Existence and stability for Fokker–Planck equations with log-concave reference measure

**Luigi Ambrosio · Giuseppe Savaré ·
Lorenzo Zambotti**

Abstract   We study Markov processes associated with stochastic differential equations, whose non-linearities are gradients of convex functionals. We prove a general result of existence of such Markov processes and a priori estimates on the transition probabilities. The main result is the following stability property: if the associated invariant measures converge weakly, then the Markov processes converge in law. The proofs are based on the interpretation of a Fokker–Planck equation as the steepest descent flow of the relative entropy functional in the space of probability measures, endowed with the Wasserstein distance.

Keywords   Reversible Markov processes · Log-concave probability measures · Gradient flows · Optimal transportation · Relative entropy

Mathematics Subject Classification (2000)   60J60 · 37L40 · 60G07 · 49Q20 · 35K90 · 28A33

L. Ambrosio
Scuola Normale Superiore, Pisa, Italy
e-mail: l.ambrosio@sns.it

G. Savaré
Dipartimento di Matematica, Università di Pavia, Pavia, Italy
e-mail: giuseppe.savare@unipv.it

L. Zambotti (✉)
Université Pierre et Marie Curie,
LPMA, 4 place Jussieu, 75252 Paris cedex 05, France
e-mail: lorenzo.zambotti@upmc.fr

1 Introduction and main results

In the seminal paper [24], Jordan–Kinderlehrer–Otto have given a remarkable interpretation of the solution to a linear Fokker–Planck equation as the steepest descent flow of the relative Entropy functional in the space of probability measures, endowed with the Wasserstein distance. The book [3] by Ambrosio–Gigli–Savaré has provided a general theory of gradient flows in the Wasserstein space of probability measures, including linear and non-linear PDE's, in finite and infinite dimensions.

In this paper we want to investigate the probabilistic counterpart of such results. The approach is analytical and based on techniques from calculus of variations and optimal transport problems; however, several results have important consequences on existence and in particular convergence of Markov processes being reversible with respect to a *log-concave* probability measure.

Following the ideas of [3,4], we interpret the solution $(\mu_t)_{t \geq 0}$ of a Fokker–Planck equation with convex potential, as a curve in the space of probability measures, solving a suitable differential equation (or, more precisely, a *differential variational inequality*). We obtain interesting estimates on μ_t which have, to our knowledge, no direct probabilistic proof, and are very useful in the study of the stochastic process $(X_t)_{t \geq 0}$ whose one-time distributions are $(\mu_t)_{t \geq 0}$.

1.1 The main results

We consider a separable Hilbert space H, which could be finite or infinite dimensional, whose scalar product and norm will be respectively denoted by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$. We denote by $\mathscr{P}(H)$ the set of all probability measures on H.

We consider a probability measure γ on H with the following property:

Assumption 1.1 γ is *log-concave*, i.e. for all pairs of open sets B, $C \subset H$

$$\log \gamma \left((1-t)B + tC \right) \geq (1-t) \log \gamma(B) + t \log \gamma(C) \qquad \forall \, t \in (0, 1), \qquad (1.1)$$

where $(1 - t)B + tC := \{(1 - t)b + tc, b \in B, c \in C\}$.

The class of log-concave probability measures includes all measures of the form (here $\mathscr{L}^k$ stands for Lebesgue measure)

$$\gamma := \frac{1}{Z} e^{-V} \mathscr{L}^k, \quad \text{where } V : H = \mathbb{R}^k \to \mathbb{R} \text{ is convex and } Z := \int_{\mathbb{R}^k} e^{-V} \, dx < +\infty,$$
$$(1.2)$$

all Gaussian measures, all Gibbs measures on a finite lattice with convex Hamiltonian. Borell [8] has proved that if γ is log-concave on H then there exists $\varepsilon > 0$ such that

$$\int e^{\varepsilon \|x\|} \gamma(dx) < +\infty. \qquad (1.3)$$

See Proposition 2.2 and the Appendix for more information on the class of log-concave probability measures.

We denote the support of γ by $K = K(\gamma)$ and the smallest closed affine subspace of H containing K by $A = A(\gamma)$. We write canonically

$$A = H^0 + h^0, \quad h^0 \in A, \quad \|h^0\| \leq \|k\| \quad \forall \, k \in A, \tag{1.4}$$

so that $h^0 = h^0(\gamma)$ is the element of minimal norm in A and $H^0 = H^0(\gamma)$ is a closed linear subspace of H. As in the Gaussian case, we will say that γ in *non-degenerate* if $H^0(\gamma) = H$.

We want to consider a stochastic processes with values in $A(\gamma)$ and reversible with respect to γ. We now state a first result which determines such process in a canonical way. We denote by $C_b(H)$ the space of bounded continuous functions in H and by $C_b^1(A(\gamma))$ the space of all $\Phi : A(\gamma) \mapsto \mathbb{R}$ which are bounded, continuous and Fréchet differentiable with bounded continuous gradient $\nabla \Phi : A(\gamma) \mapsto H^0(\gamma)$ (notice that all functions in $C_b^1(A(\gamma))$ are Lipschitz continuous).

We set $\Omega := C([0, +\infty[; K) \subset K^{[0,+\infty[}$, and we denote by $X_t : K^{[0,+\infty[} \to K$ the coordinate process $X_t(\omega) := \omega_t$, $t \geq 0$. We shall endow Ω with the Polish topology of uniform convergence on bounded subsets of $[0, +\infty[$, and the relative Borel σ-algebra. On $K^{[0,+\infty[}$ we shall consider the canonical σ-algebra generated by cylindrical sets and, for probability measures in $K^{[0,+\infty[}$, the convergence induced by the duality with continuous cylindrical functions of the form $f(X_{t_1}, \ldots, X_{t_n})$, with $f \in C_b(K^n)$.

