Wasserstin-2 Analysis of the non-symmetric Fokker–Planck equation and the Trotter Product Formula

Igor Stojković
Matematisch Instituut
Leiden University
The Netherlands
e-mail: stojkov@math.leidenuniv.nl

October 5, 2010

Abstract

We study the nonsymmetric Fokker-Plamck equation from the point of view of the gradient flows on the Wasserstein-2 space of probability measures on $\mathbb{R}^d$. Even though this equation is not a gradient flow, we do show that under suitable conditions it does have properties rather analogous to those that flows induceed by maximal monotone operators on Hilbert spaces have. Besides that that we also show that the flow has a regularizing property, i.e. the semigroup is differentiable. One of the main tools that we use is the Trotter-Kato product formula for our equation with recpect to the Wasserstein-2 metric, which we prove in Section 3. This result is by itself new even if we consider the symmetric Fokker-Planck equation.

1 Introduction

In this monograph we study the Fokker-Planck equation on $\mathbb{R}^d$.

$$\partial_t \mu_t = \Delta \mu_t + \nabla \cdot (b\mu_t) \quad \mu_0 \in \mathcal{P}(\mathbb{R}^d) \tag{1.1}$$

in the sense of distributions on $\mathcal{D}((0,\infty) \times \mathbb{R}^d$ where the solution $t \mapsto \mu_t$ assumes values in the space $\mathcal{P}(\mathbb{R}^d)$ of Borel probability measures on $\mathbb{R}^d$ and is weakly continuous. The drift coefficient b of (1) is assumed to satisfy the following three conditions: it is locally Lipschits, i.e.

$$\forall R > 0 \ \ \exists c_R > 0 \text{ such that } |b(x) - b(y)| \leqslant c_R |x - y| \text{ if } |x|, |y| \leqslant R \tag{1.2}$$

it satisfies the linear growth condition

$$\exists c > 0 \text{ such that } |b(x)| \leqslant c(1 + |x|) \quad \text{ for } x \in \mathbb{R}^d \tag{1.3}$$

and is moreover monotone, i.e.

$$\langle b(x) - b(y), x - y \rangle \geqslant 0 \quad \text{for } x, y \in \mathbb{R}^d \tag{1.4}$$

Equation (1.1) is the Kolmogorov forward equation of the associated stochastic differential equation (SDE)

$$dX_t = -b(X_t)dt + \sqrt{2}dW_t \tag{1.5}$$

where $(W_t)_{t\geqslant 0}$ in a standard Brownian Motion with values in $\mathbb{R}^d$. For equations of this type the strong existence of uniqueness hold provided (1.2) and (1.3) hold. Moreover the equation (1.5) induces a family of Markov transition kernels, and the paths of the dual semigroup (i.e. the action of transition kernels on measures) are the solutions of (1.1) for any initial $\rho_0 \in \mathcal{P}(\mathbb{R}^d)$ (interested reader may consult monographs [11], [20], [21] and also [9] for theory on infinite dimensional spaces. Moreover [19] considers stochastic equations as (1.5) where the drift is a multyvalued maximal monotone operator).

Equations (1.1) and (1.5) have been studied extensively in the past, and even more recently, several papers have appeared on this subject (see [5] for existance, [4] for uniqueness and [3]) for the regularity under rather general growth conditions. In these monographs authors consider the second order part of the differential operator to be eliptic with coefficients depending on x, and b to be just locally bounded, and in particular in [3] authors show that any solution μ_t to (1.1) is $\mathcal{L}^1$ a.e. absolutely continuous w.r.t the Lebegue measure $\mathcal{L}^d$. Although such setting is more general than the setting in curent work, we will show that our setting is a natural one to study equation (1.1).

Furthermore in [14] authors investigate (1.1) from the point of view of Optimal Transportation and prove contractivness of the flow of solutions with respect to a class of optimal transportation distances which contain all the Wasserstin-p for $p \in [1, \infty)$ distances. Such point of view has been very successfully is investigating equations like (1.1) when the drift b coefficient is a gradient of a lower semicontinuous convex function $V : \mathbb{R}^d \to (-\infty, \infty]$, or even if when V is not convex. Seminal papers [12] and [17] have provided a foundation for a novel branch of variational analysis which has inspired many mathematicians to do their own investigations and subsequently many papers and books about these and other related topics have appeared during the past decade (for example [1], [23], [24]). We adopt this point of view and analyse equation (1.1) as an equation on the metric space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ of probability measures on $\mathbb{R}^d$ equiped with the Wasserstin-2 distance.

The existing theory about gradient flows on the Wasserstein-2 space does not suffice to yield any claim about equation (1.1), since if b is not a gradient (of a convex function on $\mathbb{R}$)) there is no reason why solutions of (1.1) should be gradient flow curves on $(\mathcal{P}_2, W_2)$ in the sense of Ambrosio–Gigli-Savaré (see [1] and also[2] for the study of the symmetric Fokker-Planck equation in the infinite dimensional setting)-as a matter a fact the author of this thesis would be much less surprissed to read a proof of the reverse claim. However in light of the continuity equations ([1] ch 8), the Wasserstein "generator" of (1.1) can be seen as a sum of the subdifferential of the Heat Entropy functional and the mapping b which is monotone along any couppling plan. These observations point out that the Wasserstein–2 space should be a natural ambient for the Fokker–Planck equation (1.1). In fact we are going to prove that solutions of equation(1.1) have the properties that we expect, i.e. just as in the case of gradient flows,there is a strong analogy between the properties of a flow induceed by a maximal monotone operator on a Hilbert space and the properties that solutions of our Non-symmetric Fokker-Planck equation have as curves in the wasserstein-2 space. Moreover we will show that the Trotter-Kato Product formula holds for (1.1)-this result is new and seems intersting by itself (but see ...for a related though less general result where authors show that this formula holds for the symmetric Foker-Planck equation, when approximation steps are both taken by a resolvent and the drift coefficient is a gradient of a C^2 convex function that has bouneded Laplacian). Next to this we will show that if the drift coefficient b is Lipschitz, the induceed semigroup exibites the regularizing effect. The fact that our solutions posses all of the above mentioned properties, indeed conform strongly that $(\mathcal{P}_2, W_2)$ is a natural ambient space for the Non-symmetric Fokker-Planck equation in authors opinion.

Plan of Chapter

In section 2. we briefly recall some basic background material concerning equation (1.1), in particular the theory of Gradient flows in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ and some absolute continuity properties of curves in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$. In section 3 we show that $\mathcal{P}_2(\mathbb{R}^d)$ is invariant under the action of the Markov semigroup of the SDE. Moreover we show that the Trotter-Kato product formula holds when both approximation steps are obtained by the semigroups, provided b satisfies (1.3) and (1.4). In particular product formula holds when $v = \nabla V$ and this is by itself a novel result! As our composition of two steps is a contraction, the Trotter-Kato product formula will be a very powerful tool to show that (1.1) has the regularizing effect, which we adress in section 4. In this section we characterize the domain of the Wasserstin "generator" of (1.1) and show that just as a gradient flow, our semigroup posseses the regularizing effect. That is for any initial measure $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$, the solution μ_t at time $t \geqslant 0$ of (1.1) is a member of the domain of this "generator". and $t \mapsto \mu_t$ is Lipshitz on $[\varepsilon, \infty)$ for each $\varepsilon > 0$. We conclude the section with a theorem that states all the properties of our flow; it does not take long to observe the great analogy with the properties of a flow induced by a Maximal Monotone Operator on Hilbert space (c.f [6] Theoreme 3.1). Finaly in section 5 we look at the invariant measure under the uniform elipticity condition on b, and explain that even in the case where b is a matrix, there are infinitely many positive definite matrices that have the same invariant measure. Hence contrary to the gradient flow case, the invariant measure does not contain sufficient information about the flow of (1.1) in general. For in the case of gradient flow, the resolvents and subsequently the induceed semigroup are fully determined by the relative entropy functional with respect to the invariant measure.

2 Preliminaries

Any maximal monotone subset b of $\mathbb{R}^d$ with domain equal to $\mathbb{R}^d$ induces a contraction semigroup $(S_t)_{t\geqslant 0}$ and it's paths $t \mapsto S_t x$ are Lipschits on $[0,\infty)$ for each $x \in \mathbb{R}^d$ and the equation

$$\frac{d}{dt} S_t x = -b(S_t x)\ \mathcal{L}^1 \text{ a.e on}[0,\infty)\ \ \forall x \in \mathbb{R}^d tag2.1 \tag{1}$$

holds (see [1] Theorem 3.1). Since we assume our monotone operator $b : \mathbb{R}^d \mapsto \mathbb{R}^d$ to be single-valued and at least continuous [1] Proposition 2.4 implies that it is always maximal. We than also have that the function

$$t \mapsto |b(S_t x)| \tag{2.2}$$

is decreasing for each $x \in \mathbb{R}^d$.

For any map $b : \mathbb{R}^d \to \mathbb{R}^d$ that satisfies (1.2) and (1.3) we have the strong existence and uniqueness for the SDE (1.5) and the solutions are Markov processes. We denote $(Q_t)_{t\geqslant 0}$ to be it's Markov semigroup, i.e.

$$Q_t f(x) = Ef((X_t^x), \quad f \in B_b(\mathbb{R}^d) \quad t \geqslant 0 \tag{2.3}$$

and by $(R_t)_{t\geqslant 0}$ the dual semigroup, i.e.

$$R_t\mu(B) := Q_t^*\mu(B) = \int_{\mathbb{R}^d} Q_t 1_B d\mu(x) \quad \mu \in \mathcal{P}(\mathbb{R}^d). \tag{2.4}$$

In this $\mathcal{P}(\mathbb{R}^d)$ denotes the set of Borel probability measures on $\mathbb{R}^d$ and X_t^x denotes the solution of (1.5) with $X_0 := x$ a.s.

