PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 149, Number 8, August 2021, Pages 3583–3596
https://doi.org/10.1090/proc/15477
Article electronically published on May 18, 2021

QUADRATIC TRANSPORTATION INEQUALITIES FOR SDES WITH MEASURABLE DRIFT

KHALED BAHLALI, SOUFIANE MOUCHTABIH, AND LUDOVIC TANGPI

(Communicated by Zhen-Qing Chen)

Abstract. Let X be the solution of a stochastic differential equation in Euclidean space driven by standard Brownian motion, with measurable drift and Sobolev diffusion coefficient. In our main result we show that when the drift is measurable and the diffusion coefficient belongs to an appropriate Sobolev space, the law of X satisfies Talagrand's inequality with respect to the uniform distance.

1. Introduction and main results

1.1. Background. The concentration of measure phenomenon, initiated by [24], quantifies the deviation of a (Lipschitz continuous) function f of a random vector; $f(\xi^1, \dots, \xi^n)$ to its mean. It can be seen as a vast improvement of the classical Chebyshev inequality in large deviation theory; see e.g. the texts by [21] and [5] for modern presentations. [22] showed that the transportation inequality first established by [31] can be efficiently used to explain the concentration of measure phenomenon. The quadratic transportation inequality (or Talagrand inequality) compares the Wasserstein distance and the Kullback-Leibler divergence: Given a constant C, a probability measure μ is said to satisfy Talagrand's $T_2(C)$ inequality (or quadratic transportation inequality) if

$$\mathcal{W}_2(\mu,\nu) \le \sqrt{CH(\nu|\mu)} \text{ for all probability measures } \nu.$$

We define the (second order) Wassertein distance and the Kullback-Leibler divergence respectively by

$$\mathcal{W}_2(\mu,\nu) := \Big(\inf_\pi \int_{\Omega\times\Omega} \|\omega-\eta\|_\infty^2\,\pi(\mathrm{d}\omega,\mathrm{d}\eta)\Big)^{1/2} \quad \text{and} \quad H(\nu|\mu) := \int \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\log\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\mu$$

where the infimum is taken over all, couplings π of μ and ν, that is, probability measures on $\Omega\times\Omega$ such that if $(\zeta,\eta)\sim\pi$ then $\zeta\sim\mu$ and $\eta\sim\nu$, and we used the convention $\mathrm{d}\nu/\mathrm{d}\mu = +\infty$ if ν is not absolutely continuous w.r.t. μ. The transportation inequality has since found numerous applications, for instance to isoperimetric problems, to randomized algorithms [11], or to quantitative finance [19, 32] and to various problems of probability in high dimensions [9, 20, 23]. Transportation

Received by the editors March 26, 2020, and, in revised form, August 3, 2020, August 18, 2020, September 22, 2020, and October 24, 2020.

2020 *Mathematics Subject Classification.* Primary 60E15, 60H20, 60J60, 28C20.

Key words and phrases. Quadratic transportation inequality, stochastic differential equations, singular drifts, Sobolev regularity.

The second author was supported by PHC Toubkal 18/59. The third author was supported by NSF grant DMS-2005832.

inequalities are also related to other functional inequalities as Poincaré inequality, log-Sobolev inequality, inf-convolution and hypercontractivity; see [4, 25].

1.2. **Main result.** Our objective is to investigate transportation inequalities for stochastic differential equations (SDE) of the form

$$(1)\qquad X(t) = x + \int_0^t b(s, X(s))\mathrm{d}s + \int_0^t \sigma(s, X(s))\mathrm{d}B(s) \quad \text{for } t \in [0,T],\ x \in \mathbb{R}^d$$

under minimal regularity assumptions on the coefficients $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$. Note that this equation is understood in the $\mathbb{P}$-almost sure sense on the canonical probability space $(\Omega, \mathcal{F}, \mathbb{P})$ of the d-dimensional Brownian motion denoted by B and equipped with the $\mathbb{P}$-completion of the raw filtration $\sigma(B(s), s \leq t)$ generated by B. That is $\Omega = C([0,T], \mathbb{R}^d)$ endowed with the supremum norm, and $B(t,\omega) = \omega(t)$. Further denote by $\mathcal{P}(\Omega)$ the set of all Borel probability measures on Ω. In all of the article we assume that the diffusion coefficient σ satisfies the uniform ellipticity condition

$$(2)\qquad \xi^\top \sigma(t,x)\xi \geq \Lambda_\sigma |\xi|^2 \quad \text{for all} \quad (t,x,\xi) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^d,\ \text{for some } \Lambda_\sigma > 0.$$

To state our main result, let us define the following functional spaces. For $p \geq 1$ denote by $L^p_{\mathrm{loc}}([0,T]) := L^p_{\mathrm{loc}}([0,T] \times \mathbb{R}^d)$ the (Lebesgue) space of classes of locally integrable functions and for every $m_1, m_2 \in \mathbb{Z}_+$, let $W^{m_1,m_2}_p([0,T]) := W^{m_1,m_2}_p([0,T] \times \mathbb{R}^d)$ be the usual Sobolev space of weakly differentiable functions $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ such that

$$\|f\|_{W^{m_1,m_2}_p} := \sum_{i=0}^{m_1} \|\partial_t^i f\|_{L^p} + \sum_{|\alpha| \leq m_2} \|\partial_x^\alpha f\|_{L^p} < \infty$$

where α is a d-dimensional multiindex, i.e. a d-tuple $\alpha = (\alpha_1, \ldots, \alpha_d)$ and $|\alpha| := \alpha_1 + \cdots + \alpha_d$. Denote by $W^{m_1,m_2}_{p,\mathrm{loc}}([0,T])$ the space of weakly differentiable functions $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ such that

$$\|f\|_{L^p_{\mathrm{loc}}} + \sum_{|\alpha| \leq m_1} \|\partial_t^\alpha f\|_{L^p_{\mathrm{loc}}} + \sum_{|\alpha| \leq m_2} \|\partial_x^\alpha f\|_{L^p_{\mathrm{loc}}} < \infty.$$

Further let $L^p_q([0,T]) := L^q\big([0,T], L^p(\mathbb{R}^d)\big)$ be the space of (classes of) measurable functions $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\|f\|_{L^q_p} := \Big(\int_0^T \Big(\int_{\mathbb{R}^d} |f(s,x)|^p \mathrm{d}x\Big)^{q/p} \mathrm{d}s\Big)^{1/q} < \infty.$$

The aim of this note is to prove the following two results:

Theorem 1. *Let $\sigma \in W^{0,1}_{2(d+1),\mathrm{loc}}([0,T]) \cap L^\infty([0,T] \times \mathbb{R}^d)$ and $b \in W^{0,1}_{(d+1),\mathrm{loc}}([0,T]) \cap L^\infty([0,T] \times \mathbb{R}^d)$. Assume that σ satisfies (2). Then, equation (1) admits a unique strong solution X, and*

the law μ_x of X satisfies $T_2(C)$

with constant

$$C := \inf_{0<\varepsilon<1} 2\|\sigma\|^2_\infty \exp\Big(7T\frac{32+\varepsilon}{\varepsilon(1-\varepsilon)}\Big)\frac{1}{1-\varepsilon}.$$

This result gives the transportation inequality for SDEs with coefficients in some Sobolev spaces. It moreover constitute one essential argument in the proof of the next result where we combine it with some gradient estimates for solutions of partial differential equations (PDEs) established by [17] to establish transportation inequality for SDEs with measurable drifts.

