where $h$ is uniformly bounded in both arguments and both $\partial_y h$ and $\partial_y^2 h$ are Lipschitz in $x$ uniformly in $y$. Under these conditions

$$\mu^y(dx)=Z(y)^{-1}e^{-\beta_X V(x,y)}dx \text{ with } Z(y)=\int_{\mathcal{X}} e^{-\beta_X V(x,y)}dx$$

is a Gaussian measure with covariance matrix $\beta_X Q$ and mean $g(y)$ perturbed by a bounded factor $e^{-\beta_X h(x,y)}$. As such it satisfies a Logarithmic Sobolev inequality with respect to the usual square field operator $|\nabla|^2$ with constant

$$c_L^0=(\beta_X\lambda_Q)^{-1}e^{\beta_X\,\mathrm{osc}(h)}\quad\text{with }\ \mathrm{osc}(h)=\sup h-\inf h$$

and $\lambda_Q$ is the smallest eigenvalue of $Q$. In particular, $\mu^y$ satisfies a Logarithmic Sobolev inequality with constant

$$c_L = \varepsilon \lambda_Q^{-1} e^{\beta_X \operatorname{osc}(h)}$$

with respect to $\Gamma^X = \varepsilon^{-1}\beta_X^{-1}|\nabla|^2$.

We have

$$\begin{aligned}
&-(x_1-x_2)^T(\nabla_x V(x_1,y) - \nabla_x V(x_2,y)) \\
&\quad = -(x_1-x_2)^T Q(x_1-x_2) - (x_1-x_2)^T(\nabla_x h(x_1,y) - \nabla_x h(x_2,y)) \\
&\quad \le -\lambda_Q|x_1-x_2|^2 + |x_1-x_2|\|\nabla_x h\|_\infty \\
&\quad \le -\lambda_Q|x_1-x_2|^2 + \frac{\|\nabla_x h\|_\infty}{4\lambda_Q}
\end{aligned}$$

so that we can choose

$$\kappa_X = \varepsilon^{-1}\lambda_Q, \qquad\qquad \alpha = \varepsilon^{-1}\frac{\|\nabla_x h\|_\infty}{4\lambda_Q}.$$

We also have trivially

$$\lambda_X = \Lambda_X = \bar{\lambda}_X = \varepsilon^{-1}\beta_X^{-1}, \qquad \Lambda_Y = \beta_Y^{-1}, \qquad \kappa_Y = \|\nabla_x b_Y\|_\infty$$

and the separation of timescales is

$$\gamma = \frac{{\kappa_X}^2\lambda_Y}{\Lambda_X m {\kappa_Y}^2} = \varepsilon^{-1}\frac{{\lambda_Q}^2\beta_Y^{-1}}{\|\nabla_x b_Y\|_\infty^2 \beta_X^{-1}}.$$

If $\gamma > \frac{1}{(\sqrt{3}-\sqrt{2})^2} \approx 9.899$ we can apply Theorem 19 with $p=1$ to get

$$\mathbb{E}\left[\sup_{0\le t\le T}|Y_t - \bar{Y}_t|\right]^2 \le \varepsilon C_1\left(27(c_P(\varepsilon)/c_L)^2 T + C_2\mathbb{E}\int_0^T \Psi(Y_t)dt\right)\exp\left(2p' C_3 T + 2\|\bar{b}\|_{\mathrm{Lip}}T\right)$$

with

$$\begin{aligned}
C_1 &= \varepsilon^{-1} m {\kappa_Y}^2 \Lambda_X {c_L}^2 = m{\kappa_Y}^2\beta_X^{-1}\lambda_Q^{-2}e^{2\beta_X \operatorname{osc}(h)}, \\
C_2 &= \frac{2}{4 - {c_L}^2\Lambda_X(m{\kappa_Y}^2 + 3{c_V}^2)} \le 1 \text{ for } \varepsilon \le \frac{2\lambda_Q e^{-\beta_X \operatorname{osc}(h)}\beta_X}{\|\nabla_X b_Y\|_\infty^2 + 3{c_V}^2}, \\
C_3 &= \frac{\kappa_X(\alpha + n\bar{\lambda}_X)}{\gamma\Lambda_X} = \frac{\|\nabla_x b_Y\|_\infty^2(\frac{\|\nabla_x h\|_\infty}{4\lambda_Q} + n\beta_X^{-1})}{\beta_Y^{-1}\lambda_Q},
\end{aligned}$$