We write $\mu_t := Law(X_t)$ the distribution of the solution of time $t \geqslant 0$, and remark that by an easy application of the Ito formula one sees that the weakly continuous curve $t \mapsto \mu_t$ of Borel probability measures is the solution of the Fokker-Planck equation (1.1). Next to this we consider the following two semigroups and their dual semigroups:

$$P_t^1 f(x) := f(S_t x) \quad t \geqslant 0 \quad x \in \mathbb{R}^d \quad f \in C_0(\mathbb{R}^d) \tag{2.5}$$

$$P_t^2 f(x) = \frac{1}{\sqrt{(2\pi)^d t^d}} \int f(y) e^{-|x-y|^2/2t} dy \quad t \geqslant 0 \quad x \in \mathbb{R}^d \quad f \in C_0(\mathbb{R}^d) \tag{2.6}$$

$$R_t^1 \mu := (P_t^1)^* \mu \quad \mu \in \mathcal{P}(X) \tag{2.7}$$

$$R_t^2 \mu := (P_t^2)^* \mu \quad \mu \in \mathcal{P}(X) \tag{2.8}$$

Observe that we have

$$R_t^2 \mu(B) = \frac{1}{\sqrt{(2\pi)^d t^d}} \int_B e^{-|x-y|^2/2t} d\mu(x) \tag{2}$$

for

$$B \in \mathcal{B}(\mathbb{B}^d), \quad \mu \in \mathcal{P}(\mathbb{R}^d), \quad t \geqslant 0. \tag{3}$$

In the next section we will show that $(R_t^1)_{t\geqslant 0}$ and $(R_t^2)_{t\geqslant 0}$ and $(R_t)_{t\geqslant 0}$, are c_0-semigroups on $C_0(\mathbb{R}^d)$-the space of continuous functions on $\mathbb{R}^d$ that vanish at infinity. Subsequently we are going to use these facts in order to show that the Trotter-Kato product formula

$$(R^2_{t/n}R^1_{t/n})^n\mu \xrightarrow{W_2} R_t\mu \text{ for } t \geqslant 0 \quad \mu \in \mathcal{P}_2(\mathbb{R}^d) \tag{2.9}$$

holds, where W_2 denotes the Wasserstein–2 distance, defined in Chapter 1 ???. Having done that we can conclude immedeately that $(R_t)_{t\geqslant 0}$ is W_2–contractive. This result will turn out to be a very habndy tool to analyse the path properties of $(R_t)_{t\geqslant 0}$ (see section 4). [1]

Let us recall some basics and some notation about the Wasserstein-2 space of probability measures on $\mathbb{R}^d$. $\mathcal{P}_2(\mathbb{R}^d)$ will denote the set of probability measures on $\mathbb{R}^d$ such that $\int |x|^2 dp(x) < +\infty$. One can define a metric on $\mathcal{P}_2(\mathbb{R}^d)$ by

$$.W_2^2(\mu_1,\mu_2) := \inf_{\gamma\in\Gamma(\mu_1,\mu_2)} \int |x-y|^2 d\gamma(x,y) \tag{2.18}$$

where $\Gamma(\mu_1,\mu_2)$ is a subset of probability measures on $\mathbb{R}^{2d}$ such that the push forwards $\pi^1_{\#}\gamma$ and, $\pi^1_{\#}\gamma$ of the projection $\pi^1(x,y) := x$ and $\pi^2(x,y) := y$ which map $\mathbb{R}^d \times \mathbb{R}^d$ onto $\mathbb{R}^d$, are μ_1 and μ_2 respectively. One can show that W_2 is complete separable metric space (see [1] Remark 7.1.7). Moreover the infimum in (2.18) is actually a minimum, and the set of minimizers is called the set of optimal transport plans and it is denoted by $\Gamma_0(\mu_1,\mu_2)$. If $\mu_1 \ll \mathcal{L}^d$ ($\mathcal{L}^d$ denotes the Lebesgue measure on. $\mathbb{R}^d$) then $\Gamma_0(\mu_1,\mu_2)$ is a one point set and there is a mapping $r \in L^2(\mu_1;\mathbb{R}^d)$ such that $r_{\#}\mu_1 = \mu_2$ and

$$W_2^2(\mu_1,\mu_2) = \int |x-r(x)|^2 d\mu_1(x) \tag{2.19}$$

Moreover there is a convex function ψ on $\mathbb{R}^d$ such that $r = \nabla\psi$ for μ_1 a.e. $x \in bbR^d$. Such map r is called the optimal transport map

Next we recall some facts about absolutely continuous curves in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$. A curve $(a,b) \ni t \mapsto \mu_t \in (\mathcal{P}_2(\mathbb{R}^d), W_2)$ is said to be of class $AC^2((a,b),(\mathcal{P}_2(\mathbb{R}^d),W_2))$ if there is a $m \in L^2((a,b),dx)$ so that

$$W_2(\mu_{t_1},\mu_{t_2}) \leqslant \int_{t_1}^{t_2} m(r)dr \quad \text{ for } t_1,t_2 \in [a,b] \tag{2.10}$$

[1]The author of this thesis wishes to emphasize two things at this point in time. First he didnt know about the work of Natile-Peletier-Savare [15] where authors prove contractivity of the semigroup of the Non-symmetric Fokker-Planck equation (even with reaspect to a rather wide class of transportation distances) before he completed his investigations. Secondly even though in the above mentioned work a more general result seem to have obtained , the proof there is a different one than the proof presented here. Therefore the present result should be consederd as a new different piece of mathematics in any case.

If (2.10) holds then the right metric derivative

$$\lim_{h\to 0}\frac{W_2(\mu_{t+h},\mu_t)}{h} =: |\dot{\mu}|(t) \tag{2.11}$$

is exists $\mathcal{L}^1$ a.e. on (a,b) and (2.10) holds with $m = |\dot{\mu}|$ and $|\dot{\mu}| \in L^2((a,b),dx)$. If $t \mapsto \mu_t$ is defined on $[c,d]$ $-\infty \leqslant c < d \leqslant +\infty$ and (2.10) holds on $[a,b]$ for all $c < a < b < d$ then we say that mu_t is of class $AC^2_{loc}([c,d],(\mathcal{P}_2(\mathbb{R}^d),W_2))$. If $\mu : (a,b) \to (\mathcal{P}_2(\mathbb{R}^d),W_2)$ is of class AC^2 then there is a family $v_t \in L^2(\mu_t;\mathbb{R}^d)$ defined $\mathcal{L}^1$ a.e. on (a,b) such that

$$(t,x) \mapsto v_t(x) \tag{4}$$

is a Borel map,

$$\int_a^b \int_{\mathbb{R}^d} (\partial_t\varphi(t,x) + \langle \nabla_x\varphi(t,x), v_t(x)\rangle) d\mu_t(x)dt = 0 \tag{2.12}$$

for all $\varphi \in C_c^\infty((a,b)\times\mathbb{R}^d)$, and

$$\int_a^b ||v_t||^2_{L^2(\mu_t;\mathbb{R}^d)} dt < +\infty \tag{2.13}$$

This claim and it's converse are proved in [1] Theorem 8.3.1, and (2.12) is called the continuity equation. More shortly (2.12) reeds

$$\partial_t\mu_t + \nabla\cdot(v_t\mu_t) = 0 \text{ in } \mathcal{D}'((a,b)\times\mathbb{R}^d). \tag{2.14}$$

The (regular) tangent space of a measure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ with $(\mathcal{P}_2(\mathbb{R}^d),W_2)$ as the ambient space is defined by

$$Tan_\mu(\mathcal{P}_2(\mathbb{R}^d)) := \overline{\{\nabla\varphi | \varphi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu;\mathbb{R}^d)} \tag{2.15}$$

and when $\mu \ll \mathcal{H}^{d-1}$-the $d-1$ dimensional Haussdorf meassure on $\mathbb{R}^d$ this definition coincides with the abstract definition of the Euclidean tangent cone to a metric space with nonpositive curvature in the sense of Alexandrov. [2]

Moreover there is a $\mathcal{L}^1$ a.e. unique choice of a vector field $(v_t)_{t\in(a,b)}$ in (2.12) s.t. $v_t \in Tan_{\mu_t}(\mathcal{P}_2(\mathbb{R}^d))$ $\mathcal{L}^1$ a.e. on $[a,b]$.A path of such curve can also be locally "linearized", a claim comparable with the first order Taylor expansion (see [1] Proposition 8.4.6 and Theorem 8.4.7).

As we alreday explained in our introductory Capter 1, lower semicontinuous mappings

$$\varphi : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty,+\infty] \tag{2.16}$$

that satisfy conditions $\varphi \not\equiv +\infty$, φ is bounded from bellow on same ball in $\mathcal{P}_2(\mathbb{R}^d)$ and is convex along generalized geodesics (see [1] Chapter 9) inducees

[2] The basic theory of geometry on spaces with curvature bounded bellow or above in the sence of Alexandrov can be found in [7]. The proof of the fact that $\mathcal{P}^2(\mathbb{R}^d)$ is nonpositively curved in this sense is given in [1] Theorem 7.3.2

a gradient flow on $(\mathcal{P}_2(\mathbb{R}^d, W_2)$. This is a semigroup which enjoys certainThe inequalities follow by path regularity properties analogous to the classical gradient flows on Hilbert spaces. This abstract and somewhat technical theory turns to be a very usefull way to interpret of several classes of PDE's, such as the heat equation, the symmetric Fokker–Planck equation, and the porous medium equation. In the sequal we will consider the relative Entropy functional (relative to $\mathcal{L}^d$) which induces solutions of the heat equation:

$$\mathcal{H}(\mu) := \begin{cases} \int \frac{d\mu}{dx} \log \frac{d\mu}{dx} dx & \text{if } \mu \ll \mathcal{L}^d \\ +\infty & \text{otherwise} \end{cases} \tag{2.17}$$

For the full treatment of these and related topics we refer to [1], [**?**],[17], [23], [24], and the numerous references therein.

3 Construction of the semigroup on $(\mathcal{P}_2, W_2)$ – The Trotter-Kato product formula

We want to show that (2.9) holds, and our strategy is as follows. We first show that convergence of iterations as in(2.9) holds in the weak* sense in $C_0(\mathbb{R}^d)^*$. Next we obtain bounds for the fourth order moments of iterations $(R^2_{t/n}, R^1_{t/n})^k$ for $\mu \in \mathcal{P}_4(\mathbb{R}^d) = \{\mu \in \mathcal{P}(\mathbb{R}^d) | \int |x|^4 d\mu(x) < +\infty\}$, and finally use that to show (2.9). Once we prove these statements we will be able to show that $\mathcal{P}_2(\mathbb{R}^d)$ is $(R_t)_{t\geqslant 0}$ invariant and that this semigroup is W_2–contracting.