Theorem 2. *Assume that σ satisfies* (2) *and that one of the following conditions is satisfied:*

(A) $\sigma, b \in L^\infty([0,T]\times\mathbb{R}^d)$, *the function σ is continuous in (t,x) and belongs to* $W^{0,1}_{2(d+1),\mathrm{loc}}([0,T])$.

(B) $\sigma \in W^{0,1}_{2(d+1),\mathrm{loc}}([0,T]) \cap L^\infty([0,T]\times\mathbb{R}^d)$, *$\sigma$ is uniformly continuous in x. The function b satisfies $b \in L^p_q([0,T])$ for some p,q such that $d/p+2/q<1$, $2(d+1)\le p$ and $q>2$.*

Then there exists $T>0$ small enough so that equation (1) *admits a unique strong solution X with continuous paths and*

$$\text{the law } \mu_x \text{ of } X \text{ satisfies } T_2(C)$$

for some constant C depending on the data, namely $\|b\|_{L^q_p}, \|\sigma\|_\infty, T, x, d, p$ and q.

Since b is merely assumed measurable, Theorem 2 gives transportation inequality for *singular* SDEs as $\mathrm{d}X(t) = \mathrm{sgn}(X(t))\mathrm{d}t + \mathrm{d}B(t)$, or for "regime switching" models as

$$\mathrm{d}X(t) = \big\{b_1(t,X(t))1_A(X(t)) + b_2(t,X(t))1_{A^c}(X(t)\big\}\mathrm{d}t + \sigma(t,X(t))\mathrm{d}B(t)$$

with A a measurable subset of $\mathbb{R}^d$. Other examples are discussed at the end of the article.

1.3. **Related literature.** [31] proved a quadratic transportation inequality for the multidimensional Gaussian distribution with optimal constant $C=2$. Using stochastic analysis techniques, notably Girsanov's theorem, Talagrand's work was then extended to Wiener measure on the path space by [14]. The case of SDEs was first analyzed by [10] using a technique based on Girsanov's transform that we also employ here. Their results gave rise to an interesting literature, including the papers [7, 26, 27, 33] on SDEs driven by Brownian motion and [29, 30] on SDEs driven by abstract Gaussian noise. Note that almost all the aforementioned works on SDEs assume that the coefficients are Lipschitz-continuous or satisfy a dissipative condition, with the exception of [7, 27] who assume a contraction condition of the form $(b(t,x)-b(t,y))\cdot(x-y)\le K\|x-y\|^2$. It is worth noting that [27] additionally deal with equations with reflections.

The extension of [10] to diffusions with non-smooth coefficients was started by [3] where it is proved that $T_2(C)$ holds for *one-dimensional* equations, if b is measurable in space and *differentiable* in time and σ *Lipschitz* continuous. The idea of [3] is based on a transformation that is tailor-made for the one-dimensional case. The present paper deals with the multidimensional case and further weakens the regularity requirements imposed in [3]. In this case we use Zvonkin's transformation which is well-known in SDE theory; see for instance [1, 34, 36]. Note that considering multidimensional equations is, for instance, fundamental for applications to concentration and asymptotic results on *interacting particle systems*; see e.g. [9, Section 5] and the various examples we give in the final section. The proof of Theorem 2 is given in the next section, and Section 3 presents some examples.

2. Proofs of the main results

We explain our strategy to prove Theorem 2. Roughly speaking, in both situations (A) and (B), the path remains the same and it consists in reducing equation (1) to one without drift. When assumption (A) is satisfied, we follow the path developed by [1] in the proof of strong uniqueness for SDEs with measurable drift and a locally Sobolev diffusion coefficient: We first derive the transportation inequality for SDEs with coefficients belonging to some Sobolev spaces, this is the goal of the next subsection. After that, we use Zvonkin's transformation [36] to deal with the case where the drift is only measurable. When assumption (B) is satisfied, i.e. the case where the drift belongs to some $L^p_q([0,T])$-space, we also remove the drift but the proof is more elaborated: We slightly modify the method of [17] in order to derive suitable gradient estimates for singular second order parabolic PDEs. Thus, along with gradient estimates for solutions of singular PDEs, Theorem 1 is an essential building block for the proof of the main result.

2.1. **Technical lemmas.** The aim of this section is to present two lemmas that will be used in the proof of Theorem 1. The proof of the first lemma can be found in Step 1 of the proof of [10, Theorem 5.6].

Lemma 3. *Let $\nu \in \mathcal{P}(\Omega)$ be such that $\nu \ll \mu_x$ and $H(\nu|\mu_x) < \infty$, and let X be the solution of* (1)*. Then, the probability measure $\mathbb{Q}$ given by*

$$\mathrm{d}\mathbb{Q} := \frac{\mathrm{d}\nu}{\mathrm{d}\mu_x}(X)\mathrm{d}\mathbb{P}$$

satisfies

$$H(\nu|\mu_x) = \mathbb{E}_{\mathbb{Q}}\Big[\frac{1}{2}\int_0^T |q(s)|^2\mathrm{d}s\Big] \tag{3}$$

for some progressively measurable, square integrable process q such that $\tilde{B} := B - \int_0^\cdot q(s)\mathrm{d}s$ is a $\mathbb{Q}$-Brownian motion.

Lemma 4 is a crucial point in the proof of Theorem 1, it provides the argument allowing to go around the uniform Lipschitz (or dissipativity) condition usually imposed on the coefficients. The main idea probably originated from [34] and further developed in [1] in order to establish the pathwise uniqueness of multidimensional SDEs when the diffusion coefficient is locally in some Sobolev space, that is when the usual Gronwall's lemma cannot be directly used.