$$
\begin{aligned}
\Psi(y) &= \frac{3m{\kappa_Y}^2(\alpha+n\bar\lambda_X)}{2\kappa_X} + \tfrac{3}{2}|\bar b(y)|^2 + \tfrac12 \sum_{i,j=1}^m a_Y^{ij}(y)^2 + \sum_{i,j=1}^m a_Y^{ij}(y)\operatorname{Cov}_{\mu^y}(\partial_{y_i}\beta_X V, \partial_{y_j}\beta_X V)\\
&= \frac{3\|\nabla_x b_Y\|_\infty^2(\frac{\|\nabla_x h\|_\infty}{4\lambda_Q}+n\beta_X^{-1})}{2\lambda_Q} + \tfrac12 m\beta_Y^{-2} + \beta_Y^{-1}\beta_X^2\sum_i \operatorname{Var}_{\mu^y}(\partial_{y_i}V) + \tfrac32|\bar b(y)|^2\\
&\le \frac{3\|\nabla_x b_Y\|_\infty^2(\frac{\|\nabla_x h\|_\infty}{4\lambda_Q}+n\beta_X^{-1})}{2\lambda_Q} + \tfrac12 m\beta_Y^{-2} + \beta_Y^{-1}\beta_X^2 c_L^0\sum_i \|\partial_{y_i}V\|_{\mathrm{Lip}}^2 + \tfrac32|\bar b(y)|^2\\
&= \frac{3\|\nabla_x b_Y\|_\infty^2(\frac{\|\nabla_x h\|_\infty}{4\lambda_Q}+n\beta_X^{-1})}{2\lambda_Q} + \tfrac12 m\beta_Y^{-2} + \beta_X(\beta_Y\lambda_Q)^{-1}e^{\beta_X\operatorname{osc}(h)}\sum_i \|\partial_{y_i}V\|_{\mathrm{Lip}}^2 + \tfrac32|\bar b(y)|^2,
\end{aligned}
$$

$$
2 < p' = \frac{1}{1-\frac12\left(1+\sqrt{\frac{2}{\gamma}}\right)} < \frac{2}{3-\sqrt2\sqrt3} \approx 3.633
$$

and

$$
{c_V}^2 = \sup_y \left(\sum_{i=1}^m \|\partial_{y_i}V(\cdot,y)\|_{\mathrm{Lip}}^2 + \sum_{i,j=1}^m \|\partial_{y_iy_j}^2 V(\cdot,y)\|_{\mathrm{Lip}}^2\right).
$$

If we suppose that $c_P(\varepsilon)/c_L$ converges to a finite limit as $\varepsilon\to0$ and that

$$
\mathbb{E}\int_0^T \bar b(Y_t)^2 dt < \infty
$$

then there exists a constant $C$ depending on $T, V, \beta_X, b_Y$ and $\beta_Y$ such that for $\varepsilon$ sufficiently small

$$
\mathbb{E}\sup_{0\le t\le T}|Y_t-\bar Y_t| \le \sqrt{\varepsilon}C.
$$

In other words, we obtain a strong averaging principle of order 1/2 in $\varepsilon$.

## 6.2 Temperature-Accelerated Molecular Dynamics

In [MV06] the authors introduced the TAMD process $(X_t, Y_t)$ and its averaged version $\bar Y_t$ defined by