Lemma 3.1 *$\mathcal{P}_2(\mathbb{R}^d)$ is invariant under the action of the semigroup $(R^1_t)_{t\geqslant 0}$, and this action is W_2 contracting. Moreover for any $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ and $0 < t_1 < t_2$*

$$W_2^2(\mu_{t_1}, \mu_{t_2}) \leqslant (t_2 - t_1)^2 \int |b(x)|^2 d\mu_0(x) \tag{3.1}$$

where $\mu_t := R_t \mu_0$.the content of

Proof Clearly

$$\begin{aligned} \int |x|^2 d\mu_t = \int |S_t x|^2 d\mu_0(x) \leqslant |S_t 0|^2 + \int |S_t x - S_t 0|^2 d\mu_0(x) \\ \leqslant |S_t 0| + \int |x|^2 d\mu_0(x) < +\infty \end{aligned}$$

so $\mathcal{P}_2(\mathbb{R}^d)$ is $(R^1_t)_{t\geqslant 0}$ invariant. Next for $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$ pick a $\gamma \in \Gamma_0(\mu_1, \mu_2)$. As $(S_t, S_t)_\# \gamma \in \Gamma(S_t \mu_1, S_t \mu_2)$ for each $t \geqslant 0$ we estimatethe content of

$$\begin{aligned} W_2^2(R^1_t \mu_1, R^1_t \mu_2) &\leqslant \iint |S_t x - S_t y|^2 d\gamma(x,y) \leqslant \\ &\leqslant \iint |x - y|^2 d\gamma(x,y) = W_2^2(\mu_1, \mu_2) \end{aligned} \tag{3.2}$$

hence $(R^1_t)_{t\geqslant 0}$ is W_2–contactive.

Finaly for $0 < t_1 < t_2$ the map $\psi(x) := (S_{t_1}x, S_{t_2}x)$ has property that $\psi_{\#}\mu_0 \in \Gamma(S_{t_1}\mu_0, S_{t_2}\mu_0)$ for any $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ hence we estimate

$$\begin{aligned} W_2^2(R^1_{t_1}\mu_0, R^1_{t_2}\mu_0) &\leqslant \int |S_{t_1}x - S_{t_2}x|^2 d\mu_0(x) = \\ &= \int_{\mathbb{R}^d} \left| \int_{t_1}^{t_2} -b(S_1 x) ds \right|^2 d\mu_0(x) \leqslant |t_0 - t_1|^2 \int |b(x)|^2 d\mu_0(x) \end{aligned} \tag{3.3}$$

using (2.1) and (2.2).

□

We want to show that $(P^1_t)_{t\geqslant 0}$ and $(P^2_t)_{t\geqslant 0}$ are c_0–semigroups on $C_0(\mathbb{R}^d)$. For $(P^1_t)_{t\geqslant 0}$ we first observe that since b satisfies (1.2) and (1.3) we can solve the ODE system

$$\tilde{S}_t x - x = \int_0^t b(\tilde{S}_s x) ds \tag{3.5}$$

on $[0, \infty) \times \mathbb{R}^d$ and extend $(S_t)_{t\geqslant 0}$ to a group on $\mathbb{R}^d$ by

$$S_t x := \tilde{S}_{-t} x \quad t < 0 \quad x \in \mathbb{R}^d \tag{3.6}$$

which in particular means that for each $t \geqslant 0$, S_t has a continuous inverse. This in turn implies that for $f \in C_0(\mathbb{R}^d)$ and $x_n \to +\infty$, $P^1_t f(x_n) = f(S_t x_n) \to 0$, i.e. $P^1_t(C_0(\mathbb{R}^d)) \subset C_0(\mathbb{R}^d)$ for $t \geqslant 0$.

For the strong continuity property of $(P^1_t)_{t\geqslant 0}$ i.e. $\lim_{t\downarrow 0} P^1_t f = f$ in $C_0(\mathbb{R}^d)$ we appeal to [20] Lemma III 6.7, and observe that indeed for $x \in \mathbb{R}^d$ and $f \in C_0(\mathbb{R}^d)$ $P^1_t f(x) = f(S_t x) \to f(x)$ as $t \downarrow 0$. As for $(P^2_t)_{t\geqslant 0}$ an easy calculation yields the same conclusions. We therefore have the following

Proposition 3.2 *The restriction of $(P^1_t)_{t\geqslant 0}$ and $(P^2_t)_{t\geqslant 0}$ to $C_0(\mathbb{R}^d)$ are contractive c_0 semigroups.*

Proof In light of the above discussion we only need to show contractivness. But this is easy, for $t \geqslant 0$ and $f \in C_0(\mathbb{R}^d)$ we have

$$|P^1_t f(x)| = |f(S_t x)| \leqslant |f|_\infty$$

and

$$\begin{aligned} |P^2_t f(x)| &= \left| \frac{1}{\sqrt{(2\pi)^d t^d}} \int f(y) e^{-|x-y|^2/et} dy \right| \leqslant \\ &\leqslant |f|_\infty \frac{1}{\sqrt{(2\pi)^d t^d}} \int e^{-|x-y|^2/2t} dy = |f|_\infty \end{aligned}$$

□

We also have that $Q_t(C_0(\mathbb{R}^d)) \subset C_0(\mathbb{R}^d)$. Indeed this claim holds since (1.2) insures that the condition (5.3.6) of Example 5.3.3 in [14]. Moreover solutions of (1.5) have continuous paths a.s. hence $Q_t f(x) = \mathbb{E}f(X_t^x) \to \mathbb{E}f(X_0^x) = f(x)$ as $t \to 0$, and again by [20] Lemma III 6.7 we conclude that $(Q_t)_{t\geqslant 0}$ is a C_0 semigroup on $C_0(\mathbb{R}^d)$. Observe that the generators od P_t^1, P_t^2, P_t are

$$\mathcal{A}_1 f := -\langle \nabla f, b\rangle \quad f \in \mathcal{D}(\mathcal{A}_1) \tag{3.7}$$

$$\mathcal{A}_2 f := \Delta f \quad f \in \mathcal{D}(\mathcal{A}_2) \tag{3.8}$$

$$\mathcal{A} f := \Delta f - \langle \nabla f, b\rangle \quad f \in \mathcal{D}(\mathcal{A}) \tag{3.9}$$

respectively.

Since $(Q_t)_{t\geqslant 0}$ is clearly a contraction on $C_0(\mathbb{R}^d)$, [10] Theorem 3.15 yields that $\overline{Im(\lambda - \mathcal{A})} = C_0(\mathbb{R}^d)$ for $\lambda > 0$. But than since $(P_t^1)_{t\geqslant 0}$ and $(P_t^2)_{t\geqslant 0}$ are also contraction on $C_0(\mathbb{R}^d)$, [10] Corollary 5.8 gives the followong result.

Proposition 3.3 *For any* $f \in C_0(\mathbb{R}^d)$

$$\lim_n (P_{t/n}^2 P_{t/n}^1)^n f = Q_t f \ \textit{in}\ C_0(\mathbb{R}^d) \tag{3.10}$$

and

$$\lim_n (P_{t/n}^1 P_{t/n}^2)^n f = Q_t f \ \textit{in}\ C_0(\mathbb{R}^d) \tag{3.11}$$

uniformly on compact time intervals.

Corollary 3.4 *For* $\mu \in \mathcal{M}(\mathbb{R}^d) = C_0(\mathbb{R}^d)^*$

$$(R_{t/n}^2 R_{t/n}^1)^n \mu \to R_t \mu \ \textit{for}\ t \geqslant 0 \tag{3.12}$$

$$(R_{t/n}^1 R_{t/n}^2)^n \mu \to R_t \mu \ \textit{for}\ t \geqslant 0 \tag{3.13}$$

the convergence being in the weak sense of* $C_0(bbR^d)^*$.

In next two lemmas we obtain estimates for the fourth order moments of the successive iterations of a measure $\mu_0 \in \mathcal{P}_4(\mathbb{R}^d)$.

Lemma 3.5 *There is a constant* $c > 0$ *depending only on the linear growth constant of* b *in (1.3), such that for any* $\mu_0 \in \mathcal{P}_4(\mathbb{R}^d)$ *and for any* $t > 0$

$$\int |x|^4 dR_t^1 \mu_0 dx \leqslant e^{ct} \int |x|^4 d\mu_0(x) + cte^{ct} \tag{3.14}$$

Proof We simply estimate (using (1.3), (2.2))

$$\int |x|^4 dR^1_t\mu_0(x) = \int |S_t x|^4 d\mu_0(x) = \int \left| x - \int_0^t b(S_s x)ds \right|^4 d\mu_0 =$$

$$= \int \left(|x|^2 - 2\left\langle x, \int_0^t b(S_s x)ds \right\rangle + \left| \int_0^t b(S_s x)ds \right|^2 \right)^2 d\mu_0(x)$$

$$\leqslant \int [|x|^4 + 4|x|^3 tc(1+|x|) + 4|x|(tc)^3(1+|x|)^3 +$$

$$+ 2|x|^2(tc)^2(1+|x|)^2 + 2|x|^2(tc)^2(1+|x|)^2 + (tc)^4(1+|x|)^4] d\mu_0(x)$$

$$\leqslant [|x|^4 + 4tc|x|^3 + (4tc)^2|x|^2 + (4tc)^3|x| + (4tc)^4 +$$

$$+ (4tc + (4tc)^2 + (4tc)^3 + (4tc)^4)|x|^4] d\mu_0(x)$$

As for $k = 1, 2, 3$

$$\int |x|^k d\mu_0(x) = \int_{|x|\geqslant 1} + \int_{|x|<1} |x|^k d\mu_0(x) \leqslant 1 + \int |x|^4 d\mu_0(x)$$

we obtain

$$\int |x|^4 dR^1_t\mu_0(x) \leqslant \int \sum_{k=0}^{4} (4tc)^k |x|^4 d\mu_0(x) + \sum_{k=0}^{4} (4tc)^k$$

$$\leqslant e^{ct} \int |x|^4 d\mu_0(x) + cte^{ct}$$

(for a different but finite constant c).