Theorem 1.2 (Markov process and Dirichlet form associated to γ) *Let γ be a log-concave probability measure on H and let K be its support. Then*:

(a) *The bilinear form $\mathcal{E} = \mathcal{E}_{\gamma, \|\cdot\|}$ given by*

$$\mathcal{E}(u, v) := \int_K \langle \nabla u, \nabla v \rangle_{H^0(\gamma)} \, d\gamma, \quad u, v \in C_b^1(A(\gamma)), \tag{1.5}$$

is closable in $L^2(\gamma)$ and its closure $(\mathcal{E}, D(\mathcal{E}))$ is a symmetric Dirichlet Form. Furthermore, the associated semigroup $(P_t)_{t \geq 0}$ in $L^2(\gamma)$ maps $L^\infty(\gamma)$ in $C_b(K)$.

(b) *There exists a unique Markov family $(\mathbb{P}_x : x \in K)$ of probability measures on $K^{[0,+\infty[}$ associated with $\mathcal{E}$. More precisely, $\mathbb{E}_x[f(X_t)] = P_t f(x)$ for all bounded Borel functions and all $x \in K$. Moreover, $x \mapsto \mathbb{P}_x$ is continuous.*

(c) *For all $x \in K$, $\mathbb{P}_x^*(C(]0, +\infty[; H)) = 1$ and $\mathbb{E}_x[\|X_t - x\|^2] \to 0$ as $t \downarrow 0$. Moreover, $\mathbb{P}_x^*(C([0, +\infty[; H)) = 1$ for γ-a.e. $x \in K$.*

(d) *$(\mathbb{P}_x : x \in K)$ is reversible with respect to γ, i.e. the transition semigroup $(P_t)_{t \geq 0}$ is symmetric in $L^2(\gamma)$; moreover γ is invariant for (P_t), i.e. $\gamma(P_t f) = \gamma(f)$ for all $f \in C_b(K)$ and $t \geq 0$.*

In (c), $\mathbb{P}_x^*$ denotes the outer measure defined by $\mathbb{P}_x$ on all subsets of $K^{[0,+\infty[}$.

An example in $H := \mathbb{R}^k$ of the above setting is provided by (1.2) when the potential $V : \mathbb{R}^k \to \mathbb{R}$ is convex with Lipschitz continuous gradient $\nabla V : \mathbb{R}^k \to \mathbb{R}^k$. Then γ

is log-concave, see Proposition 2.2, and the process X is a solution of the stochastic differential equation (SDE):

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2}\,dW, \qquad X_0(x) = x, \tag{1.6}$$

where W is an $\mathbb{R}^k$-valued Brownian motion. One can also consider a convex $V \in C^{1,1}(U)$, set of $V \in C^1(U)$ with Lipschitz gradient, where $U \subset \mathbb{R}^k$ is a smooth convex open set, and $V \equiv +\infty$ on $\mathbb{R}^k \setminus U$. Then X solves the SDE with reflection at the boundary ∂U of U:

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2}\,dW + \mathbf{n}(X_t)\,dL_t, \qquad X_0(x) = x, \tag{1.7}$$

where $\mathbf{n}$ is an inner normal vector to ∂U and L is a continuous monotone non-decreasing process which increases only when $X_t \in \partial U$. Equations like (1.6) and (1.7) with convex potentials arise in the theory of random interfaces. The invariant measure γ is typically a Gibbs measure on a lattice. Interesting infinite-dimensional examples include stochastic PDEs. See Sect. 1.2 for an overview of the literature.

Before stating the next theorem, we define the relative entropy functional; for all probability measures μ on H we set:

$$\mathcal{H}(\mu|\gamma) := \int_H \rho \, \log \rho \, d\gamma \tag{1.8}$$

if $\mu = \rho \, \gamma$ for some $\rho \in L^1(\gamma)$, and $+\infty$ otherwise. We recall that $\mathcal{H}(\cdot|\gamma) \geq 0$ by Jensen's inequality.

We also define the Wasserstein distance: given two probability measures μ, ν on H, we set

$$W_2(\mu, \nu) := \inf \left\{ \left[\int_{H \times H} \|y - x\|^2 \, d\Sigma \right]^{\frac{1}{2}} : \Sigma \in \Gamma(\mu, \nu) \right\}. \tag{1.9}$$

Here $\Gamma(\mu, \nu)$ is the set of all *couplings* between μ and ν: it consists of all probability measures Σ on $H \times H$ whose first and second marginals are respectively, μ and ν, i.e. $\Sigma(B \times H) = \mu(B)$ and $\Sigma(H \times B) = \nu(B)$ for all $B \in \mathscr{B}(H)$. We set

$$\mathscr{P}_2(H) := \left\{ \mu \in \mathscr{P}(H) : \int_H \|x\|^2 \, d\mu(x) < \infty \right\}.$$

It turns out that $W_2(\cdot, \cdot)$ is a distance on $\mathscr{P}_2(H)$ and that $(\mathscr{P}_2(H), W_2)$ is a complete and separable metric space, whose convergence implies weak convergence, see for instance [3, Proposition 7.1.5]. Then, we have the following result:

Theorem 1.3 (Estimates on transition probabilities) *Let γ be a log-concave probability measure on H and let X be as in Theorem 1.2. Fix $x \in K$ and denote the law of X_t under $\mathbb{P}_x$ by ν_t^x. Then $[0, +\infty[\times K \ni (t, x) \mapsto \nu_t^x \in \mathscr{P}_2(H)$ is continuous and*

$$\mathcal{H}(\nu_t^x | \gamma) \leq \inf_{\sigma \in \mathscr{P}_2(H)} \left\{ \frac{1}{2t} \int_H \|y - x\|^2 \, d\sigma(y) + \mathcal{H}(\sigma | \gamma) \right\} < +\infty \quad \forall t > 0,$$

so that $\nu_t^x \ll \gamma$ for all $t > 0$, $x \in K$. Moreover,

$$W_2(\nu_t^x, \nu_s^x) \leq \sqrt{2 \mathcal{H}(\nu_\varepsilon^x | \gamma)} \sqrt{|t - s|}, \quad t, s \geq \varepsilon, \ x \in K.$$

Notice that the estimates given in Theorem 1.3 do not contain any constant depending on H or on γ and appear to be of a structural nature.

We consider now a sequence (γ_n) of log-concave probability measures on H such that γ_n converge weakly to γ. We denote $K_n := K(\gamma_n)$, $A_n := A(\gamma_n)$, $H_n := H^0(\gamma_n)$, in the notation of (1.4). We want to consider situations where each H_n is an Hilbert space endowed with a scalar product $\langle \cdot, \cdot \rangle_{H_n}$ and an associated H-continuous norm $\|\cdot\|_{H_n}$ possibly different from the scalar product and the norm induced by H. In order to ensure that this family of norms converges (in a suitable sense) to the norm of H as $n \to \infty$, we will make the following assumption:

Assumption 1.4 There exists a constant $\kappa \geq 1$ such that

$$\frac{1}{\kappa} \|h\|_H \leq \|h\|_{H_n} \leq \kappa \|h\|_H \quad \forall h \in H_n, \ n \in \mathbb{N}. \tag{1.10}$$

Furthermore, denoting by $\pi_n : H \to H_n$ the orthogonal projections induced by the scalar product of H, we have

$$\lim_{n \to \infty} \|\pi_n(h)\|_{H_n} = \|h\|_H \quad \forall h \in H. \tag{1.11}$$

This assumption guarantees in some weak sense that the geometry of H_n converges to the geometry of H; the case when all the scalar products coincide with $\langle \cdot, \cdot \rangle_H$, $H_n \subset H_{n+1}$ and $\cup_n H_n$ is dense in H is obviously included and will play an important role in the paper.

Let $(\mathbb{P}_x^n : x \in K_n)$ (respectively $(\mathbb{P}_x : x \in K)$) be the Markov process in $[0, +\infty[^{K_n}$ associated to γ_n (resp. in $[0, +\infty[^K$ associated to γ) given by Theorem 1.2. We denote by $\mathbb{P}_{\gamma_n}^n := \int \mathbb{P}_x^n \, d\gamma_n(x)$ (resp. $\mathbb{P}_\gamma := \int \mathbb{P}_x \, d\gamma(x)$) the associated stationary measures.

With an abuse of notation, we say that a sequence of measures $(\mathbf{P}_n)$ on $C([a, b]; H)$ converges weakly in $C([a, b]; H_w)$ if, for all $m \in \mathbb{N}$ and $h_1, \ldots, h_m \in H$, the process $(\langle X_\cdot, h_i \rangle_H, i = 1, \ldots, m)$ under $(\mathbf{P}_n)$ converges weakly in $C([a, b]; \mathbb{R}^m)$ as $n \to \infty$.

In this setting we have the following stability and tightness result:

Theorem 1.5 (Stability and tightness) *Suppose that γ_n converge weakly to γ and that the norms of H_n satisfy Assumption 1.4. Then, for all $x_n \in K_n$ such that $x_n \to x \in K$ in H:*

(a) $\mathbb{P}^n_{x_n} \to \mathbb{P}_x$ *weakly in* $H^{[0,+\infty[}$ *as* $n \to \infty$;

(b)   *for all* $0 < \varepsilon \le T < +\infty$, $\mathbb{P}^n_{x_n} \to \mathbb{P}_x$ *weakly in* $C([\varepsilon, T]; H_w)$;

(c)   *for all* $0 \le T < +\infty$, $\mathbb{P}^n_{\gamma_n} \to \mathbb{P}_\gamma$ *weakly in* $C([0, T]; H_w)$.

This stability property means that the weak convergence of the invariant measures γ_n and a suitable convergence of the norms $\|\cdot\|_{H_n}$ to $\|\cdot\|_H$ imply the convergence in law of the associated processes, starting from *any* initial condition. Notice also statement (b) makes sense, because Theorem 1.2(c) gives that our processes have continuous modifications in $C(]0, +\infty[; H)$ (however, we are able to prove tightness only for the weak topology of H).

Finally, our approach yields naturally the following.

Theorem 1.6 (Uniqueness in $\mathscr{P}_2(H)$ of the invariant measure) *If $\mu \in \mathscr{P}_2(H)$ is an invariant measure of $(P_t)_{t \ge 0}$, i.e. $\mu(P_t f) = \mu(f)$ for all $f \in C_b(K)$ and $t \ge 0$, then $\mu = \gamma$.*

Notice that necessarily a log-concave γ belongs to $\mathscr{P}_2(H)$, by (1.3).

1.2 Motivations and a survey of the literature

Existence and uniqueness for stochastic equations like (1.6) and (1.7) in finite dimension are classical problems in probability theory, starting from [32] and [35]. In [12], existence and uniqueness of strong solutions are proved for general convex potential V. For non-convex domains see [26]. The Dirichlet form approach is detailed in [18].