$$
\begin{aligned}
dX_t &= -\tfrac{1}{\varepsilon}\nabla_x U(X_t,Y_t)dt + \sqrt{2(\beta\varepsilon)^{-1}}dB_t^X, \quad X_0 \sim e^{-\beta U(x,y_0)}dx\\
dY_t &= -\tfrac{1}{\bar\gamma}\kappa(Y_t-\theta(X_t))dt + \sqrt{2(\bar\beta\bar\gamma)^{-1}}dB_t^Y, \quad Y_0 = y_0\\
d\bar Y_t &= \bar b(\bar Y_t)dt + \sqrt{2(\bar\beta\bar\gamma)^{-1}}dB_t^Y, \quad \bar Y_0 = y_0\\
U(x,y) &= V(x) + \tfrac{\kappa}{2}|y-\theta(x)|^2,\\
\bar b(y) &= Z(y)^{-1}\int -\bar\gamma^{-1}\kappa(y-\theta(x))e^{-\frac{\kappa}{2}|y-\theta(x)|^2}e^{-V(x)}dx, \quad Z(y) = \int e^{-\frac{\kappa}{2}|y-\theta(x)|^2}e^{-V(x)}dx
\end{aligned}
$$

with $X_t \in \mathbb{R}^n$, $Y_t, \bar{Y}_t \in \mathbb{R}^m$, a Lipschitz-continuous function $V(x)$, constants $\kappa, \varepsilon, \beta, \bar{\beta}, \bar{\gamma} > 0$ and independent standard Brownian motions $B^X$, $B^Y$ on $\mathbb{R}^n$ and $\mathbb{R}^m$.

Let $D \subset \mathbb{R}^m$ be a compact set and define the stopping time $\tau = \inf\{t \geq 0 : Y_t \notin D\}$.

We will show that under some additional assumptions, a strong averaging principle with rate $1/2$ holds in the sense that for any fixed $T$ and $\varepsilon$ sufficiently small but fixed, there exists a constant $C$ not depending on $\varepsilon$ such that

$$\sup_{0 \leq t \leq T} |Y_{t\wedge\tau} - \bar{Y}_{t\wedge\tau}| \leq C\varepsilon^{1/2}.$$

We need the following extra assumptions on the TAMD process:

$$\begin{gathered}
0 < \lambda_\theta \mathrm{Id}_m < D\theta(x) D\theta(x)^T < \Lambda_\theta \mathrm{Id}_m < \infty, \\
-(x_1 - x_2)^T (\nabla_x(\theta(x_1) - y)^2 - \nabla_x(\theta(x_2) - y)^2) \leq -\kappa_\theta |x_1 - x_2|^2 + \alpha_\theta \\
\lim_{|x|\to\infty} |\theta(x)| = \infty \\
\lambda_\theta \kappa > \Lambda_\theta \beta^{-1}.
\end{gathered}$$

In order to apply Theorem 19 we also need to suppose that Assumption 2 holds for the TAMD process.

We will now briefly comment on the form of $\bar{Y}_t$. Let

$$\mu(dx) = Z_0^{-1} e^{-V(x)} dx, \quad Z_0 = \int e^{-V(x)} dx$$

so that

$$\begin{aligned}
\bar{b}(y) &= \frac{Z_0}{Z(y)} \int -\bar{\gamma}^{-1} \kappa(\theta(x) - y) e^{-\frac{\kappa}{2}|\theta(x) - y|^2} \mu(dx) \\
&= \frac{Z_0}{Z(y)} \bar{\gamma}^{-1} \int -\kappa(z - y) e^{-\frac{\kappa}{2}|z-y|^2} \theta_\# \mu(dz) \\
&= \frac{Z_0}{Z(y)} \bar{\gamma}^{-1} \nabla_y \int e^{-\frac{\kappa}{2}|z-y|^2} \theta_\# \mu(dz)
\end{aligned}$$

where $\theta_\# \mu$ denotes the image measure of $\mu$ by $\theta$. Now note that

$$\frac{Z(y)}{Z_0} = \int e^{-\frac{\kappa}{2}|\theta(x)-y|^2} \mu(dx) = \int e^{-\frac{\kappa}{2}|z-y|^2} \theta_\# \mu(dz)$$

so that

$$\bar{b}(y) = \bar{\gamma}^{-1} \nabla_y \log \int e^{-\frac{\kappa}{2}|z-y|^2} \theta_\# \mu(dz) = \nabla_y \log(\theta_\# \mu * \mathcal{N}(0, \kappa^{-1}))(y).$$

In the last expression, $*$ denotes convolution, $\mathcal{N}(0, \kappa^{-1})$ denotes the Gaussian measure with mean 0 and variance $\kappa^{-1}$ and we identify through an abuse of notation measures and their densities which we suppose to exist.