□

In a similar fashion we have

Lemma 3.6 *There is a constant $c > 0$ such that for*

$$\mu_0 \in \mathcal{P}_4(\mathbb{R}^d) \text{ and } t > 0$$

$$\int |x|^4 dR^2_t\mu_0(x) \leqslant e^{ct} \int |x|^4 d\mu_0(x) + ct \tag{3.15}$$

Proof This is a well known fact. However it follows for instace by [22]. In that monograph a similar estimate is proven for the Wasserstein resolvent associated to the relative entropy functional of any initial measure with finite fourth order moment, and than the exponential formula (4.0.11) in [1] implies our claim.

□

Lemma 3.7 *For any $T > 0$ there is a constant $d > 0$ depending only on b such that for each $\mu_0 \in \mathcal{P}_4(\mathbb{R}^d)$ $0 \leqslant t \leqslant T$ $n \in \mathbb{N}$ and $0 \leqslant k \leqslant n$*

$$\int |x|^4 d(R^2_{t/n}R^1_{t/n})^k\mu_0(x) \leqslant e^{2cT} \int |x|^4 d\mu_0(x) + d \tag{3.16}$$

$$\int |x|^4 d(R^1_{t/n}R^2_{t/n})^k\mu_0(x) \leqslant e^{2cT} \int |x|^4 d\mu_0(x) + d \tag{3.17}$$

Proof By induction we have that for $k = 1, 2, ..., n$

$$\begin{aligned}\int |x|^4 d(R^2_{t/n} R^1_{t/n})^k \mu_0(x) &\leqslant e^{k \cdot 2ct/n} \int |x|^4 d\mu_0(x) + ct \sum_{j=1}^{2k} e^{jct/n} \\ &\leqslant e^{2cT} \int |x|^4 d\mu_0(x) + cT \sum_{k=1}^{2n} e^{kcT/n} \\ &= e^{2cT} \int |x|^4 d\mu_0(x) + \frac{cT}{n} e^{cT/n} \frac{1 - e^{2cT}}{1 - e^{cT/n}}\end{aligned}$$

and on the function $x \mapsto \frac{xe^x}{e^x - 1}$ in continuous on $[0, T]$ and $0 < \frac{T}{n} \leqslant T$ for $n \in \mathbb{N}$ we easily obtain constant $d > 0$ such that (3.16) holds. (3.17) is obtained in likewise fashion.

□

With Corollary 3.4 and Lemma 3.7 at hand we are now able to show that the Product formula holds in our case. Notice that in particular it holds when b is a gradient of a convex l.s.c. function V on $\mathbb{R}^d$ (provided it satisfies (1.2) and (1.3)).

Theorem 3.8 *For $t > 0$ and $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ we have $R_t\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ and*

$$(R^2_{t/n} R^1_{t/n})^n \mu_0 \xrightarrow{W_2} R_t\mu_0 \quad (R^1_{t/n} R^2_{t/n})^n \mu_0 \xrightarrow{W_2} R_t\mu_0 \tag{3.18}$$

Moreover $(R_t)_{t \geqslant 0}$ in W_2 - contractive.

Proof We will show the first claim in (3.18) only; the second one follows by symmetry of the situation. Fix $T \geqslant 0$, and pick any $\mu_0 \in \mathcal{P}_4(\mathbb{R}^d)$. By Corollary 3.4 we have that $(R^2_{t/n} R^1_{t/n})^n \mu_0$ converges to $R_t\mu_0$ in the weak* sense in $C_0(\mathbb{R}^d)^*$. Moreover by Lemma 3.7 the fourth order moments of this iterations are uniformly bounded for $n \in \mathbb{N}$ and $t \in [0, T]$, hence the second moments are uniformly integrable. This means that for $t \in [0, T]$ any subsequence of $\{(R^2_{t/n} R^1_{t/n})^n\}_n$ has further subsequence which converges w.r.t. W_2 distance to some measure in $\mathcal{P}_2(\mathbb{R}^d)$. But since W_2 convergence implies weak* in duality with $C_b(\mathbb{R}^d)$, it also implies weak* convergence in $C_0^*(\mathbb{R}^d)$ as the last space contains less functions for which the convergence of the integrals must be tested. By the previous argument there can be only one limit and that limit is $R_t\mu_0$. This implies that for $R_t\mu_0 \in \mathcal{P}_4(\mathbb{R}^d)$, $(R^2_{t/n} R^1_{t/n})^n \mu_0 \xrightarrow{W_2} R_t\mu_0$. Moreover as for each n $(R^2_{t/n} R^1_{t/n})^n : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ is W_2-contracting, so is $R_t : \mathcal{P}_4(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$.

Next pick a general $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ pick $\varepsilon > 0$ and $\mu_1 \in \mathcal{P}_4(\mathbb{R}^d)$ such that $W_2(\mu_0, \mu_1) < \varepsilon$-we can always find such μ_1.

Since $(R^2_{t/n} R^1_{t/n})^n \mu_1 \xrightarrow{W_2} R_t\mu_1$, it is a Caushy sequence, hence there is an $N \in \mathbb{N}$ so that for $n, m \geqslant N$, $W_2((R^2_{t/n} R^1_{t/n})^n \mu_1, (R^2_{t/m} R^1_{t/m})^m \mu_1) < \varepsilon$.

But then since $(R^2_{t/n}, R^1_{t/n})^n$ is W_2- contractive, the triangle inequality yields

$$W_2((R^2_{t/n}R^1_{t/n})^n\mu_0, (R^2_{t/m}R^1_{t/m})^m\mu_0) < 3\varepsilon \quad n, m \geqslant N$$

i.e. this sequence is Caushy. The limit must be the same as the weak* limit $R_t\mu_0$ of these itterations, and we have proven that the product formula holds. The contracting property is now obvious, as $(R^2_{t/n}R^1_{t/n})^n$ is contracting for all $t > 0$ and $n \in \mathbb{N}_0$.

□

Remark Arguing as in Lemma 3.5, Lemma 3.6 and Lemma 3.7 we can obtain that for each $T > 0$ and $p \in \mathcal{P}_2(\mathbb{R}^d$

$$\sup\{S|x|^2 d(R^2_{t/n}R^1_{t/n})^k d\mu(x) | n \in \mathbb{N} \;\; k = 1, 2, ..., n \;\; t \in [0, T]\} < +\infty$$

i.e. such set of iterations is contained in a ball. We omit the proof as it is entirely analogous to what we already have done, but remark that we will use this fact in Section 4.

4 Absolute continuity of paths and the regularizing effect

As it is shown in [1] gradient flow paths of a semigroup associated to a functional φ on $\mathcal{P}_2(\mathbb{R}^d)$ that enjoys appropriate convexity and non-degeneracy conditions together with the lower seminontinuity property, exibits a very strong regularity properties. Precisely for any initial point $\mu_0 \in \overline{D(\varphi)} \subset \mathcal{P}_2(\mathbb{R}^d)$ the curve $t \mapsto S_t\mu_0 = \mu_t$ is Lipshitz on $[\varepsilon, +\infty)$ for all $\varepsilon > 0$, and $\mu_t \in D(|\partial\varphi|)$ for all $t > 0$. In such situation the analogy with semigroups induced by a l.s.c. convex functional on a Hilbert space, or even with linear c_0–semigroups on a Banach space is the following: the domain of the slope corresponds to the domain of the subdifferential respectively the domain of the generator. Also a path that is localy Lipschitz on $[0, +\infty)$ corresponds to a path that is strongly differentiable for $t > 0$ in a Hilbert space. Moreover since b is not assumed to be a gradient there is no reason why our semigroup $(R_t)_{t\geqslant 0}$ should be a gradient flow. But if b satisfies (1.3) the paths of the semigroup $(R^1)_t$ are lipschitz as we compute for $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ and for $t \geqslant s \geqslant 0$

$$\frac{1}{(t-s)^2} W_2^2(R^1_s\mu_0, R^1_t\mu_0) \leqslant \frac{1}{(t-s)^2} \int \left| \int_s^t b(S_r x) dr \right|^2 d\mu_0(x) \leqslant \int |b(x)|^2 d\mu_0(x) \tag{4.1}$$

since $r \mapsto |b(S_r x)|$ is a decreasing function c.ff (2.2) and $(S_s, S_t)_{\#}\mu_0 \in \Gamma(R^1_s\mu_0), R^1_t\mu_0$. We interpret as this observation as its "generator" being smooth. Now it should be reasonable to expect that perturbing (in the sence of "addition") of "a generator" with "a smooth one' should not alter any regularity properties. Precisely we expect that R_t has the same locally Lipschitz properties as $(R^2_t)_{t\geqslant 0}$ does, and that for any initial measure μ_0, $\mu_t = R_t\mu_0 \in D(\partial\mathcal{H})$ for eacht $t \geqslant 0$ where $\mathcal{H}$ is the relative entropy functional defined in (2.17). We are now going to prove this under the assumption that b is globally Lipschitz. Through we have to put some effort to this, the result seems to justify our efforts. Just as in the case where $b = \nabla V$ as considered in [12] and [1] these regularity properties, the contractivity and the form of the derivative along a gradient flow path in the Wanesstein sence, indicates that the Wanesstein-2 space in the natural environment for our equation (1.1).

Before proceeding we state two results from [24]

Theorem 4.1 *Assume that b is Lipschitz, and denote $l := Lip(b)$. Than for any $t > 0$, $t_2 > t_1 > 0$ and $\mu_0 \in D(\mathcal{H})$, $\sigma \in D(\mathcal{H})$ we have*

$$\mathcal{H}(R^1_t\mu_0) \leqslant \mathcal{H}(\mu_0) + tl \tag{4.2}$$

$$\mathcal{H}(R_t\mu_0) \leqslant \mathcal{H}(\mu_0) + tl \tag{4.3}$$

$$\frac{1}{\alpha}W_2^2(\mu_{t_2},\sigma) - \frac{1}{\alpha}W_2^2(\mu_{t_1},\sigma) \leqslant -\int_{t_1}^{t_2}\int \langle b_0, r^2_\mu - i\rangle d\mu_1 ds + \int_{t_1}^{t_2} \mathcal{H}(\sigma) - \mathcal{H}(\mu_1) ds \tag{4.4}$$

Proof See for i) see [16] Proposition 2 For ii) and iii) see [16] Theorem 2.