Lemma 4. *Let X_1, X_2 be two square integrable semi-martingales of the form*

$$X_i(t) = x_i + \int_0^t \beta_i(u)\mathrm{d}u + \int_0^t \alpha_i(u)\mathrm{d}B(u), \quad i = 1,2$$

where β_i, α_i are bounded adapted processes with $\xi^\top \alpha_i(t)\xi \geq \Lambda|\xi|^2$ for all $(t,\xi) \in [0,T]\times\mathbb{R}^d$. For $N > 0$, we put

$$\tau^N := \inf\{t > 0 : |X_1(t)| > N \text{ or } |X_2(t)| > N\} \wedge T, \quad N \in \mathbb{Z}_+.$$

(i) *Let $f \in W^{0,1}_{2(d+1),loc}$ and $A^N_f(s)$ be the increasing process defined by*

$$A^N_f(t) := \int_0^{t\wedge\tau^N}\int_0^1 |\partial_x f(s, \lambda X_1(s) + (1-\lambda)X_2(s))|^2\mathrm{d}\lambda\mathrm{d}s.$$

Then, for every stopping time τ with values in $[0,T]$ it holds

$$\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N} |f(s,X_1(s))-f(s,X_2(s))|^2\mathrm{d}s\Big] \le 6\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N} |X_1(s)-X_2(s)|^2\mathrm{d}A_f^N(s)\Big]$$

(ii) Let $g\in W^{0,1}_{(d+1),loc}$ and $A_g^N(s)$ be defined as follows

$$A_g^N(t) := \int_0^{t\wedge\tau^N}\int_0^1 |\partial_x g(s,\lambda X_1(s)+(1-\lambda)X_2(s))|\mathrm{d}\lambda\mathrm{d}s.$$

Then, for every stopping time τ with values in $[0,T]$ it holds

$$\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N} |X_1(s)-X_2(s)||g(s,X_1(s))-g(s,X_2(s))|\mathrm{d}s\Big] \le \mathbb{E}\Big[\int_0^{\tau\wedge\tau^N} |X_1(s)-X_2(s)|^2\mathrm{d}A_g^N(s)\Big].$$

Proof. It must be noted that, in contrast to the Lipschitz case, the difference ratios of the functions f and g are not bounded in our situation, not even locally. We will see that, since the coefficients α_i of X_i are non degenerate, Krylov's estimate [18, Theorem 2.2.4] allows us to avoid this difficulty. Using Krylov's estimate, one can show that the processes $A_f^N(s)$ and $A_g^N(s)$ are well defined. We denote by K_N the ball $\{x\in\mathbb{R}^d, |x|\le N\}$ and put $K:=[0,T]\times K_N$. Let $f_n\in C^\infty(K)$ be such that,

$$\|f_n-f\|_{W^{0,1}_{2(d+1)}(K)}\to 0, \tag{4}$$

see e.g. [12, Theorem 5.3.2] for the existence of such a sequence. Using subsequently the classical inequality $(a+b+c)^2\le 3(a^2+b^2+c^2)$, Krylov's estimate and Taylor's formula with integral remainder, we have that

$$\begin{aligned}
I &:= \mathbb{E}\Big[\int_0^{\tau\wedge\tau^N} |f(s,X_1(s))-f(s,X_2(s))|^2\mathrm{d}s\Big]\\
&\le 3\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N} |f(s,X_1(s))-f_n(s,X_1(s))|^2\mathrm{d}s\Big]\\
&\quad + 3\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N} |f(s,X_2(s))-f_n(s,X_2(s))|^2\mathrm{d}s\Big]\\
&\quad + 3\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N} |f_n(s,X_1(s))-f_n(s,X_2(s))|^2\mathrm{d}s\Big]\\
&\le C_{T,N,d}\,\|f-f_n\|_{L^{2(d+1)}(K)} + 3\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N} |f_n(s,X_1(s))-f_n(s,X_2(s))|^2\mathrm{d}s\Big]\\
&\le C_{T,N,d}\|f-f_n\|_{L^{2(d+1)}(K)} + 3\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)|^2\int_0^1|\partial_x f_n(s,Z(s,\lambda))|^2\mathrm{d}\lambda\mathrm{d}s\Big]
\end{aligned} \tag{5}$$

with $Z(s,\lambda):=\lambda X_1(s)+(1-\lambda)X_2(s)$, and where $C_{T,N,d}$ is a positive constant which depends on T,N and d. Using Fubini's theorem and Krylov's estimate once again, we get

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)|^2\int_0^1|\partial_x f_n(s,Z(s,\lambda))|^2\mathrm{d}\lambda\mathrm{d}s\Big]\\
&\le \mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)|^2\int_0^1 2|\partial_x f_n(s,Z(s,\lambda))\\
&\qquad\qquad -\partial_x f(s,Z(s,\lambda))|^2+2|\partial_x f(s,Z(s,\lambda))|^2\mathrm{d}\lambda\mathrm{d}s\Big]\\
&\le 8N^2\int_0^1\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|\partial_x f_n(s,Z(s,\lambda))-\partial_x f(s,Z(s,\lambda))|^2\mathrm{d}s\Big]\mathrm{d}\lambda\\
&\quad+2\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)|^2\int_0^1|\partial_x f(s,Z(s,\lambda))|^2\mathrm{d}\lambda\mathrm{d}s\Big]\\
&\le 8N^2C_{T,N,d}\|\partial_x f_n-\partial_x f\|_{L^{2(d+1)}(K)}+2\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)|^2\mathrm{d}A_f^N(s)\Big].
\end{aligned}$$

Putting this together with the estimate (5) for I leads to

$$\begin{aligned}
I&\le C_{T,N,d}\|f_n-f\|_{L^{2(d+1)}(K)}+24N^2C_{T,N,d}\|\partial_x f_n-\partial_x f\|_{L^{2(d+1)}(K)}\\
&\quad+6\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)|^2\mathrm{d}A_f^N(s)\Big]\\
&\le C'_{T,N,d}\|f_n-f\|_{W^{0,1}_{2(d+1)}(K)}+6\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)|^2\mathrm{d}A_f^N(s)\Big]
\end{aligned}$$

for some constant $C'_{T,N,d}$. Hold N fixed and take the limit as n goes to infinity in the last inequality then use (4) to obtain the first claim.