Thus,

$$d\bar{Y}_t = \bar{\gamma}^{-1}\nabla_y \log(\theta_{\#}\mu * \mathcal{N}(0,\kappa^{-1}))(\bar{Y}_t)dt + \sqrt{2(\bar{\beta}\bar{\gamma})^{-1}}dB_t^Y.$$

In physical terms, $\bar{Y}_t$ evolves at an inverse temperature of $\bar{\beta}$ on the energy landscape corresponding to the image measure of $\mu$ by $\theta$ convolved with a Gaussian measure of variance $\kappa^{-1}$.

We proceed to establish a Logarithmic Sobolev inequality for $\mu^y$ via the Lyapunov function method. From [CG17] Theorem 1.2 it follows that a sufficient condition for a Logarithmic Sobolev inequality to hold for an elliptic, reversible diffusion process with generator $L$ and reversible measure $\mu$ is: there exist constants $\lambda > 0$, $b > 0$, a function $W \geq w > 0$, a function $V(x)$ such that $V$ goes to infinity at infinity, $|\nabla V(x)| \geq v > 0$ for $|x|$ large enough and such that $\mu(e^{aV}) < \infty$ verifying

$$LW(x) \leq -\lambda V(x)W(x) + b.$$

Fix $y$ and let $F(x,y) = \frac{1}{2}|\theta(x) - y|^2$. In order to establish a Logarithmic Sobolev inequality for $\mu^y$ we are going to show that the preceding condition holds for $V(x) = F(x,y)$ and $W(x) = e^{F(x,y)}$. We have

$$\begin{aligned}
\nabla_x F(x) &= D\theta(x)^T(\theta(x) - y),\\
\lambda_\theta|\theta(x) - y|^2 \leq |\nabla_x F|^2 &\leq \Lambda_\theta|\theta(x) - y|^2,\\
\Delta F &= n\bar{\lambda}_\theta + (\Delta\theta)^T(\theta - y).
\end{aligned}$$

Furthermore

$$\begin{aligned}
\varepsilon L^X F &= -\nabla_x V_0^T \nabla_x F - \kappa|\nabla_x F|^2 + \beta^{-1}\Delta F\\
&= -\nabla_x V_0^T D\theta^T(\theta - y) - \kappa|D\theta^T(\theta - y)|^2 + \beta^{-1}n\bar{\lambda}_\theta + \beta^{-1}\Delta\theta^T(\theta - y)\\
&\leq |\nabla_x V_0|\sqrt{\Lambda_\theta}|\theta - y| - \kappa\lambda_\theta|\theta - y|^2 + \beta^{-1}n\bar{\lambda}_\theta + \beta^{-1}|\Delta\theta||\theta - y|\\
&\leq -\kappa\lambda_\theta F + \frac{(|\nabla_x V_0|\sqrt{\Lambda_\theta} + \beta^{-1}|\Delta\theta|)^2}{2\kappa\lambda_\theta} + \beta^{-1}n\bar{\lambda}_\theta\\
&= -\kappa\lambda_\theta F + G(x)
\end{aligned}$$

where we used the fact that $-ax^2 + bx + c \leq -\frac{1}{2}ax^2 + \frac{b^2}{2a} + c$ for the second inequality.

Let $W(x,y) = e^{F(x,y)}$. Now,

$$\begin{aligned}
\varepsilon L^X W(x,y) &= \varepsilon L^X F(x,y)W(x,y) + \beta^{-1}|\nabla_x F(x,y)|^2 W(x,y)\\
&\leq -(\lambda_\theta\kappa - \Lambda_\theta\beta^{-1})F(x,y)W(x,y) + \|G\|_\infty W(x,y)\\
&= -((\lambda_\theta\kappa - \Lambda_\theta\beta^{-1})F(x,y) - \|G\|_\infty)W(x,y).
\end{aligned}$$