Proposition 4.2 *Assume that b monotone and globally Lipschitz. Than for any $\mu_0 \in D(\mathcal{H})$, the curve $t \mapsto \mu_t := R_t\mu_0$ is Lipschitz on $[\varepsilon, +\infty)$ for all $\varepsilon > 0$ and of class $AC^2_{loc}((0,+\infty); CalP_2(bbR^d))$. Moreover $\mu_t \in D(|\partial\mathcal{H}|)$ for each $t > 0$.*

Proof Pick a $t > 0$, and set $t_n := t \cdot 2^{-n}$ for $n \in \mathbb{N}$ and denote $l := Lip(b)$. In the sequal we also abrevate $R^{2,1}_s = R^2_s R^1_s$ for $s > 0$ in order to simplify the notation. Define piecewise constant curves on $[0, t]$

$$\mu^n_0 := \mu_0 \quad \mu^n_{(k+1)t_n} := R^{2,1}_{t_n}\mu^n_{kt_n}, \quad k = 0, ..., 2^n - 1 \tag{4.5}$$

$$\mu^n_{kt_n+h} := \mu^n_{(k+1)t_n} \quad \text{for } 0 < h < 2^{-n} \tag{4.6}$$

The Evolution variational inequality (EVI) (4.0.13) in [1] gives (by integrating and (2.4.26) in [1] which gives that $t \mapsto \mathcal{H}(R^2_t\nu)$ in decreasing for $\nu \in \mathcal{P}_2(\mathbb{R}^d)$)

$$\frac{W_2^2(R^{2,1}_{t_n}\mu_0, R^1_{t_n}\mu_0)}{2t_n} \leqslant \varphi(R^1_{t_n}\mu_0) - \varphi(R^{2,1}_{t_n}\mu_0) \tag{4.7}$$

so that by (4.2)

$$\frac{W_2^2(R^{2,1}_{t_n}\mu_0, R^1_{t_n}\mu_0)}{2t_n} \leqslant \varphi(\mu_0) - \varphi(R^{2,1}_{t_n}\mu_0) + t_n l \tag{4.8}$$

Moreover the properties of $s \mapsto S_s x$ give that

$$\begin{aligned} W_2^2(R^1_{t_n}\mu_0,\mu_0) &\leqslant \int_{\mathbb{R}} |x - S_{t_n}x|^2 d\mu_0(x) \\ &\leqslant t_n^2 \int_{\mathbb{R}^d} |b(x)|^2 d\mu_0(x) \leqslant c^2 t_n^2 (1 + W_2^2(\mu_0,\delta_0)) \end{aligned} \tag{4.9}$$

where c is a linear growth constant of b. As

$$W_2^2(R^{2,1}_{t_n}\mu_0,\mu_0) \leqslant 2W_2^2(R^{2,1}_{t_n}\mu_0, R^1_{t_n}\mu_0) + 2W_2^2(R^1_{t_n}\mu_n,\mu_0)$$

combining (4.8) and (4.9) we arrive

$$\frac{W_2^2(R^{2,1}_{t_n}\mu_0,\mu_0)}{t_n} \leqslant 4(\varphi(\mu_0) - \varphi(R^{2,1}_{t_n}\mu_0)) + 4t_n l + 4c^2 t_n (1 + W_2^2(\mu_0,\delta_0)). \tag{4.10}$$

Recalling the definition (4.5) we then have for $k = \overline{0, 2^n - 1}$

$$\frac{W_2^2(\mu^n_{kt_n},\mu^n_{(k+1)t_n})}{t_n} \leqslant 4(\varphi(\mu^n_{kt_n}) - \varphi(\mu^n_{(k+1)t_n})) + 4t_n(l + c^2) + 2t_n c^2 W_2^2(\mu^n_{kt_n},\delta_0) \tag{4.11}$$

Summing for over k we obtain

$$\sum_{k=0}^{2^n-1} \frac{W_2^2(\mu^n_{kt_n},\mu^n_{(k+1)t_n})}{t_n} \leqslant 4(\varphi(\mu^n_0) - \varphi(\mu^n_t)) + 4(l + c^2) + \sum_{k=0}^{2^n-1} 2c^2 t_n W_2^2(\mu^n_{kt_n},\delta_0) \tag{4.12}$$

As we observed in Remark 3.8 all measures μ^n_s $n \in \mathbb{N}$ $s \in [0,t]$ are contained in a ball. And the proof of [1] Theorem 11.2.5 and [8] Lemma 4.1 imply that $\inf\{\mathcal{H}(\mu^n_t) | n \in \mathbb{N}\} > -\infty$. Hence there is a constant $K < +\infty$ such that the right side in (4.12) is bounded by K for all $n \in \mathbb{N}$. Next, we introduce the piecewise constant (hence measurable) and decreasing (as $R^{2,1}_s$ is a contraction for $s > 0$) functions u^n on $[0,t)$ for $n \in \mathbb{N}$

$$u^n(s) := \frac{W_2(\mu^n_{kt_n},\mu^n_{(k+1)t_n})}{t_n}, \quad u^n(0) := 0 \tag{4.13}$$

if $s = kt_n + h$, $0 < h \leqslant 2^{-n}$, $k = 0, ..., 2^n - 1$.

By definition, for $0 \leqslant s_1 < s_2 \leqslant t$.

$$W_2^2(\mu^n_{s_1},\mu^n_{s_2}) \leqslant \int_{s_1}^{s_2} u^n(s) ds, \quad n \in \mathbb{N} \tag{4.14}$$

Morevover (4.22) says that

$$\sup_n \int_0^t (u^n(r))^2 dr < \infty \tag{4.15}$$

Now by Helly's theorem (see [1] Thm. 3.3.3) these is a subsequence $\{u^n\}$ which we again denote $\{u^k\}$ and a nonincreasing function $u : [0,t] \to [-\infty, \infty]$ s.d.

$$u^n(t) \to u(r) \quad \text{for each } r \in [0,t] \tag{4.15}$$

By taking a further subsequence if neccessary we may assume that $u^n \to r$ weakly in $L^2(0,t)$. Fix $s_1, s_2 \in (0,t)$ of the form $s_1 = k_1 2^{-n_0} t$, $s_2 = k_2 2^{-n_0} t$ for some $k_1, k_2, n_0 \in \mathbb{N}$. We than have

$$\begin{aligned} \mu^n_{s_j} &= (R^{2,1}_{t/2^n})^{k_j 2^{n-n_0} t} \mu_0 = \left(R_{\frac{tk_j 2^{n_0}}{k_j 2^{n-n_0}}} \right)^{k_j 2^{n-n_0}} \mu_0 = \\ &= (R^{2,1}_{s_j / k_j 2^{n-n_0}})^{k_j 2^{-n_0}} \mu_0 \to R_{s_j} \mu_0, \ quad j = 1, 2 \end{aligned} \tag{4.17}$$

hence as $u^n \to r$ weakly in $L^2(0,t)$, we also have

$$W_2(\mu_{s_1}, \mu_{s_2}) \leqslant \int_{s_1}^{s_2} u(r) dr. \tag{4.18}$$

Recall that our aim is to show that for any $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ the curve $t \mapsto \mu_t$ is locally Lipschitz with $\partial|\mathcal{H}|(\mu_t) < +\infty$. Now by Proposition 4.2 it is enough to show that $\mathcal{H}(\mu_t) < +\infty$ for $t > 0$; indeed having proved that, we will know that the differentiability properties hold for $s \mapsto \mu_{t+s}$ and since $t > 0$ is arbitrary here, the claim regularizing effect property follows.

Proposition 4.3 *Assume that b is globally Lipschitz and monotone. Then for any $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ the curve $t \mapsto \mu_t = R_t \mu$ is Lipschitz on $[\varepsilon, +\infty)$ for any $\varepsilon > 0$, and $|\partial \mathcal{H}|(\mu_t) < +\infty$ for each $t > 0$.*

Proof As we explained above it is enough to show that $\mathcal{H}(\mu_t) < +\infty$ for $t > 0$. Pick a sequence $\{\mu_{0,k}\}_k$ converging to μ_0 (w.r.t. W_2) such that for all $k \in \mathbb{N}$ $|\partial\mathcal{H}|(\mu_{0,k}) < +\infty$; this is always possible as $\overline{D|\partial\mathcal{H}|} = \overline{D(\mathcal{H})} = \mathcal{P}_2(\mathbb{R}^d)$. Denote further $\mu_{t,k} =: R_{t\mu_{0,k}}$. By (4.3)

$$\mathcal{H}(\mu_{\alpha,k} \leqslant \mathcal{H}(\mu_{0.k}) + tl \quad . \tag{4.23}$$

But than for $t > s > 0$ and $k \in \mathbb{N}$

$$\mathcal{H}(\mu_{t,k}) - tl \leqslant \mathcal{H}(\mu_{sk}) - sl \tag{4.24}$$

which implies

$$t\mathcal{H}(\mu_{t,k}) - tl \leqslant \int_0^t (\mathcal{H}(\mu_{s,k}) - sl) ds \quad , \tag{4.25}$$

and this in turn implies

$$t\mathcal{H}(\mu_{t,k}) \leqslant \int_0^t \mathcal{H}(\mu_{s,k}) + tl \quad . \tag{4.26}$$

On the other hand (4.4) for $\sigma \in D(\mathcal{H})$ gives

$$\begin{aligned}\int_0^t \mathcal{H}(\mu_{s,k})ds &\leqslant t\mathcal{H}(\mu_{0,k}) + \int_0^t \int_{\mathbb{R}^d} \langle \sigma_s r^{\sigma}_{\mu_{s,k}} - i\rangle d\mu_{s,k} \\ &+ \frac{1}{2}W_2^2(\mu_{0,k},\sigma) - \frac{1}{\alpha}W_2^2(\mu_{t,k},\sigma) \\ &\leqslant t(\mathcal{H}(\mu_0,k) + \frac{1}{2}W_2^2(\mu_{s,k},\sigma)) \\ &+ \frac{1}{2}\int_0^t (\int |b(x)|^2 d\mu_{s,k}(k) + W_2^2(\mu_{s,k},\sigma))ds\end{aligned} \tag{4.27}$$