The second inequality is obtained similarly. Indeed, let (g_n) be a sequence of functions such that $g_n\in C^\infty(K)$ and

$$\|g_n-g\|_{W^{0,1}_{(d+1)}(K)}\to 0.$$

We have

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)||g(s,X_1(s))-g(s,X_2(s))|\mathrm{d}s\Big]\\
&\le \mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)|\{|g(s,X_1(s))-g_n(s,X_1(s))|+|g_n(s,X_1(s))-g_n(s,X_2(s))|\\
&\quad+|g(s,X_2(s))-g_n(s,X_2(s))|\}\mathrm{d}s\Big]\\
&\le 2N^2C_{T,N,d}\|g-g_n\|_{L^{(d+1)}(K)}+\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)|^2\int_0^1|\partial_x g_n(s,Z(s,\lambda))|\mathrm{d}\lambda\mathrm{d}s\Big]\\
&\le 2N^2C_{T,N,d}\|g-g_n\|_{W^{0,1}_{(d+1)}(K)}+\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)|^2\int_0^1|\partial_x g(s,Z(s,\lambda))|\mathrm{d}\lambda\mathrm{d}s\Big]\\
&=2N^2C_{T,N,d}\|g-g_n\|_{W^{0,1}_{(d+1)}(K)}+\mathbb{E}\Big[\int_0^{\tau\wedge\tau^N}|X_1(s)-X_2(s)|^2\mathrm{d}A_g^N(s)\Big].
\end{aligned}$$

Again, hold N fixed and let n go to infinity to obtain the second claim. □

2.2. **Proof of Theorem 1.** That (1) admits a unique strong solution follows from e.g. [1, Theorem 2.1]. Let $\nu\in\mathcal{P}(\Omega)$ be absolutely continuous with respect to μ_x.

We can assume without loss of generality that $H(\nu|\mu_x) < \infty$. Let $\mathbb{Q}$ and q be as in Lemma 3. Under the probability measure $\mathbb{Q}$, the SDE (1) takes the form

$$\mathrm{d}X(t) = \sigma(t, X(t))\mathrm{d}\tilde{B}(t) + \big\{\sigma(t, X(t))q(t) + b(t, X(t))\big\}\mathrm{d}t, \quad \text{with} \quad X(0) = x \tag{6}$$

and the law of X under $\mathbb{Q}$ is ν. Furthermore, the SDE

$$\mathrm{d}Y(t) = \sigma(t, Y(t))\mathrm{d}\tilde{B}(t) + b(t, Y(t))\mathrm{d}t, \quad \text{with } Y(0) = x$$

admits a unique solution (see [1]) and the law of Y under $\mathbb{Q}$ is μ_x. That is, (X, Y) under $\mathbb{Q}$ is a coupling of (ν, μ_x). Thus,

$$\mathcal{W}_2^2(\nu, \mu_x) \le \mathbb{E}_{\mathbb{Q}}\Big[\sup_{0\le t\le T} |X(t) - Y(t)|^2\Big]. \tag{7}$$

We now estimate the right hand side above. By Itô's formula, we have

$$\begin{aligned}|X(t) - Y(t)|^2 = &\int_0^t 2(X(s) - Y(s))\sigma(s, X(s))q(s) + |\sigma(s, X(s)) - \sigma(s, Y(s))|^2\mathrm{d}s\\ &+ \int_0^t 2(X(s) - Y(s))(b(s, X(s)) - b(s, Y(s)))\mathrm{d}s\\ &+ \int_0^t 2(X(s) - Y(s))(\sigma(s, X(s)) - \sigma(s, Y(s)))\mathrm{d}\tilde{B}(s),\end{aligned} \tag{8}$$

where we simply denote by ab the inner product between two vectors a and b. The difficulty is to deal with the terms $\sigma(s, X(s))-\sigma(s, Y(s))$ and $b(s, X(s))-b(s, Y(s))$. This was done in Lemma 4. As in that lemma, we introduce the following random times: First consider the sequence of stopping times

$$\tau^N := \inf\{t > 0 : |X(t)| > N \text{ or } |Y(t)| > N\} \wedge T.$$

It is clear that $\tau^N \uparrow T$. For each λ in $[0, 1]$ and t in $[0, T]$, we put $Z(t, \lambda) := \lambda X(t) + (1 - \lambda)Y(t)$. For every $N \ge 0$ and $t \in [0, \infty)$, define

$$A^N(t) := \int_0^{t\wedge\tau^N}\int_0^1 \Big(|\partial_x\sigma(s, Z(s, \lambda))|^2 + |\partial_x b(s, Z(s, \lambda))|\Big)\mathrm{d}\lambda\mathrm{d}s \tag{9}$$

$$:= A_\sigma^N(t) + A_b^N(t), \tag{10}$$

where the (weak) derivative operator acts on the spacial variable. Notice that the function $t \mapsto A^N(t)$ is not necessarily strictly increasing. Thus, consider the process $k^N(t) := t + A^N(t)$. It is continuous, strictly increasing and satisfies $k^N(0) = 0$. Moreover, k^N maps $[0, \infty)$ onto itself. We denote by γ^N the unique inverse map of k^N. Observe that for each N, we have $\gamma_t^N \uparrow \infty$ as $t \uparrow \infty$. Using (8), Cauchy-Schwarz and Burkholder-Davis-Gundy inequalities, one can show that for each $t \in [0, \infty)$ it holds that

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\Big[\sup_{s\in[0,\gamma_t^N\wedge\tau^N]}|X(s)-Y(s)|^2\Big]\\
&\quad\leq \mathbb{E}_{\mathbb{Q}}\Big[\int_0^{\gamma_t^N\wedge\tau^N}|X(s)-Y(s)|^2\mathrm{d}s+\|\sigma\|_\infty^2\int_0^{\gamma_t^N\wedge\tau^N}|q(s)|^2\mathrm{d}s\Big]\\
&\qquad+\mathbb{E}_{\mathbb{Q}}\Big[\int_0^{\gamma_t^N\wedge\tau^N}|\sigma(s,X(s))-\sigma(s,Y(s))|^2\\
&\qquad\qquad\qquad\qquad+2|X(s)-Y(s)||b(s,X(s))-b(s,Y(s))|\mathrm{d}s\Big]\\
&\qquad+2C_{BDG}\mathbb{E}_{\mathbb{Q}}\Big[\Big(\int_0^{\gamma_t^N\wedge\tau^N}|X(s)-Y(s)|^2|\sigma(s,X(s))-\sigma(s,Y(s))|^2\mathrm{d}s\Big)^{1/2}\Big]\\
&\quad:= I_1+I_2+I_3, \qquad (11)
\end{aligned}
$$

for a (universal) constant $C_{BDG}>0$.