Natural generalizations of (1.6) to the infinite dimension are provided by stochastic partial differential equations (SPDEs): see Chap. 8 of [15] and [14]. SPDEs with reflection, which generalize (1.7), have also been studied: see [13,28,37,38]. Unlike the finite-dimensional case, no general result of existence and uniqueness is known, and in fact it is not even clear how to define a general definition of solution.

The main result of this paper is the general stability property of this class of stochastic processes, given by Theorem 1.5: if the log-concave invariant measures γ_n converge, then the laws $\mathbb{P}^n_x$ of the associated stochastic processes also converge. In order to appreciate the strength of this result, notice that convergence of γ_n is a much weaker information than convergence (in any sense) of the drift ∇V_n in (1.6). In fact, every approach based either on the SDE or on the generator and the Dirichlet form associated with the process, seems bound to give only weaker results.

In the stability result, the limit process is identified by the associated Dirichlet form (1.5): however, in the general case, we can not write a stochastic equation for the limit, although this can be (and has been) done in many interesting situations. Our approach yields existence of stochastic processes associated with any Dirichlet form of the gradient type (1.5) with log-concave reference measure: this also seems to be a new result (see [1]).

Stochastic equations of the form (1.6) and (1.7) are used as models for the random evolution of interfaces; in these cases the invariant measure is typically a Gibbs measure on a lattice with convex interaction: see [21,33] and [19] for the physical background.

In many interesting cases, the Gibbs measure converges, under a proper rescaling, to a non-degenerate Gaussian (or related) measure on some function or distribution space. Convergence in law of the associated stationary dynamics to the solution of an infinite dimensional stochastic equation is interpreted as convergence of the equilibrium fluctuations of the interface around its macroscopic hydrodynamic limit: see [20,22].

Such convergence results are obtained only in the stationary case and the proofs use very particular properties of the model. For instance, the techniques of [20] are based on monotonicity properties and can not be applied to many interesting situations. Our Theorem 1.5 extends the convergence result to more general initial conditions and is based only on the log-concavity of the probability measures: see [41,42] for some recent applications. For a different (and weaker) approach based on infinite dimensional integration by parts, see [39,40].

Finally, we notice that log-concave measures are still widely used as models for random interfaces: see [31] and references therein.

1.3 Plan of the paper

We conclude this introduction with a short description of the plan of the paper: Sect. 2 is devoted to the introduction of some basic concepts and terminology, while in Sect. 3 we illustrate the model case when $H = \mathbb{R}^k$ and ∇V is smooth, bounded and Lipschitz: here almost no technical issue arises and the basic heuristic ideas can be presented much better. In Sect. 4 we show the basic convexity properties of the relative Entropy functional needed to build in Sect. 5, by implicit time discretization, a "Fokker–Planck" semigroup in the Wasserstein space of probability measures. Section 6 is devoted to the quite strong stability properties of this semigroup, and these are used in Sect. 7 to establish, starting from the smooth case, the link with Dirichlet forms. Finally, in Sect. 8 we canonically build our process in $K^{[0,+\infty[}$, and deduce its continuity properties from the continuity properties of its transition probabilities, provided by the Wasserstein semigroup. Finally, we adapt to our case some general results from [27] on the existence of Markov processes associated to Dirichlet forms to obtain the results stated in Theorem 1.2(c).

2 Notation and preliminary results

In this section we fix our main notation and recall the main results on Wasserstein distance and optimal couplings.

Throughout the paper we consider a real separable Hilbert space H. For $J \subset H$ closed we denote by $\mathrm{Lip}_b(J)$ the space of all bounded $\varphi : J \mapsto \mathbb{R}$ such that:

$$[\varphi]_{\mathrm{Lip(J)}} := \sup \left\{ \frac{|\varphi(x) - \varphi(y)|}{\|x - y\|} \ : \ x, \, y \in J, \ x \neq y \right\} \ < \ +\infty.$$

Measure-theoretic notation. If H is a separable Hilbert space, we shall denote by $\mathscr{B}(H)$ the Borel σ-algebra of H, and by $\mathscr{P}(H)$ the set of (Borel) probability measures

in H. Given a Borel map $r : H \to H$, the *push forward* $r_{\#}\mu \in \mathscr{P}(H)$ of $\mu \in \mathscr{P}(H)$ is defined by $r_{\#}\mu(B) := \mu(r^{-1}(B))$ for all $B \in \mathscr{B}(H)$.

The set of non-degenerate Gaussian measures on H, which all belong to $\mathscr{P}_2(H)$, will be denoted by $G(H)$. Analogously, we shall denote by $\mathscr{B}_G(H)$ the σ-ideal of *Gaussian* null sets, *i.e.* the sets $B \in \mathscr{B}(H)$ such that $\mu(B) = 0$ for all $\mu \in G(H)$. Lebesgue measure in $\mathbb{R}^k$ will be denoted by $\mathscr{L}^k$.