Choose $\sigma_k = \mathcal{J}_t\mu_{0,k} := \mathrm{argmin}_{\nu\in\mathcal{P}_2(\mathbb{R}^d)}(\frac{1}{2t}W_2^2(\mu_{0,k},\nu) + \mathcal{H}(\nu))$ in (4.27) and combine this with (4.25) to obtain

$$\mathcal{H}(\mu_{t,k}) \leqslant \mathcal{H}_t(\mu_{0,k}) + tl + \frac{1}{2}\int_0^t \int |b(x)|^2 d\mathcal{J}_t\mu_{0,k}x_d s + \frac{1}{2}\int_0^t W_2^2(\mu_{s,k},\mathcal{J}_t\mu_{0,k})ds \tag{4.28}$$

Well now for a fixed $t > 0$ the resolvent mapping $\mathcal{J}_t : \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d)$ is continuous (c.f. [1] Theorem 4.1.2 (i)) and as $|b|^2$ is continuous with quadratic growth

$$\frac{1}{2}\int_0^t \int |b(x)|^2 d\mathcal{J}_t\mu_{0,k}(x) = \frac{1}{2}t\int |b(x)|^2 d\mathcal{J}_t\mu_{0,k}(x) \to \frac{1}{2}t\int |b(x)|^2 d\mathcal{J}_t\mu_0(x) \tag{4.29}$$

as $k \to +\infty$. In particular the integrals appearing at the left side in (5.29) form a bounded sequence of positive members. Moreover by contractivity of $(R_t)_{t\geqslant 0}$

$$W_2(\mu_{s,k},\mathcal{J}_t\mu_{0,k}) \leqslant W_2(\mu_{s,k},\mu_s) + W_2(\mu_s,\mathcal{J}_t\mu_{0,k}) \leqslant W_2(\mu_{0,k},\mu_0) + W_2(\mu_s,\mathcal{J}_t\mu_{0,k})$$

and since $\mu_{0,k} \to \mu_0$, $\mathcal{J}_t\mu_{0,k} \to \mathcal{J}_t\mu_0$ and by Remark below Theorem 3.8 $\{\mu_s | s \in [0,t]\}$ is a bounded set, we see that

$$\sup_{k\in\mathbb{N}} \frac{1}{2}\int_0^t \int_{\mathbb{R}^d} |b(x)|^2 d\mathcal{J}_t p_{0,k}(x)ds + \frac{1}{2}\int_0^t W_2^2(\mu_{s,k},\mathcal{J}_t\mu_{0,k})ds < +\infty$$

We can now finish the proof since $\mathcal{H}$ is l.s.c. and $\nu \mapsto \mathcal{H}_t(s) = \min_{\sigma\in\mathcal{P}_2} \frac{1}{\partial t}W_2^2(\nu,\sigma) + \mathcal{H}(\sigma)$ is a continuous mapping (c.f. [1] Lemma 3.12), by estimating

$$\begin{aligned}\mathcal{H}(\mu_t) &\leqslant \liminf_{k\to+\infty} \mathcal{H}(\mu_{0,k}) \leqslant \\ &\leqslant \liminf_{k\to\infty}[\mathcal{H}_t(\mu_{0,k}) + tl + \frac{1}{2}\int_0^t \int |b(x)|^2 d\mathcal{J}_t\mu_{0,k}(x)ds + \frac{1}{2}\int_0^t W_2^2(\mu_{s,k},\mathcal{J}_t\mu_{0,k})ds] \leqslant \\ &\leqslant \limsup_{k\to+\infty}[\mathcal{H}_t(\mu_{0,k}) + tl + \frac{1}{2}\int_0^t \int |b(x)|^2 d\mathcal{J}_t\mu_{0,k}(x)ds + \frac{1}{2}\int_0^t W_2^2(\mu_{s,k},\mathcal{J}_t\mu_{0,k})ds] \leqslant \\ &\leqslant \mathcal{H}_t(\mu_0) + tl + \limsup_{k\to\infty} \frac{1}{2}[\int_0^t \int |b(x)|^2 d\mathcal{J}_t\mu_{0,k}(x)ds + \int_0^t W_2^2(\mu_{s,k}\mathcal{J}_t\mu_{0,k})ds] < +\infty\end{aligned} \tag{4.30}$$

To conclude this section, let us summarize the results we obtained so far in the following theorem. We need a proposition first.

Proposition 4.4 *Let X be a metric space and let $(F_t)_{t\geqslant 0}$ be a contraction semigroup on X with paths of class $AC([0,T];X)$ for each $0 < T < \infty$. Then for each $x \in X$ and for each $t \geqslant 0$ the $\mathcal{L}^1$ a.e. defined metric derivative $|\dot{u}|(t)$ of the curve $u(t) = F_t x$ has (unique nonincreasing) right continuous version $t \mapsto g(t)$. Moreover, $t \mapsto u(t)$ is right metric differentiable for each $t \geqslant 0$ and*

$$\lim_{h\downarrow 0} \frac{d(u(t+h,u(t))}{h} = g(t). \tag{4.31}$$

Proof Fix $x \in X$. Since by assumption the curve $u(t) := F_t x$ is absolutely continuous on each compact interval its metric derivative $|\dot{u}|(t)$ is $\mathcal{L}^1$ a.e. defined on $[0,+\infty)$, hence due to contractivity assumption if $0 \leqslant t_1 \leqslant t_2 < \infty$ are two time points where it is defined we have for each $h \geqslant 0$

$$\frac{d(u(t_1+h),u(t_1))}{h} \geqslant \frac{d(u(t_2+h),u(t_2))}{h}, \tag{4.32}$$

hence $t \mapsto |\dot{u}|(t)$ is $\mathcal{L}^1$ a.e. equal to a nonincreasing function. This implies that there is a unique right continuous nonincreasing function $t \mapsto g(t)$ on $[0,+\infty)$ such that g is a representative of the metric derivative of $t \mapsto u(t)$. Moreover, for each $t \geqslant 0$ and for each $h > 0$, on one hand we have

$$\frac{d(u(t+h),u(t))}{h} \leqslant \int_t^{t+h} g(s)ds,$$

hence

$$\limsup_{h\downarrow 0} \frac{d(u(t+h),u(t))}{h} \leqslant g(t). \tag{4.33}$$

And on the other hand if $t_2 > t$ is such that (4.31) holds with t replaced by t_2, (4.32) yields

$$\liminf_{h\downarrow 0} \frac{d(u(t_2+h),u(t_2))}{h} \geqslant \liminf_{h\downarrow 0} \frac{d(u(t+h),u(t))}{h} = g(t_2) \uparrow g(t) \tag{4.34}$$

for $t_2 \downarrow t$. The claim now follows by (4.33) and (4.34).

Theorem 4.5 *Let b be satisfy (1.2), (1.3), and (1.4). Then for each $\mu_0 \in D(|\partial\mathcal{H}|)$ the curve $t \mapsto \mu(t) := R_t\mu_0$ is the unique curve that is defined on $0,+\infty)$ with values in $\mathcal{P}^2(\mathbb{R}^d)$ such that*

1 $t \mapsto \mu(t)$ is of class $AC^2([0,T];\mathcal{P}^2(mathbbR^d)$ for each $T \geqslant 0$. Moreover it is Lipschitz on $[0,+\infty)$ with Lipschitz constant $\leqslant |\partial\mathcal{H}|(\mu_0)+(\int |b|^2 d\mu_0)^{1/2}$

2 $t \mapsto \mu(t)$ is a solution of equation (1.1) with initial condition $\mu(0) = \mu_0$.

3 $t \mapsto \mu(t)$ is right metrically differentiable and its metric derivative is nonincreasing.

If moreover b is golbally Lipschitz than the semigroup $(R_t)_{t\geqslant 0}$ has the regularizig effect, i.e. for any $\mu_0 \in \mathcal{P}^2(\mathbb{R}^d)$ we have

$$\mu(t) := R_t\mu_0 \in D(|\partial\mathcal{H}|) \tag{5}$$

and $t \mapsto \mu(t)$ is of class $AC^2_{loc}((0,+\infty);\mathcal{P}^2(\mathbb{R}^d)$.

Proof Uniqueness of solutions of (1.1) subject to the condition (1.3) within the class $AC^2_{loc}((0,+\infty);\mathcal{P}^2(\mathbb{R}^d)$ is a simple cosequece of [15] Theorem 1.1 (the suability condition (15) there clearly holds in this case), and as we alreday remarked a simple application of the Ito formula yields that $t \mapsto \mu(t)$ is a solution of (1.1). Now if we show that this curve is Lipschitz on $[0,+\infty)$, we will know that it is metric derivative if defined $\mathcal{L}^1 a.e.$ and since by Proposition 3.8 $(R_t)_t$ is contractive, Proposition 4.4 insures that our curve is right metrically differentiable for each $t \geqslant 0$. Claims concerning the case when b is globally Lipschitz are already proven in Theorem 4.3.