We shall estimate each of the terms I_2 and I_3 separately, since they contain the differences $\sigma(t,X(t))-\sigma(t,Y(t))$ and $b(t,X(t))-b(t,Y(t))$. To start with, notice that I_3 can be estimated in terms of I_2. In fact, letting $\varepsilon>0$, by the Young's inequality $2ab\leq\varepsilon a^2+\frac{1}{\varepsilon}b^2$, we have

$$
\begin{aligned}
I_3&\leq \varepsilon\mathbb{E}_{\mathbb{Q}}\Big[\sup_{s\in[0,\gamma_t^N\wedge\tau^N]}|X(s)-Y(s)|^2\Big]\\
&\quad+\frac{C_{BDG}^2}{\varepsilon}\mathbb{E}_{\mathbb{Q}}\Big[\int_0^{\gamma_t^N\wedge\tau^N}|\sigma(s,X(s))-\sigma(s,Y(s))|^2\mathrm{d}s\Big]\\
&\leq \varepsilon\mathbb{E}_{\mathbb{Q}}\Big[\sup_{s\in[0,\gamma_t^N\wedge\tau^N]}|X(s)-Y(s)|^2\Big]+\frac{C_{BDG}^2}{\varepsilon}I_2.
\end{aligned}
$$

We subsequently use Lemma 4 (with $f:=\sigma$ and $g:=b$) and identity (9) to get

$$
\begin{aligned}
I_2&\leq 6\mathbb{E}\Big[\int_0^{\gamma_t^N\wedge\tau^N}|X(s)-Y(s)|^2\mathrm{d}A_\sigma^N(s)+\int_0^{\gamma_t^N\wedge\tau^N}|X(s)-Y(s)|^2\mathrm{d}A_b^N(s)\Big]\\
&=6\mathbb{E}\Big[\int_0^{\gamma_t^N\wedge\tau^N}|X(s)-Y(s)|^2\mathrm{d}A^N(s)\Big].
\end{aligned}
$$

Coming back to (11), since $k^N(t):=t+A^N(t)$, we have

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\Big[\sup_{s\in[0,\gamma_t^N\wedge\tau^N]}|X(s)-Y(s)|^2\Big]\\
&\quad\leq \mathbb{E}_{\mathbb{Q}}\Big[\int_0^{\gamma_t^N\wedge\tau^N}|X(s)-Y(s)|^2\mathrm{d}k^N(s)+\|\sigma\|_\infty^2\int_0^{\gamma_t^N\wedge\tau^N}|q(s)|^2\mathrm{d}s\Big]\\
&\qquad+\varepsilon\mathbb{E}_{\mathbb{Q}}\Big[\sup_{s\in[0,\gamma_t^N\wedge\tau^N]}|X(s)-Y(s)|^2\Big]\\
&\qquad+6\frac{C_{BDG}^2+\varepsilon}{\varepsilon}\mathbb{E}_{\mathbb{Q}}\Big[\int_0^{\gamma_t^N\wedge\tau^N}|X(s)-Y(s)|^2\mathrm{d}k^N(s)\Big].
\end{aligned}
$$

The time change $s \equiv \gamma^N_r$ (which is equivalent to $r = k^N(s)$) gives

$$(1-\varepsilon)\mathbb{E}_{\mathbb{Q}}\Big[\sup_{s\in[0,\gamma^N_t\wedge\tau^N]}|X(s)-Y(s)|^2\Big] \le \|\sigma\|^2_\infty\mathbb{E}_{\mathbb{Q}}\Big[\int_0^{\gamma^N_t\wedge\tau^N}|q(s)|^2\mathrm{d}s\Big] + 7\frac{C^2_{BDG}+\varepsilon}{\varepsilon}\mathbb{E}_{\mathbb{Q}}\Big[\int_0^{t\wedge\tau^N}|X(\gamma^N_r)-Y(\gamma^N_r)|^2\mathrm{d}r\Big]$$

and using the fact that the function γ^N is increasing, we then have

$$\begin{aligned}(1-\varepsilon)\mathbb{E}_{\mathbb{Q}}\Big[&\sup_{s\in[0,\gamma^N_t\wedge\tau^N]}|X(s)-Y(s)|^2\Big]\\ &\le \|\sigma\|^2_\infty\mathbb{E}_{\mathbb{Q}}\Big[\int_0^{\gamma^N_t\wedge\tau^N}|q(s)|^2\mathrm{d}s\Big]\\ &\quad + 7\frac{C^2_{BDG}+\varepsilon}{\varepsilon}\mathbb{E}_{\mathbb{Q}}\Big[\int_0^{t}\sup_{s\in[0,\gamma^N_r\wedge\tau^N]}|X(s)-Y(s)|^2\mathrm{d}r\Big].\end{aligned}$$

Choosing $\varepsilon < 1$ and using Gronwall's lemma, we get

$$\mathbb{E}_{\mathbb{Q}}\Big[\sup_{s\in[0,\gamma^N_t\wedge\tau^N]}|X(s)-Y(s)|^2\Big] \le \frac{1}{1-\varepsilon}\|\sigma\|^2_\infty\mathbb{E}_{\mathbb{Q}}\Big[\int_0^T|q(s)|^2\mathrm{d}s\Big]\exp\Big(7\frac{C^2_{BDG}+\varepsilon}{\varepsilon(1-\varepsilon)}T\Big)$$

where we also used the fact that $\tau^N \wedge \gamma^N_t \le T$. Letting successively t then N go to infinity, it follows by Fatou's lemma, $\gamma^N_t \uparrow \infty$, $\tau^N \uparrow T$ and the continuity of X and Y that

$$\mathbb{E}_{\mathbb{Q}}\Big[\sup_{s\in[0,T]}|X(s)-Y(s)|^2\Big] \le \exp\Big(7\frac{C^2_{BDG}+\varepsilon}{\varepsilon(1-\varepsilon)}T\Big)\frac{1}{1-\varepsilon}\|\sigma\|^2_\infty\mathbb{E}_{\mathbb{Q}}\Big[\int_0^T|q(s)|^2\mathrm{d}s\Big].$$

Hence, we conclude from (3) and (7) that

$$\mathcal{W}^2_2(\mu_x,\nu) \le 2\exp\Big(7\frac{C^2_{BDG}+\varepsilon}{\varepsilon(1-\varepsilon)}T\Big)\frac{1}{1-\varepsilon}\|\sigma\|^2_\infty H(\nu|\mu_x).$$

A close reading of the proof of Burkholder-Davis-Gundy inequality (see e.g. [15, Theorem 3.3.1]) shows that $C_{BDG} = 4\sqrt{2}$. This concludes the proof. □