Wasserstein distance, optimal couplings and maps. We have already defined the class of *couplings* between two probability measures μ and ν on H and the Wasserstein distance $W_2(\mu, \nu)$: see (1.9). Existence of a minimizing Σ in (1.9) is a simple consequence of the tightness of $\Gamma(\mu, \nu)$; the class of *optimal couplings* will be denoted by $\Gamma_o(\mu, \nu)$:

$$\Gamma_o(\mu, \nu) := \left\{ \Sigma \in \Gamma(\mu, \nu) : \int_{H \times H} \|y - x\|^2 \, d\Sigma = W_2^2(\mu, \nu) \right\}. \qquad (2.1)$$

In the special case when μ vanishes on all Gaussian null sets (that corresponds to absolute continuity with respect to Lebesgue measure in finite dimensions) it has been proved in Theorem 6.2.10 of [3] that there exists a unique optimal coupling Σ, and it is induced by an optimal *transport map* t, namely $\Sigma = (i \times t)_{\#}\mu$ (the proof is based on the fact that the non-Gateaux differentiability set of a Lipschitz function in H is Gaussian null, see e.g. Theorem 5.11.1 in [6]). We shall denote this optimal transport map by t_μ^ν. This is one of the infinite-dimensional generalizations (see also [17] for another result in Wiener spaces) of the finite-dimensional result ensuring that whenever $\mu \in \mathscr{P}_2(\mathbb{R}^k)$ is absolutely continuous with respect to $\mathscr{L}^k$, then there exists a unique optimal transport map that is also the gradient of a convex function.

When we have a sequence $(\gamma_n) \subset \mathscr{P}_2(H)$ as in Assumption 1.4, we can introduce Wasserstein distances in $\mathscr{P}_2(A_n)$ using two different scalar products: $\langle \cdot, \cdot \rangle_H$ and $\langle \cdot, \cdot \rangle_{H_n}$. The Wasserstein distance with respect to the former one is indicated in the standard way $W_2(\cdot, \cdot)$, while we introduce the notation:

$$W_{2,H_n}^2(\mu, \nu) := \inf \left\{ \int_{A_n \times A_n} \|y - x\|_{H_n}^2 \, d\Sigma : \Sigma \in \Gamma(\mu, \nu) \right\}, \quad \mu, \nu \in \mathscr{P}_2(A_n);$$

$$(2.2)$$

notice that if $x, y \in A_n$ then $x - y \in H_n$, so that $\|x - y\|_{H_n}$ makes sense. If μ, ν are supported in A_n we also denote the class of optimal couplings in $\Gamma(\mu, \nu)$ with respect to the H_n-distance by $\Gamma_{H_n, o}(\mu, \nu)$. By (1.10) the two distances are equivalent.

Convergence of measures. We will use two notions convergence of measures: first the *weak* convergence in $\mathscr{P}(H)$, induced by the duality with $C_b(H)$; second, the convergence in $\mathscr{P}_2(H)$ induced by the Wasserstein distance. The two definitions are related by the following result (see [3], Theorem 5.1.13 and Remark 7.1.11):

Lemma 2.1 *If $(\mu_n) \subset \mathscr{P}_2(H)$, then $\mu_n \to \mu$ in $\mathscr{P}_2(H)$ if and only if $\mu_n \to \mu$ weakly and*

$$\lim_{n\to\infty} \int_H \|x\|^2 \, d\mu_n = \int_H \|x\|^2 \, d\mu. \tag{2.3}$$

Notice that, for weakly converging sequences (μ_n), the convergence of the second moments (2.3) is easily seen to be equivalent to

$$\lim_{R\uparrow\infty} \limsup_{n\to\infty} \int_{\{\|x\|\geq R\}} \|x\|^2 \, d\mu_n = 0. \tag{2.4}$$

We recall that weak convergence of μ_n to μ implies

$$\liminf_{n\to\infty} \int_H f \, d\mu_n \geq \int_H f \, d\mu, \tag{2.5}$$

for every lower semicontinuous function $f : H \to (-\infty, +\infty]$ bounded from below. We shall also often use the following extension, involving integration with respect to a variable function: if f_n are uniformly bounded from below and equi-continuous, we have

$$\liminf_{n\to\infty} \int_H f_n \, d\mu_n \geq \int_H \liminf_{n\to\infty} f_n \, d\mu. \tag{2.6}$$

The proof immediately follows by (2.5), with the monotone approximation with the continuous functions $g_k = \inf_{n\geq k} f_n$.

Log-concave probability measures and Entropy. The concept of log-concavity has been introduced in Assumption 1.1. Since this concept is crucial in this paper, we recall the following result.

Proposition 2.2 ([3,9], Theorem 9.4.11) *Let $H = \mathbb{R}^k$. Then $\gamma \in \mathscr{P}(H)$ is log-concave if and only if it admits the following representation*:

$$\gamma(B) = \int_{B\cap\{V<+\infty\}} e^{-V} \, d\mathscr{H}^d \quad \forall B \in \mathscr{B}(\mathbb{R}^k), \tag{2.7}$$

where $V : \mathbb{R}^k \to (-\infty, +\infty]$ is a suitable convex and lower semicontinuous function, $d \geq 0$ is the dimension of $A(\gamma)$, and $\mathscr{H}^d$ is the d-dimensional Hausdorff measure.

If the dimension of H is infinite, then $\gamma \in \mathscr{P}(H)$ is log-concave if and only if all the finite dimensional projections of γ are log-concave and therefore admit the representation (2.7) for some V and d.

If γ is log-concave, the relative Entropy functional $\mathcal{H}(\cdot|\gamma)$ (1.8) enjoys a crucial convexity property in terms of Wasserstein distance, which has been discovered by McCann in [29] and further extended to the infinite dimensional case in [3].