Lets us show that our curve is Lipschitz. To this aim fix $t_1, t_2 \geqslant 0$ such that for some $k,m \in \mathbb{N}$ we have $t_1 = k2^{-m}t_2$. Denoting $k_n := k2^{n-m} \to +\infty$ for $n \in \mathbb{N}$ we have

$$\frac{W_2((R^{1,2}_{t_2 2^{-n}})^{k2^{n-m}}\mu_0, (R^{1,2}_{t_2 2^{-n}})^{2^n}\mu_0)}{t_2 - t_1} \leqslant \frac{1}{t_2 - t_1}\Sigma_{j=k_n}^{2^n-1} W_2((R^{1,2}_{t_2 2^{-n}})^j\mu_0, (R^{1,2}_{t_2 2^{-n}})^{j+1}\mu_0)$$
$$\leqslant \frac{2_n - k_n}{t_2 - t_1} W_2(\mu_0, R^{1,2}_{t_2 2^{-n}}\mu_0) \leqslant \frac{1}{t_2 2^{-n}}(W_2(\mu_0, R^1_{t_2 2^{-n}}\mu_0) + W_2(\mu_0, R^2_{t_2 2^{-n}}\mu_0)) \tag{6}$$

and due to assumption on μ_0 we have $\limsup_{n\to\infty} \frac{1}{t_2 2^{-n}} W_2(\mu_0, R^1_{t_2 2^{-n}}\mu_0) = |\partial\mathcal{H}|(\mu_0)$ as well as

$$\limsup_{n\to\infty} \frac{1}{t_2 2^{-n}} W_2(\mu_0, R^2_{t_2 2^{-n}}\mu_0) \leqslant \int |b|^2 d\mu_0$$

Now by Theorem 3.8 $\lim_n (R^{1,2}_{t_1/k_n})^{k_n}\mu_0 = \mu(t_1)$ and $\lim_n (R^{1,2})^{2^n}\mu_0 = R_{t_2}\mu_0$, our semigroup $(R_t)_{t\geqslant 0}$ is easily seen to be weakly* continuous, and by [1] Proposition 7.1.3 W_2 is sequantially lowersemicontinuos with respect to the weak* convergence we conclude that $t \mapsto \mu(t)$ is Lipschitz with Lipschitz constant $\leqslant |\partial\mathcal{H}|(\mu_0) + (\int |b|^2 d\mu_0)^{1/2}$.

□

5 Invariant measure and asymptotic behavior - symmetric versus non-symmetric case

In the first part of this section investigate the existence of invariant measure and the rate of convergence to it of the flow. In the second part of this section we compare the amount of information that the invariant measure contains about the flow in cases when the drift is a gradient versus when it is just a monotone operator on $\mathbb{R}^d$.

It is well known that SDE (1.5) has a unique invariant measure provided that b is α monotone for $\alpha > 0$, i.e.

$$\langle b(x) - b(y), x - y\rangle \geqslant \alpha|x-y|^2 \quad x, y \in \mathbb{R}^d \tag{5.1}$$

(see for instance [13] Theorem 6.32). However we can obtain existence and uniqueness of invariant measure, for the semigroup R_t by a simple fixed point argument, provided we show that R_t is a strict contraction for $t > 0$. Moreover we than have exponential convergence do the invariant measure with respect to W_2–metric. Uniqueness and exponential rate of convergence has already been established in [19] corollary 1.3 where authors consider class of transportation distances on probability measures. Yet me bring our observations forward too since among other things the authors [14] did not construct the semigroup and in particular do not have existence of invariant measure.

Proposition 5.1 *Assume that b is such that (1.2), (1.3), and (5.1) hold. Then for each $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$ we have*

$$W_2(R_t\mu_1, R_t\mu_2) \leqslant e^{-\alpha t} W_2(\mu_1, \mu_2) \tag{5.2}$$

for all $t \geqslant 0$.

Proof First of all notic that (5.1) implies that for $x, y \in \mathbb{R}^d$

$$|S_t x - S_t y|^2 \leqslant e^{-\alpha t}|x-y|^2, \tag{5.3}$$

i.e., $(S_t)_{t\geqslant 0}$ is an α-contraction on $\mathbb{R}^d$. Next fix $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$ and estimate for $t \geqslant 0$, $n \in \mathbb{N}$, and $\gamma \in \Gamma_0(\mu_1, \mu_2)$,

$$\begin{aligned} W_2^2(R_t^1\mu_1, R_2^1\mu_2) \leqslant & \int\int |S_{t/n}x - S_{t/n}y|^2 \, d\gamma \\ e^{-2\alpha t/n} & \int\int |x-y|^2 \, d\gamma = e^{-2\alpha t/n} W_2^2(\mu_1, \mu_2) \end{aligned} \tag{5.4}$$

and then we have that

$$W_2(R_{t/n}^{2,1}\mu_1, R_{t/4}^{2,1}\mu_2) \leqslant e^{-\alpha t/n} W_2(\mu_0, \mu_1). \tag{5.5}$$

Now the claim follows easily by induction and passing to the limit:

$$\begin{aligned} W_2((R_{t/n}^{2,1})^n\mu_1, (R_{t/n}^{2,1})^n\mu_2) &\leqslant (e^{-\alpha t/n})^n W_2(\mu_1, \mu_2) \\ &= e^{-\alpha t} W_2(\mu_1, \mu_2) \end{aligned} \tag{5.6}$$

by Theorem 3.8.

□

Now since for $t > 0$ R_t is a strict contraction due to the banach fixed poit theorem it has a unique fixed point $\mu_{\infty,t}$. Let us show that for $s, t \geqslant 0$, $\mu_{s,\infty} = \mu_{t,\infty}$. Well if $s = kt$ for $k \in \mathbb{N}$ than $R_s\mu_{t,\infty} = (R_t)^k \mu_{t,\infty} = \mu_{t,\infty}$, hence as R_s can have only one invariant measure, we must have $\mu_{s,\infty} = \mu_{t,\infty}$. Same kind of argument yealds that $\mu_{s,\infty} = \mu_{t,\infty}$ if $t = ks$ for $k \in \mathbb{N}$, so that if $t/s \in \mathbb{Q}$ we must have

$$\mu_{s,\infty} = \mu_{t,\infty} \tag{111}$$

as well. Finally since all paths of $(R_t)_{t\geqslant 0}$ are continuous by Theorem 4.5, (111) follows for all $s, t \geqslant 0$, i.e. the semigroup $(R_t)_{t\geqslant 0}$ has a unique invariant measure which we denote by μ_∞. Since $t \mapsto R_{t\mu_\infty}$ satisfies (12), and by Proposition 4.5 $\mu_\infty = R_t\mu_\infty \in D(|\partial\mathcal{H}|)$, i.e. $\mu_\infty = \rho_\infty \cdot \mathcal{L}^d$ $\rho_\infty \in W^{1,1}(\mathbb{R}^d)$, $\frac{\nabla\rho_\infty}{\rho_\infty} \in L^2(\mu_\infty; \mathbb{R})$ and the continuity equation holds, i.e.

$$\partial_t\mu_\infty + \nabla \circ \left(- \left(\frac{\nabla\rho_\infty}{\rho_\infty} + P_{Tan_{\mu_\infty}} b \right) \mu_\infty \right) = -\nabla \circ \left(\frac{\nabla\rho_\infty}{\rho_\infty} + P_{Tan_{\mu_\infty}} b \right) \mu_\infty \tag{5.9}$$

Since invariant measure exists and is unique we have weak existence and uniqueness of the eliptic problem

$$\int \langle \nabla\varphi, \left(\frac{\nabla S_\infty}{S_\infty} + b \right) \rangle S dx = 0 \quad \forall\varphi \in C_c^\infty(\mathbb{R}^d) \tag{5.10}$$

(within the class of manegdive L^1 functions in $L^1(dx)$ and the solution to (5.10) denoted ρ_∞ as is the density of the invariant measure for SDE (1) (and our semigroup $(R_t)_{t\geqslant 0}$) If b is a gradient then paths of R_t are gradient flow paths, for the entropy functional

$$\mathcal{H}(\mu|\mu_\infty) = \begin{cases} \int \frac{d\mu}{d\mu_\infty} \log \frac{d\mu}{d\mu_\infty} & \text{if } \mu \ll \mu_\infty \\ +\infty & \text{otherwise} \end{cases} \tag{5.11}$$

(see for instance [15] or [**?**], and μ_∞ contains all information about the flow, as R_t in the limit of the variational minimization scheme. However, if b is not a gradient this is not the case as we are about to explain. Let us assume that $b = A$ ia a matrix in $M^d \cong \mathbb{R}^{d^2}$, and let it be strict positive definite but not symmetric (recall that a positive matrices is a gradient iff it is symmetric and if it is symmetric that $A = \nabla\psi$ for the convex C^∞ function $\psi(x) := \frac{1}{2}\langle Ax, x\rangle$). Such matrix always satisfies (5.1). Moreover [13] section (6.2.1) and [16] provide an explicit form of μ_∞. It is described as follows: define for $t > 0$

$$Q_t := \int_0^t S_s S_s^* ds \tag{5.12}$$

a positive definite symmetric matrix. Then for any $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ $R_t\mu_0 = \mu_0 \star \mathcal{N}(0, Q_t)$ where $\mathcal{N}(0.Q_t)$ is the d-dimensional Gaussian measure with mean 0 and covariance matrix Q_t. Moreover the invariant measure is $\mu_\infty = W(0, Q_\infty)$ where

$$Q_\infty = \int_0^\infty S_t S_t^* dt \tag{5.13}$$

By (5.10) we have for $\varphi \in C_c^\infty(\mathbb{R}^d)$

$$0 = \int \langle \nabla\varphi(x)(-Q_\infty^{-1} + A)(x)\rangle \frac{e^{-\langle Q_\infty^{-1}x,x\rangle/2}}{\sqrt{(2\pi)^d det\ Q_\infty}} dx \tag{5.14}$$

which amounts to

$$P_{Tan_{\mathcal{N}(0,Q_\infty)}} A = Q_\infty^{-1} \tag{5.15}$$

Partial integration in (5.14) gives (and we write $\rho_\infty(x) = \frac{1}{\sqrt{(2\pi)^d det\ Q_\infty}} e^{-\langle Q_\infty^{-1}x,x\rangle/2}$

$$0 = \int \varphi(x)(\operatorname{tr} A - \operatorname{tr} Q_\infty^{-1})\rho_\infty(x)dx - \int \varphi(x)\langle Q_\infty^{-1}x, (A - Q_\infty^{-1})x\rangle \rho_\infty(x)dx \tag{5.16}$$

for all $\varphi \in C_c^\infty(\mathbb{R}^d)$ so that

$$\langle Q_\infty^{-1}x, (A - Q_\infty^{-1})x\rangle = \operatorname{tr} A - \operatorname{tr} Q_\infty^{-1} \quad \forall x \in \mathbb{R}^d \tag{5.17}$$

must hold. Evaluating the left side in (5.13) for $x = cy$, $c \in \mathbb{R}$ implies

$$\operatorname{tr} A = \operatorname{tr} Q_\infty^{-1} \quad \langle Q_\infty^{-1}x, (A - Q_\infty^{-1})x\rangle \equiv 0 \quad \text{ on } \mathbb{R}^d \tag{5.18}$$

and if (5.18) holds then (5.15) and (5.10) hold for $b = A$ and $\rho_\infty \cdot \mathcal{L}^d$ is the invariant measure. Recall the $\langle\cdot,\cdot\rangle_C$ inner product on $M^d \cong \mathbb{R}^{d^2}$ for a matrix $C \in V_t = \{\text{symmetric positive definite matrices}\}$ defined by

$$\langle U, V\rangle_C := \operatorname{tr} C^{-1}UV \text{ for } U, V \in M^d \tag{5.19}$$

and the C inner product on $\mathbb{R}^d$ defined by

$$\langle x, y\rangle_C := \langle C^{-1}x, y\rangle \text{ for } x, y \cong\in \mathbb{R}^d \tag{5.20}$$