Since $F$ goes to infinity at infinity, for $x$ outside a compact set

$$-(\lambda_\theta\kappa - \Lambda_\theta\beta^{-1})F(x,y) + \|G\|_\infty \leq -\tfrac{1}{2}(\lambda_\theta\kappa - \Lambda_\theta\beta^{-1})F(x,y)$$

so that

$$\varepsilon L^X W(x,y) \le -\tfrac{1}{2}(\lambda_\theta \kappa - \Lambda_\theta \beta^{-1}) F(x,y) W(x,y) + K$$

for some constant $K$. This establishes a Log-Sobolev inequality for the measure $\mu^y$ with respect to $\varepsilon\Gamma^X$ in the sense that

$$\int f^2 \log f^2 d\mu^y \le 2c_L^y \int \varepsilon\Gamma^X d\mu^y$$

for some constant $c_L^y$ depending on $y$. Let $c_L = \sup_{y\in D} c_L^y$ so that

$$\int f^2 \log f^2 d\mu^y \le 2\varepsilon c_L \int \Gamma^X d\mu^y.$$

This shows that a Log-Sobolev inequality with a constant $\varepsilon c_L$ holds for each measure $\mu^y, y \in D$.

It remains to estimate $\kappa_X, \kappa_Y, \|\partial_{y_i} U\|^2_{\mathrm{Lip}}, \|\partial^2_{y_i} U\|^2_{\mathrm{Lip}}$ and $\bar{b}(y)^2$.

We have $b_X = -\varepsilon^{-1}\nabla_x V(x) - \varepsilon^{-1}\frac{\kappa}{2}\nabla_x|\theta(x) - y|^2$ and we want to find $\kappa_X$ such that

$$(x_1 - x_2)^T (b_X(x_1, y) - b_X(x_2, y)) \le -\kappa_X |x_1 - x_2|^2 + \alpha \text{ for all } x_1, x_2 \in \mathbb{R}^n, y \in \mathbb{R}^m.$$

Since $|\nabla_x V|$ is bounded and using the assumption on $\theta$, we get

$$\begin{aligned}
&(x_1 - x_2)^T (b_X(x_1, y) - b_X(x_2, y)) \\
&\quad = -\varepsilon^{-1}(x_1 - x_2)^T (\nabla_x V(x_1) - \nabla_x V(x_2)) - \varepsilon^{-1}\tfrac{\kappa}{2}(x_1 - x_2)^T (\nabla_x|\theta(x_1) - y|^2 - \nabla_x|\theta(x_2) - y|^2) \\
&\quad \le -\varepsilon^{-1}\tfrac{\kappa}{2}\kappa_\theta |x_1 - x_2|^2 + 2\varepsilon^{-1}|x_1 - x_2| \|\nabla_x V(x)\|_\infty + \varepsilon^{-1}\alpha_\theta \\
&\quad \le -\varepsilon^{-1}\frac{\kappa\kappa_\theta}{4}|x_1 - x_2|^2 + 4\varepsilon^{-1}\frac{\|\nabla_x V\|_\infty}{\kappa\kappa_\theta} + \varepsilon^{-1}\alpha_\theta
\end{aligned}$$

so that we can identify

$$\kappa_X = \varepsilon^{-1}\frac{\kappa\kappa_\theta}{4} \qquad\qquad \alpha = 4\varepsilon^{-1}\frac{\|\nabla_x V\|_\infty}{\kappa\kappa_\theta} + \varepsilon^{-1}\alpha_\theta.$$

We have

$$b_Y^i(x,y) = -\nabla_{y_i} U(x,y) = -\kappa(y_i - \theta_i(x))$$

so that

$$\nabla_x b_Y^i(x,y) = \kappa \nabla_x \theta_i(x)$$

and

$${\kappa_Y}^2 = \frac{1}{m}\sum_{i=1}^m \kappa^2 \|\nabla_x \theta_i(x)\|^2_\infty \le \kappa^2 \Lambda_\theta.$$