Proposition 2.3 (Displacement convexity of the relative entropy) *Let $\gamma \in \mathscr{P}(H)$ be log-concave and let μ^0, $\mu^1 \in \mathscr{P}_2(H)$ with finite relative entropy. Then there exists an optimal coupling $\Sigma \in \Gamma_o(\mu^0, \mu^1)$ such that the curve in $\mathscr{P}_2(H)$*

$$\mu^t := \left((1-t)\pi^0 + t\pi^1\right)_{\#} \Sigma, \ \left(\pi^i : (x^0, x^1) \in H \times H \mapsto x^i \in H, \ i = 0, 1\right) \ \ (2.8)$$

satisfies

$$\mathcal{H}(\mu^t|\gamma) \le (1-t)\mathcal{H}(\mu^0|\gamma) + t\mathcal{H}(\mu^1|\gamma) \ \ \forall t \in [0, 1]. \tag{2.9}$$

When $H = \mathbb{R}^k$ is finite dimensional and μ^0 is absolutely continuous w.r.t. the Lebesgue measure then the optimal coupling $\Sigma = (\boldsymbol{i} \times \boldsymbol{t})_{\#}\mu^0$ is unique, so that

$$\mu^t = ((1-t)\boldsymbol{i} + t\,\boldsymbol{t})_{\#}\,\mu^0. \tag{2.10}$$

3 From Fokker–Planck equation to Wasserstein gradient flows

In this section we illustrate the known connections between solutions of the SDE (1.6), solutions to Fokker–Planck equations, Dirichlet semigroups and Wasserstein gradient flows in the model case when the drift term in the SDE is the bounded gradient ∇V of a smooth function $V : \mathbb{R}^k \to \mathbb{R}$ satisfying:

$$\|\nabla V(x) - \nabla V(y)\| \le L\|x - y\| \ \ \forall x, y \in \mathbb{R}^k$$

for some $L > 0$. We shall also assume that all derivatives of V are bounded and that $\gamma = \exp(-V)\,\mathscr{L}^k$ is a log-concave probability measure in $\mathbb{R}^k$. Notice that this implies that V is convex, and also (see Appendix A) that there exist constants $A \in \mathbb{R}$ and $B > 0$ such that $V(x) \ge A + B\|x\|$ for all $x \in \mathbb{R}^k$.

All theories mentioned above have a much larger realm of validity [for instance, much less regular drift terms in the SDE (1.6) are allowed], but for our purposes it suffices to show connections and a few a priori estimates in the smooth, bounded, Lipschitz case: more general cases will follows thanks to the stability Theorem 1.5 (or its Wasserstein counterpart Theorem 6.1).

Let us fix k independent standard Brownian motions $\{W^1, \ldots, W^k\}$ on a probability space. We consider the $\mathbb{R}^k$-valued Brownian motion $(W_t)_{t \ge 0}$, where $W = (W^1, \ldots, W^k)$. Since ∇V is bounded and Lipschitz continuous, it is well known that, for all $x \in \mathbb{R}^k$, there exists a unique solution $(X_t(x) : t \ge 0)$ of the SDE

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2}\,dW, \quad X_0(x) = x. \tag{3.1}$$

Notice that $(X_t(x) - X_t(y), t \ge 0)$ solves almost surely an ordinary differential equation, since the stochastic terms dW cancel out; then one easily obtains from the

convexity of V that $t \mapsto \|X_t(x) - X_t(y)\|^2$ is non-increasing in $[0, +\infty[$ almost surely. As a consequence, a.s.

$$\|X_t(x) - X_t(y)\|^2 \leq \|x - y\|^2, \quad \forall x, y \in \mathbb{R}^k, \ t \geq 0. \tag{3.2}$$

For all $x \in \mathbb{R}^k$, $t \geq 0$ and $\mu_0 \in \mathscr{P}(\mathbb{R}^k)$ we set:

$$\nu_t^x := \text{law of } X_t(x), \quad \mu_t := \int \nu_t^x \, d\mu_0(x) \in \mathscr{P}(\mathbb{R}^k). \tag{3.3}$$

By (3.2), the map $x \mapsto \nu_t^x$ is weakly continuous, and therefore μ_t is well defined. Moreover, the continuity of the process $(X_t(x))_{t \geq 0}$ yields weak continuity of $t \mapsto \nu_t^x$ and $t \mapsto \mu_t$.

It is a trivial consequence of Itô's formula that μ_t solves the Fokker–Planck equation in the sense of distributions in $]0, +\infty[\times \mathbb{R}^k$:

$$\frac{d}{dt} \mu_t = \Delta \mu_t + \nabla \cdot (\nabla V \mu_t); \tag{3.4}$$

this means that

$$\frac{d}{dt} \int_{\mathbb{R}^k} \varphi \, d\mu_t = \int_{\mathbb{R}^k} (\Delta \varphi - \langle \nabla V, \nabla \varphi \rangle) \, d\mu_t \quad \forall \varphi \in C_c^\infty(\mathbb{R}^k), \tag{3.5}$$

and the initial condition at $t = 0$ is attained in the following weak sense:

$$\lim_{t \downarrow 0} \int_{\mathbb{R}^k} \varphi \, d\mu_t = \int_{\mathbb{R}^k} \varphi \, d\mu_0, \quad \forall \varphi \in C_c^\infty(\mathbb{R}^k). \tag{3.6}$$

Equivalently, (3.5) and (3.6) can be grouped by saying that for every $T > 0$ and $\varphi \in C_c^\infty([0, T] \times \mathbb{R}^k)$ we have

$$\int_{\mathbb{R}^k} \varphi_T \, d\mu_T = \int_{\mathbb{R}^k} \varphi \, d\mu_0 + \int_0^T \int_{\mathbb{R}^k} (\partial_t \varphi_t + \Delta \varphi_t - \langle \nabla V, \nabla \varphi_t \rangle) \, d\mu_t \, dt. \tag{3.7}$$

We have now the following standard result on Fokker–Planck equations.