We have the following

Proposition 5.2 *Let $A \in M_t^d = \{$strictly positive definite matrices on $\mathbb{R}^d\}$ and denote the corresponding invariant measure be $\mathcal{N}(0, Q_\infty) = \mu_\infty$ of (1.5) M with Q_∞ defined by (5.13). Then $\tilde{A} := P_{Tan_{\mu_\infty}} A$ equals the orthogonal projection $P^{V_+}_{\langle\cdot,\cdot\rangle_{Q_\infty}} A$ of A onto the positive cone $V_+ \cup \{0\}$ w.r.t. the inner product $\langle\cdot,\cdot\rangle_{Q_\infty}$.*

Proof Denote $\tilde{A}_1 := P^{V_+}_{\langle\cdot,\cdot\rangle_{Q_\infty}} A$. We will show that $\tilde{A}_1 = \tilde{A}$. By definition $0 = \operatorname{tr} Q_\infty^{-1}(A - \tilde{A}_1)D^* = \operatorname{tr} Q_\infty^{-1}(A - \tilde{A}_1)D$ for each $D \in V = \{\text{symmetric matrices on } \mathbb{R}^d\}$ which holds if and only if $Q_\infty^{-1}(A - \tilde{A}_1)$ is antysymmetric i.e.

$$Q_\infty^{-1}(A - \tilde{A}_1) = -(A^* - \tilde{A}_1^*) = (A^* - \tilde{A}_1)(Q_\infty^{-1})^* \tag{5.21}$$

Define $\psi(x) := \langle Q_\infty^{-1}x, (A - \tilde{A}_1)x\rangle$, and observe that

$$\nabla\psi(x) = Q_\infty^{-1}(A - \tilde{A}_1)x + (A^* - \tilde{A}_1)Q_\infty^{-1}x = 0 \quad \forall x \in \mathbb{R}^d, \tag{5.22}$$

and as $\psi(0) = 0$ we must have

$$\langle Q_\infty^{-1}x, (A - \tilde{A}_1)x\rangle = 0 \quad \forall x \in \mathbb{R}^d \tag{5.23}$$

Furthermore as Q_∞ is an element of $V^+ \cup \{0\}$ too $0 = \operatorname{tr} Q_\infty^{-1}(A - \tilde{A}_1)Q_\infty = \operatorname{tr}(A - \tilde{A}_1)$, which together with (5.23) amounts to (5.18). Hence (5.15) holds and we have proved the claim.

□

Remark Arguing in exactly the same way we can show that for any two ... measure $\mathcal{N}(0, Q) = \mu_0$ and $\mathcal{N}(a, Q) = \mu_1$ $P_{Tan_{\mu_0}} A = P^{V_+}_{\langle\cdot,\cdot\rangle_Q} A$ and $P_{Tan_{\mu_1}} A = P_{Tan_{\mu_0}}(A - P_{Tan_{\mu_0}} A)a$

Now we are able to explain why the invariant measure $\mu_\infty = \mathcal{N}(0, Q_\infty)$ does not contain all the information about the flow $(R_t)_{t\geqslant 0}$ if A is not a gradient. Indeed suppose that $A' = Q_\infty^{-1} + B$ is a positive matrix such that $B \in V^\perp$ where the orthogonal complement is taken with respect to $\langle\cdot,\cdot\rangle_{Q_\infty}$. Then by the previous analysis $Q_t\mu_t = \Delta\mu_t + \nabla\cdot(A\mu_t)$ and $Q_t\mu_t = \Delta\mu_t + \nabla\cdot(A'\mu_t)$ both have some invariant measure $\mathcal{N}(0, Q_\infty)$ (A' satisfies (5.1) if it is satisfies positive definite). In order to see that there are many such $A' \in M^+$ notice firstly that M^+ is a positive cone; and secondly that setting for $U \in M^+$ $\alpha_U := \inf_{|x|=1}\langle Ux, x\rangle$ so that $\langle Ux, x\rangle \geqslant \alpha_U|x|^2$ for $x \in \mathbb{R}^d$. Now given any $A \in M^+$ if $B \in M^d$ is any matrix s.t. $|A - B| < \alpha_A$ then $\langle Bx, x\rangle = \langle Ax, x\rangle + \langle (B - A)x, x\rangle \geqslant \alpha_A|x|^2 - |B - A||x|^2$ hence $B \in M^+$; i.e. M^+ is an open set for the standard inner product on $\mathbb{R}^{d^2}$, and then by the equivalence of measures on finite dimensional spaces for any inner product. Now pick any $D \in V_+^\perp$, the orthogonal complement being take with respect to $\langle\cdot,\cdot\rangle_{Q_\infty}$, such that $|D|_{Q_\infty} < \alpha_{Q_\infty^{-1}}$. Then we have that $A' := Q_\infty^{-1} - D \in M^+ \cap V_+^\perp$.

References

[1] L. Amrosio, N. Gigli, G. Savaré, "Gradient flows in metric spaces and in the space of probability measures".Second edition. Lectures in Mathematics ETH Zurich.Birkhuser Verlag, Basel, 2008. x+334 pp. ISBN: 978-3-7643-8721-1

[2] L. Ambrosio, G. Savaré, Zamboti, "Existence and stability for Fokker-Planck equations with log–concave reference measure", *Probab. Theory* **145** (2009), 517–564.

[3] V. I. Bogachev, G. Da Prato, M. Röckner, "On regularity of transition probability and invariant measures of singular diffusion under minimal conditions", *Commun. Partial Differ. Equations* **26** (2001), 2037-2080.

[4] V. I. Bogachev, G. Da Prato, M. Röckner, W. Stannob, "Uniqueness of solution to weak equations for measures", *Bull. Lond. Math. Soc.* **39** (2007), 631-640.

[5] V. I. Bogachev, G. Da Prato, M. Röckner, "On parabolic equatons for measures", *Commun. Partial Differ. Equations*, **33** (2008), 397–418.

[6] Brezis, H. Oprateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert' (French). North-Holland Mathematics Studies, No. 5. Notas de Matemtica (50).North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973. vi+183 pp.

[7] Burago, D; Burago, Y; Ivanov, S A course in metric geometry. Graduate Studies in Mathematics, 33. American Mathematical Society, Providence, RI, 2001. xiv+415 pp. ISBN: 0-8218-2129-6 53C23

[8] P. Clement, "An Introduction to Gradent Flows on Metric spaces" Leiden University, MI Report 2009.

[9] Da Prato, G.(I-SNS); Zabczyk, J.(PL-PAN) Ergodicity for infinite-dimensional systems.London Mathematical Society Lecture Note Series, 229. Cambridge University Press, Cambridge, 1996 xii+339 pp. ISBN: 0-521-57900-760H15

[10] Engel, K-J; Nagel, R. One-parameter semigroups for linear evolution equations. Graduate Texts in Mathematics, 194. Springer-Verlag, New York, 2000. xxii+586 pp. ISBN: 0-387-98463-1

[11] J. Jacod, A.N. Shiryaev. Limit theorems for stochastic processes. Second edition.Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 288. Springer-Verlag, Berlin, 2003. xx+661 pp. ISBN: 3-540-43932-3

[12] Jordan, Richard; Kinderlehrer, David; Otto, Felix The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. 29 (1998), no. 1, 1–17.

[13] T. Kato, K. Masuda, "Trotter's product formula for nnonlinear semigroups generated by subdifferentials of convex functionals, *J. Math. Soc. Japan* **30** (1978).

[14] Lorenzi, L; Bertoldi, M. Analytical methods for Markov semigroups. Pure and Applied Mathematics (Boca Raton), 283. Chapman & Hall/CRC, Boca Raton, FL, 2007. xxxii+526 pp. ISBN: 978-1-58488-659-4; 1-58488-659-5

[15] L. Natile, G. M. Peletier, G. Savere "Contraction of general transportation costs along solutions to Fokker-Planck equations with monotone drifts",preprint 2010.

[16] L. Natile, G. Savaré preprint 2010.

[17] Otto, F. The geometry of dissipative evolution equations: the porous medium equation. Comm. Partial Differential Equations 26 (2001), no. 1-2, 101–174.

[18] Rudin, W. Real and complex analysis. Second edition. McGraw-Hill Series in Higher Mathematics. McGraw-Hill Book Co., New York-Dsseldorf-Johannesburg, 1974. xii+452 pp.

[19] Zakai, M; Snyders, J. Stationary probability measures for linear differential equations driven with white noise. J. Differential Equations 8 1970 27–33.

[20] Rogers, L. C. G.; Williams, David Diffusions, Markov processes, and martingales. Vol. 1. Foundations. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. xx+386 pp. ISBN: 0-521-77594-9

[21] Rogers, L. C. G.; Williams, D. Diffusions, Markov processes, and martingales. Vol. 2. It calculus. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. xiv+480 pp. ISBN: 0-521-77593-0

[22] Tudorascu, A. On the Jordan-Kinderlehrer-Otto variational scheme and constrained optimization in the Wasserstein metric. Calc. Var. Partial Differential Equations 32 (2008), no. 2, 155–173.

[23] Villani, C. Topics in optimal transportation. Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, RI, 2003. xvi+370 pp. ISBN: 0-8218-3312-X 90-02

[24] Villani, C. Optimal transport. Old and new. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 338 Springer-Verlag, Berlin, 2009. xxii+973 pp. ISBN: 978-3-540-71049-3