2.3. **Proof of Theorem 2.**

(A): We start by the case when assumption (A) is satisfied. By [1, Theorem 3.1] equation (1) admits a unique strong solution. As in [1], the idea consists in using Zvonkin's transform in order to transform equation (1) into an SDE without drift then using Theorem 1 to conclude. In the rest of the paper, we denote by $\mathcal{L}$ the differential operator defined by

$$\mathcal{L} = \sum_{i=1}^d b_i\frac{\partial\phi}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^d a_{ij}\frac{\partial^2\phi}{\partial x_i\partial x_j},$$

where $(a_{ij})_{i,j=1,\dots,d}$ is the matrix $\sigma^\top\sigma$. According to [36, Theorem 2], there exists a $T > 0$ small enough such that the PDE

$$\begin{cases}\partial_t\varphi + \mathcal{L}\varphi = 0\\ \varphi(T,x) = x\end{cases}$$

admits a unique solution $\varphi \in W^{1,2}_{p,\mathrm{loc}}([0,T])$ such that: for every t, the function $x \mapsto \varphi(t,x)$ is one-to-one from $\mathbb{R}^d$ onto $\mathbb{R}^d$, both φ and its inverse ψ belong to

$W^{1,2}_{p,\mathrm{loc}}([0,T])$ for each $p>1$, both $\varphi(t,\cdot)$ as well as its inverse $\psi(t,\cdot)$ are Lipschitz continuous, with Lipschitz constants depending on $d, T, \|b\|_\infty$ and $\|\sigma\|_\infty$.

Applying Itô-Krylov's formula, see [18, Theorem 2.10.1] to $\varphi(t,X(t)) := Y(t)$, it follows that Y satisfies the drift-less SDE

$$Y(t) = Y(0) + \int_0^t \tilde{\sigma}(s,Y(s))\mathrm{d}B(s)$$

with $\tilde{\sigma}(t,y) := (\sigma^\top \partial_x\varphi)(t,\psi(t,y))$. Since σ belongs to $W^{0,1}_{2(d+1),\mathrm{loc}}([0,T])$, and φ belongs to $W^{1,2}_{p,\mathrm{loc}}([0,T])$ for each $p>1$ and both φ and ψ are Lipschitz, it follows that $\tilde{\sigma} \in W^{0,1}_{2(d+1),\mathrm{loc}}([0,T])$, more details can be found in [1, page 14, proof of Theorem 3.1]. Hence, by Lemma 1, the law μ_y of Y satisfies $T_2(C)$, where C is the constant in Theorem 1. But $X(t) = \psi(t,Y(t))$ and ψ is Lipschitz continuous. Thus, the result follows from [10, Lemma 2.1].

(B): We now assume that condition (B) is fulfilled. We need to introduce the following Banach spaces: For every $k \geq 0$ and $m \geq 1$, let $H^k_m := (I-\Delta)^{-k/2}L^m$ be the usual space of Bessel potentials on $\mathbb{R}^d$ and denote

$\mathbb{H}^{2,q}_p([0,T]) := L^q([0,T],H^2_p)$ and $H^{2,q}_p([0,T]) := \{u : [0,T] \to H^2_p$ and $\partial_t u \in L^q_p([0,T])\}$.

The space H^2_p is equipped with the norm

$$\|u\|_{H^2_p} := \|(I-\Delta)u\|_{L^p}$$

making it isomorphic to the Sobolev space $W^2_p(\mathbb{R}^d)$.

Under assumption (B), the existence and uniqueness of X follow e.g. from [35, Theorem 1.1]. We now show that the law μ_x of X satisfies $T_2(C)$ for some $C>0$. By [17, Theorem 10.3 and Remark 10.4], the PDE

$$\begin{cases} \partial_t u_i + \mathcal{L}u_i + b_i = 0 \\ u_i(T,x) = 0 \end{cases}$$

admits a unique solution $u_i \in H^{2,q}_p([0,T])$ and this solution satisfies

$$\|\partial_t u_i\|_{L^q_p} + \|u_i\|_{\mathbb{H}^{2,q}_p([0,T])} \leq C_1\|b_i\|_{L^q_p}$$

for some constant C_1 depending on d,p,q,T and $\|b\|_{L^q_p}$. Furthermore, since $d/p + 2/q < 1$, it follows by [17, Lemma 10.2] that

$$|\partial_x u_i| \leq C_2 T^{\varepsilon/2}\|u_i\|^{1-1/q-\varepsilon/2}_{\mathbb{H}^{2,q}_p([0,T])}\|\partial_t u_i\|^{1/q+\varepsilon/2}_{L^q_p}$$

with for $\varepsilon \in (0,1)$ such that $\varepsilon + d/p + 2/q < 1$ and C_2 a constant depending on p,q and ε. Therefore, it follows that

$$|\partial_x u_i| \leq C_1C_2T^{\varepsilon/2}\|b_i\|_{L^q_p}, \tag{12}$$

so that choosing T small enough, we have $|\partial_x u| < \frac{1}{2d}$. Now consider the function $\phi_i(t,x) := X_i + u_i(t,x)$, $i = 1,\dots,d$. It is easily checked that the function ϕ_i solves the PDE

$$\begin{cases} \partial_t\phi_i + \mathcal{L}\phi_i = 0 \\ \phi_i(T,x) = X_i. \end{cases} \tag{13}$$

Put $\phi(t,x) = (\phi_1(t,x),\dots,\phi_d(t,x))$. Due to (12), it holds that

$$\big(1-2d|\partial_x u|^2\big)|x-y|^2 \le 2|\phi(t,x)-\phi(t,y)|^2 \le 2\big(2+|\partial_x u|^2\big)|x-y|^2 \quad \text{for all } x,y\in\mathbb{R}^d.$$

As a consequence, ϕ is one-to-one (see e.g. the corollary on page 87 of [16]), and its inverse $\psi := \phi^{-1}$ is Lipschitz continuous.

Since for every t, $u(t,\cdot)$ belongs to H^2_p, then it can be seen as an element of $W^2_p(\mathbb{R}^d)$. Moreover, the derivative of u with respect to t belongs to L^p, it thus follows that u belongs to $W^{1,2}_p([0,T])$. Hence, the function $\phi(t,x) := x + u(t,x)$ belongs to $W^{1,2}_{p,loc}([0,T])$. Itô-Krylov's formula (see [18, Theorem 2.10.1]) applied to ϕ gives

$$\begin{aligned} Y(t) := \phi(t,X(t)) &= \phi(0,x) + \int_0^t (\partial_t\phi + \mathcal{L}\phi)(s,X(s))\mathrm{d}s + \int_0^t \partial_x\phi(s,X(s))\sigma \mathrm{d}B(s)\\ &= \phi(0,x) + \int_0^t \tilde\sigma(s,Y(s))\mathrm{d}B(s) \end{aligned}$$

with $\tilde\sigma(t,y) := (\sigma^\top \partial_x\phi)(t,\psi(t,y))$, and where the second equation follows by (13). The rest of the proof follows as in the case of assumption (A). □

Remark 5. The estimate (12) in the proof of Theorem 2 allows to see that when T is arbitrary, the conclusion of Theorem 2 still holds provided that $\|b\|_{L^q_p}$ is small enough.