We also have

$$\|\partial_{y_i} U\|^2_{\mathrm{Lip}} = \|b_Y^i\|^2_{\mathrm{Lip}} \le \kappa^2 \|\nabla_x \theta_i\|^2_\infty \le \kappa^2 \Lambda_\theta$$

and

$$\|\partial_{y_i}^2 U\|_{\text{Lip}}^2 = \|\partial_{y_i}\kappa\theta(x)\|_{\text{Lip}}^2 = 0$$

so that

$$c_V{}^2 = \sup_y \left( \sum_{i=1}^m \|\partial_{y_i} U(\cdot, y)\|_{\text{Lip}}^2 + \sum_{i,j=1}^m \|\partial_{y_i y_j}^2 U(\cdot, y)\|_{\text{Lip}}^2 \right) \leq m\kappa^2 \Lambda_\theta.$$

From the expression for $\varepsilon L^X F$ we get that

$$F \leq -\frac{\varepsilon}{\kappa\lambda_\theta} L^X F + \frac{G(x)}{\kappa\lambda_\theta}.$$

Now

$$\begin{aligned}
\bar{b}^2(y) &= \left( \int -\kappa(y - \theta(x))\mu^y(dx) \right)^2 \\
&\leq \kappa^2 \int F(x,y)\mu^y(dx) \\
&\leq -\frac{\kappa\varepsilon}{\lambda_\theta} \int L^X F(x,y)\mu^y(dx) + \frac{\kappa}{\lambda_\theta} \int G(x)\mu^y(dx) \\
&= \frac{\kappa}{\lambda_\theta} \int G(x)\mu^y(dx)
\end{aligned}$$

since $\mu^y$ is invariant for $L^X(\cdot, y)$.

The separation of timescales is

$$\gamma = \frac{\kappa_X{}^2 \lambda_Y}{\Lambda_x \kappa_Y{}^2} \geq \varepsilon^{-1} \frac{\kappa_\theta{}^2 (\bar{\beta}\bar{\gamma})^{-1}}{16\Lambda_\theta(\beta^{-1})}.$$

If $\gamma > \frac{1}{(\sqrt{3}-\sqrt{2})^2}$ we can now apply Theorem 19 as in the previous section to show that an averaging principle holds for the stopped TAMD process with rate $\varepsilon^{1/2}$, i.e. there exists a constant $C$ depending on $T, V, \beta_X, b_Y$ and $\beta_Y$ such that for

$$\varepsilon \leq \frac{16(\sqrt{3} - \sqrt{2})^2 \Lambda_\theta \bar{\gamma} \beta^{-1}}{\kappa_\theta^2 \bar{b}^1}$$

we have

$$\mathbb{E} \sup_{0 \leq t \leq T} |Y_{t\wedge\tau} - \bar{Y}_{t\wedge\tau}| \leq \sqrt{\varepsilon} C.$$

## Acknowledgement

The author wants to thank Tony Lelièvre for suggesting the entropy approach in section 3.

## References

[CG17] Patrick Cattiaux and Arnaud Guillin. "Hitting times, functional inequalities, lyapunov conditions and uniform ergodicity". In: *Journal of Functional Analysis* 272.6 (2017), pp. 2361–2391.

[FW12] Mark I. Freidlin and Alexander D. Wentzell. *Random Perturbations of Dynamical Systems.* 3rd edition. Springer Berlin Heidelberg, 2012.

[HP86] U. G. Haussmann and E. Pardoux. "Time Reversal of Diffusions". In: *The Annals of Probability* 14.4 (1986), pp. 1188–1205.

[Liu10] Di Liu. "Strong convergence of principle of averaging for multiscale stochastic dynamical systems". In: *Commun. Math. Sci* 8.4 (2010), pp. 999–1020.

[LLO16] Frederic Legoll, Tony Lelièvre, and Stefano Olla. "Pathwise estimates for an effective dynamics". In: *arXiv preprint arXiv:1605.02644* (2016).

[MV06] Luca Maragliano and Eric Vanden-Eijnden. "A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations". In: *Chemical Physics Letters* 426.1 (July 2006), pp. 168–175.

[PV01] E. Pardoux and A. Yu Veretennikov. "On the Poisson equation and diffusion approximation 1". In: *Annals of probability* (2001), pp. 1061–1085.