Proposition 3.1 (Uniqueness and stability of FP solutions) *For any $\mu_0 \in \mathscr{P}(\mathbb{R}^k)$, the Fokker–Planck equation (3.7) has a unique solution in the class of weakly continuous maps $t \mapsto \mu_t \in \mathscr{P}(\mathbb{R}^k)$. If $\mu_0 \in \mathscr{P}_2(\mathbb{R}^k)$, then the unique solution $[0, +\infty[\ni t \mapsto \mu_t \in \mathscr{P}_2(\mathbb{R}^k)$ is continuous. In addition, μ_t is stable: $\mu_0^n \to \mu_0$ in $\mathscr{P}_2(\mathbb{R}^k)$ implies $\mu_t^n \to \mu_t$ in $\mathscr{P}_2(\mathbb{R}^k)$ for all $t \geq 0$.*

Proof We consider first the case $\mu_0 = \delta_x$: it only remains to prove that $[0, \infty) \ni t \mapsto v_t^x \in \mathscr{P}_2(H)$ is continuous. Taking (2.4) into account, it is enough to prove that

$$\sup_{0 \leq t \leq T} \mathbb{E}\big[\|X_t(x)\|^2\big] \leq C(1 + \|x\|^2) < +\infty, \tag{3.8}$$

$$\sup_{0 \leq t \leq T} \mathbb{E}\big[\|X_t(x)\|^2 1_{\{\|X_t(x)\| \geq R\}}\big] \leq \omega(R)(1 + \|x\|^2) \quad \text{with} \quad \omega(R) \downarrow 0 \text{ as } R \uparrow +\infty, \tag{3.9}$$

where C and ω depend only on T and $\sup \|\nabla V\|$. We apply Itô's formula to $\|X_t(x)\|$ and find that

$$\|X_t\| \leq \|x\| + T \sup \|\nabla V\| + (k - 1) \int_0^t \frac{1}{\|X_s(x)\|} \, ds + \sqrt{2}\, \hat{B}_t, \quad \forall t \in [0, T],$$

where $\hat{B}$ is a standard Brownian motion in $\mathbb{R}$. We define now the process (b_t), unique non-negative solution of

$$b_t = \|x\| + T \sup \|\nabla V\| + (k - 1) \int_0^t \frac{1}{b_s} \, ds + \sqrt{2}\, \hat{B}_t, \quad \forall t \geq 0.$$

Then (b_t) is a Bessel process of dimension k, i.e. $(b_t, t \geq 0)$ is equal in law to $(\|b_0 \cdot h + \sqrt{2}\, W_t\|, t \geq 0)$, where $h \in \mathbb{R}^k$ with $\|h\| = 1$: see [30]. By Itô's formula, $\|X_t\| \leq b_t$ for all $t \in [0, T]$, almost surely. Then (3.8) and (3.9) follow from standard Gaussian estimates.

Existence of FP solutions, as we have seen, is provided by (3.3). Uniqueness can for instance be obtained by a classical duality argument: let μ_t^1, μ_t^2 be two weakly continuous solutions of (3.5), and let $\sigma := \mu_t^1 - \mu_t^2$ be their difference, satisfying

$$\int_{\mathbb{R}^k} \varphi_T \, d\sigma_T = \int_0^T \int_{\mathbb{R}^k} (\partial_t \varphi_t + \Delta \varphi_t - \langle \nabla V, \nabla \varphi_t \rangle) \, d\sigma_t \, dt \tag{3.10}$$

for every $T > 0$ and $\varphi \in C_c^\infty([0, T] \times \mathbb{R}^k)$. By a mollification technique, it is not difficult to check that (3.10) holds even for every function $\varphi \in C([0, T] \times \mathbb{R}^k)$ with $\partial_t \varphi$, $\nabla \varphi$ and $\nabla^2 \varphi$ continuous and bounded in $[0, T] \times \mathbb{R}^k$. For given $\psi \in C_c^\infty(\mathbb{R}^k)$ we consider the solution φ_t of the time reversed (adjoint) parabolic equation

$$\partial_t \varphi_t + \Delta \varphi_t - \langle \nabla V, \nabla \varphi_t \rangle = 0 \quad \text{in } (0, T) \times \mathbb{R}^k, \qquad \varphi_T = \psi. \tag{3.11}$$

Standard parabolic regularity theory (it suffices to use the maximum principle [23] and the fact that the first and second order spatial derivatives of φ solve an analogous

equation) shows that φ is sufficiently regular to be used as a test function in (3.10): this leads to $\int \psi \, d\sigma_T = 0$. As ψ is arbitrary, we obtain that $\mu_T^1 = \mu_T^2$.