3. Examples

Let us now present a few examples of multidimensional diffusion models with non-Lipschitz coefficients which fit to our framework.

3.1. **Particles interacting through their rank.** Let $B_1,\dots B_n$ be n independent Brownian motions. Rank-based interaction models are given by

$$\mathrm{d}X^{i,n}(t) = \sum_{j=1}^n \delta_j 1_{\{X^{i,n}(t)=X^{(j),n}(t)\}}\mathrm{d}t + \sigma^i(t)\mathrm{d}B_i(t) \quad X^{i,n}(0) = X_i$$

for some real numbers δ_j, some measurable, bounded functions σ^i, with $X^{(1),n}(t) \le X^{(2),n}(t) \le \dots \le X^{(n),n}(t)$ is the system in increasing order. More generally, this model can be written as

$$\mathrm{d}X^{i,n}(t) = b\Big(\frac{1}{n}\sum_{j=1}^n 1_{\{X^{n,j}(t)\le X^{n,i}(t)\}}\Big)\mathrm{d}t + \sigma^i(t)\mathrm{d}B_i(t) \quad X^{i,n}(0) = X_i$$

for a given (deterministic) functions b. This model was introduced by [13] in the context of stochastic portfolio theory. Concentration of measures results for such systems can be found in [28]. Assume that $0 < c \le \inf_{i,t}|\sigma^i(t)| \le \sup_{i,t}|\sigma^i(t)| \le C$ for some c, C and that $b\in L^\infty$ (respectively $b \in L^p(\mathbb{R}, dx)$ with appropriate p, d). Then Theorem 2 (A) (respectively Theorem 2 (B)) shows that the law of $(X^{1,n},\dots,X^{n,n})$ satisfies $T_2(C)$ for some $C > 0$, provided that T is sufficiently small or, as explained in Remark 5 provided that the norm of b is smalll enough. This result is for instance valid for the so-called (finite) Atlas model of [2] given by

$$\mathrm{d}X^{i,n}(t) = \sum_{j=1}^n \delta 1_{\{X^{i,n}(t)=X^{p_i,n}(t)\}}\mathrm{d}t + \sigma^i(t)\mathrm{d}B_i(t) \quad X^{i,n}(0) = X_i,$$

for some constant δ and a permutation $(p_1,\dots,p_n)$ of $(1,\dots,n)$.

3.2. **Particles in quantile interaction.** Quantile interaction models are given by

$$\mathrm{d}X^{i,n}(t) = b(t, X^{n,i}(t), V^{\alpha,n}(t))\mathrm{d}t + \sigma(t, X^{n,i}(t))\mathrm{d}B_i(t) \quad X^{i,n}(0) = X_i,$$

where $V^{\alpha,n}(t)$ is the quantile at level $\alpha \in [0,1]$ of the empirical measure of the system $(X^{1,n}(t), \dots, X^{n,n}(t))$. That is,

$$V^{\alpha,n}(t) := \inf\Big\{u \in \mathbb{R} : \frac{1}{n}\sum_{i=1}^n 1_{\{X^{i,n}(t)\le u\}} \ge \alpha\Big\}.$$

This model is considered for instance in [8] in connection to exchangeable particle systems. Theorem 2 can be applied to this case under integrability conditions on b and mild regularity conditions σ.

3.3. **Brownian motion with random drift.** In addition to particle systems, our main result can also allow to derive transportation inequalities for semimartingales. We illustrate this in Corollary 6. Let g be a progressive stochastic process. We call Brownian motion with drift the process

$$X(t) = x + \int_0^t g(s)\mathrm{d}s + \sigma \mathrm{d}B(t). \tag{14}$$

We have Corollary 6 of Theorem 2:

Corollary 6. *Assume that the constant matrix σ satisfies* (2). *If the drift g is bounded and T small enough, then the law μ_t^x of $X(t)$ given by* (14) *satisfies $T_2(C)$ for some $C > 0$ depending on T, σ, d and $\|g\|_\infty$.*

Proof. Consider the Borel measurable function

$$b(t,x) := \mathbb{E}[g(t)|X(t) = x].$$

By [6, Corollary 3.7], we have $\mu_t^x = \tilde{\mu}_t$, where $\tilde{\mu}_t$ is the law of the weak solution $\tilde{X}(t)$ of the SDE

$$\tilde{X}(t) = x + \int_0^t b(s, \tilde{X}(s))\mathrm{d}s + \sigma \mathrm{d}B(s). \tag{15}$$

Since g is bounded so is the function b. Thus, the SDE (15) admits a unique strong solution; see e.g. [1, 34]. Thus, $\tilde{X}$ is necessarily a strong solution and by Theorem 2 $\tilde{\mu}$ satisfies $T_2(C)$, which concludes the argument. □

Acknowledgment

The authors are grateful to an anonymous referee for a careful reading and helpful suggestions.

References

[1] Khaled Bahlali, *Flows of homeomorphisms of stochastic differential equations with measurable drift*, Stochastics Stochastics Rep. **67** (1999), no. 1-2, 53–82, DOI 10.1080/17442509908834203. MR1717807

[2] Adrian D. Banner, Robert Fernholz, and Ioannis Karatzas, *Atlas models of equity markets*, Ann. Appl. Probab. **15** (2005), no. 4, 2296–2330, DOI 10.1214/105051605000000449. MR2187296

[3] Daniel Bartl and Ludovic Tangpi, *Functional inequalities for forward and backward diffusions*, Electron. J. Probab. **25** (2020), Paper No. 94, 22, DOI 10.1214/20-ejp495. MR4136474

[4] Sergey G. Bobkov, Ivan Gentil, and Michel Ledoux, *Hypercontractivity of Hamilton-Jacobi equations*, J. Math. Pures Appl. (9) **80** (2001), no. 7, 669–696, DOI 10.1016/S0021-7824(01)01208-9. MR1846020

[5] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration inequalities*, Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence; With a foreword by Michel Ledoux. MR3185193

[6] Gerard Brunick and Steven Shreve, *Mimicking an Itô process by a solution of a stochastic differential equation*, Ann. Appl. Probab. **23** (2013), no. 4, 1584–1628, DOI 10.1214/12-aap881. MR3098443

[7] P. Cattiaux and A. Guillin, *Semi log-concave Markov diffusions*, Séminaire de Probabilités XLVI, Lecture Notes in Math., vol. 2123, Springer, Cham, 2014, pp. 231–292, DOI 10.1007/978-3-319-11970-0_9. MR3330820

[8] Dan Crisan, Thomas G. Kurtz, and Yoonjung Lee, *Conditional distributions, exchangeable particle systems, and stochastic partial differential equations* (English, with English and French summaries), Ann. Inst. Henri Poincaré Probab. Stat. **50** (2014), no. 3, 946–974, DOI 10.1214/13-AIHP543. MR3224295

[9] François Delarue, Daniel Lacker, and Kavita Ramanan, *From the master equation to mean field game limit theory: large deviations and concentration of measure*, Ann. Probab. **48** (2020), no. 1, 211–263, DOI 10.1214/19-AOP1359. MR4079435

[10] H. Djellout, A. Guillin, and L. Wu, *Transportation cost-information inequalities and applications to random dynamical systems and diffusions*, Ann. Probab. **32** (2004), no. 3B, 2702–2732, DOI 10.1214/009117904000000531. MR2078555

[11] Devdatt P. Dubhashi and Alessandro Panconesi, *Concentration of measure for the analysis of randomized algorithms*, Cambridge University Press, Cambridge, 2009. MR2547432

[12] Lawrence C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998. MR1625845

[13] E. Robert Fernholz, *Stochastic portfolio theory*, Applications of Mathematics (New York), vol. 48, Springer-Verlag, New York, 2002. Stochastic Modelling and Applied Probability. MR1894767

[14] D. Feyel and A. S. Üstünel, *Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space*, Probab. Theory Related Fields **128** (2004), no. 3, 347–385, DOI 10.1007/s00440-003-0307-x. MR2036490

[15] Nobuyuki Ikeda and Shinzo Watanabe, *Stochastic differential equations and diffusion processes*, 2nd ed., North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989. MR1011252

[16] Fritz John, *On quasi-isometric mappings. I*, Comm. Pure Appl. Math. **21** (1968), 77–110, DOI 10.1002/cpa.3160210107. MR222666

[17] N. V. Krylov and M. Röckner, *Strong solutions of stochastic equations with singular time dependent drift*, Probab. Theory Related Fields **131** (2005), no. 2, 154–196, DOI 10.1007/s00440-004-0361-z. MR2117951

[18] N. V. Krylov, *Controlled diffusion processes*, Stochastic Modelling and Applied Probability, vol. 14, Springer-Verlag, Berlin, 2009. Translated from the 1977 Russian original by A. B. Aries; Reprint of the 1980 edition. MR2723141

[19] Daniel Lacker, *Liquidity, risk measures, and concentration of measure*, Math. Oper. Res. **43** (2018), no. 3, 813–837, DOI 10.1287/moor.2017.0885. MR3846074

[20] Laurière, M. and L. Tangpi. *Backward propagation of chaos*, Preprint, 2019.

[21] Michel Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs, vol. 89, American Mathematical Society, Providence, RI, 2001. MR1849347

[22] K. Marton, *Bounding $\overline{d}$-distance by informational divergence: a method to prove measure concentration*, Ann. Probab. **24** (1996), no. 2, 857–866, DOI 10.1214/aop/1039639365. MR1404531

[23] Pascal Massart, *Concentration inequalities and model selection*, Lecture Notes in Mathematics, vol. 1896, Springer, Berlin, 2007. Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, July 6–23, 2003; With a foreword by Jean Picard. MR2319879

[24] V. D. Milman, *Asymptotic properties of functions of several variables that are defined on homogeneous spaces* (Russian), Dokl. Akad. Nauk SSSR **199** (1971), 1247–1250; English transl., Soviet Math. Dokl. **12** (1971), 1277–1281. MR0303566

[25] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400, DOI 10.1006/jfan.1999.3557. MR1760620

[26] Soumik Pal, *Concentration for multidimensional diffusions and their boundary local times*, Probab. Theory Related Fields **154** (2012), no. 1-2, 225–254, DOI 10.1007/s00440-011-0368-1. MR2981423

[27] Soumik Pal and Andrey Sarantsev, *A note on transportation cost inequalities for diffusions with reflections*, Electron. Commun. Probab. **24** (2019), Paper No. 21, 11, DOI 10.1214/19-ECP223. MR3940196

[28] Soumik Pal and Mykhaylo Shkolnikov, *Concentration of measure for Brownian particle systems interacting through their ranks*, Ann. Appl. Probab. **24** (2014), no. 4, 1482–1508, DOI 10.1214/13-AAP954. MR3211002

[29] Sebastian Riedel, *Transportation-cost inequalities for diffusions driven by Gaussian processes*, Electron. J. Probab. **22** (2017), Paper No. 24, 26, DOI 10.1214/17-EJP40. MR3622894

[30] Bruno Saussereau, *Transportation inequalities for stochastic differential equations driven by a fractional Brownian motion*, Bernoulli **18** (2012), no. 1, 1–23, DOI 10.3150/10-BEJ324. MR2888696

[31] M. Talagrand, *Transportation cost for Gaussian and other product measures*, Geom. Funct. Anal. **6** (1996), no. 3, 587–600, DOI 10.1007/BF02249265. MR1392331

[32] Ludovic Tangpi, *Concentration of dynamic risk measures in a Brownian filtration*, Stochastic Process. Appl. **129** (2019), no. 5, 1477–1491, DOI 10.1016/j.spa.2018.05.008. MR3944773

[33] Ali Suleyman Üstünel, *Transportation cost inequalities for diffusions under uniform distance*, Stochastic analysis and related topics, Springer Proc. Math. Stat., vol. 22, Springer, Heidelberg, 2012, pp. 203–214, DOI 10.1007/978-3-642-29982-7_9. MR3236093

[34] A. Ju. Veretennikov, *Strong solutions and explicit formulas for solutions of stochastic integral equations* (Russian), Mat. Sb. (N.S.) **111(153)** (1980), no. 3, 434–452, 480. MR568986

[35] Xicheng Zhang, *Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients*, Electron. J. Probab. **16** (2011), no. 38, 1096–1116, DOI 10.1214/EJP.v16-887. MR2820071

[36] A. K. Zvonkin, *A transformation of the phase space of a diffusion process that will remove the drift* (Russian), Mat. Sb. (N.S.) **93(135)** (1974), 129–149, 152. MR0336813

IMATH, Université de Toulon, EA 2134, 83957 La Garde Cedex, France
Email address: khaled.bahlali@univ-tln.fr

LIBMA, Department of Mathematics, Faculty of Sciences Semlalia, Cadi Ayyad University, 2390 Marrakesh, Morocco; and IMATH, Université de Toulon, EA 2134, 83957 La Garde Cedex, France
Email address: soufiane.mouchtabih@gmail.com

Department of Operations Research and Financial Engineering, Princeton University, Princeton, New Jersey, 08540
Email address: ludovic.tangpi@princeton.edu