The representation $\mu_t = \int \nu_t^x \, d\mu_0(x)$, given in (3.3), and the uniform estimates (3.8), (3.9) easily imply the stability property. $\qquad\square$

Notice that the measure γ provides a *stationary* solution of (3.4) (and it can be actually shown that all solutions μ_t weakly converge to γ as $t \to +\infty$); it is also natural to consider initial conditions $\mu_0 = \rho_0 \gamma \in \mathscr{P}(H)$ with $\rho_0 \in L^2(\gamma)$. In this class of initial data, one can consider the variational formulation of the FP equation induced by the symmetric Dirichlet form

$$\mathcal{E}_\gamma(\rho, \eta) := \int_{\mathbb{R}^k} \langle \nabla\rho, \nabla\eta \rangle \, d\gamma, \qquad \rho, \eta \in W_\gamma^{1,2}(\mathbb{R}^k), \tag{3.12}$$

where $W_\gamma^{1,2}(\mathbb{R}^k)$ is the weighted Sobolev space

$$W_\gamma^{1,2}(\mathbb{R}^k) := \left\{ \rho \in L^2(\gamma) \cap W_{\mathrm{loc}}^{1,1}(\mathbb{R}^k) : \int_{\mathbb{R}^k} \|\nabla\rho\|^2 \, d\gamma < +\infty \right\}. \tag{3.13}$$

Proposition 3.2 *For every $\rho_0 \in L^2(\gamma)$:*

1. *there exists a unique $\rho. \in W_{\mathrm{loc}}^{1,2}\left(]0, +\infty[; L^2(\gamma)\right)$ such that*

$$\frac{d}{dt}\langle \rho_t, \eta \rangle_{L^2(\gamma)} + \mathcal{E}_\gamma(\rho_t, \eta) = 0, \quad \forall \eta \in W_\gamma^{1,2}(\mathbb{R}^k), \ \lim_{t \downarrow 0} \rho_t = \rho_0 \ \text{in } L^2(\gamma); \tag{3.14}$$

   if $\rho_{\min} \le \rho_0 \le \rho_{\max}$, then ρ_t satisfies the same uniform bounds;
2. *if $\rho_0 \ge 0$ and $\int \rho_0 \, d\gamma = 1$, then $\mu_t = \rho_t \gamma \in \mathscr{P}(H)$ provides the unique solution of (3.5) starting from $\mu_0 = \rho_0 \gamma$;[-7pt]*
3. *if $\rho_0 \in C_b(\mathbb{R}^k)$, then, for all $t \ge 0$, the function*

$$P_t \rho_0(x) := \mathbb{E}(\rho_0(X_t(x))) = \int \rho_0 \, d\nu_t^x, \quad \forall x \in \mathbb{R}^k, \tag{3.15}$$

   provides a continuous version of ρ_t, i.e. $P_t \rho_0(x) = \rho_t(x)$ for γ-a.e. x; moreover, P_t acts on $\mathrm{Lip}_b(\mathbb{R}^k)$ and

$$[P_t \rho]_{\mathrm{Lip}(\mathbb{R}^k)} \le [\rho]_{\mathrm{Lip}(\mathbb{R}^k)} \quad t \ge 0, \quad \rho \in \mathrm{Lip}_b(\mathbb{R}^k); \tag{3.16}$$

4. *$(P_t)_{t \ge 0}$ has an extension to a symmetric strongly continuous semigroup in $L^2(\gamma)$.*

Proof Existence of a unique solution of (3.14) follows by the well-known theory of variational evolution equations, as well as the uniform lower and upper bounds on ρ_t, being $\mathcal{E}_\gamma$ a Dirichlet form: this proves point 1. Now, for $\varphi \in C_c^\infty(\mathbb{R}^k)$ we can choose

$\eta = \varphi \exp(V)$ in (3.14) and integrate by parts in space to obtain (3.5): this shows that μ_t is the unique solution of the FP equation, as stated in point 2. Continuity of $P_t \rho_0$ follows from (3.2); in order to prove that $\rho_t = P_t \rho_0$ γ-a.e., we can reduce by linearity to the case $\rho_0 \gamma \in \mathscr{P}(\mathbb{R}^k)$ and then point 3 follows from point 2. Point 4 follows from standard $L^2(\gamma)$ estimates for Eq. (3.14).   □

Using the convexity inequality $\mathcal{E}_\gamma(\eta, \eta) \geq \mathcal{E}_\gamma(\rho, \rho) + 2\mathcal{E}_\gamma(\rho, \eta - \rho)$, it is not difficult to show that an equivalent formulation of (3.14) is (this kind of formulation first appeared in [5], in connection with nonlinear evolution problems in Banach spaces)

$$\frac{d}{dt} \frac{1}{2} \|\rho_t - \eta\|^2_{L^2(\gamma)} + \frac{1}{2} \mathcal{E}_\gamma(\rho_t, \rho_t) \leq \frac{1}{2} \mathcal{E}_\gamma(\eta, \eta), \quad \forall \eta \in W^{1,2}_\gamma(\mathbb{R}^k). \tag{3.17}$$

We are going to show an analogous property of the solutions μ_t of the Fokker–Planck equation (3.7), obtained just replacing L^2 norm with Wasserstein distance and $\frac{1}{2}\mathcal{E}_\gamma(\rho, \rho)$ with the relative Entropy functional $\mathcal{H}(\rho\gamma|\gamma)$ with respect to γ. This provides the key connection with the variational theory of gradient flows in Wasserstein spaces. To this aim, let us first establish the analogue of the convexity inequality for the relative Entropy:

Lemma 3.3 (Energy inequality) *Let* $\rho \in C^1(\mathbb{R}^k) \cap W^{1,2}_\gamma(\mathbb{R}^k)$, *with* $0 < \rho_{\min} \leq \rho \leq \rho_{\max} < +\infty$. *Then*:

$$\mathcal{H}(\eta\gamma|\gamma) \geq \mathcal{H}(\rho\gamma|\gamma) + \int_{\mathbb{R}^k} \langle \nabla\rho, \boldsymbol{t} - \boldsymbol{i} \rangle \, d\gamma, \quad \forall \eta\gamma \in \mathscr{P}_2(\mathbb{R}^k), \tag{3.18}$$

where $\boldsymbol{t}$ *is the optimal transport map between* $\rho\gamma$ *and* $\eta\gamma$.

Proof We just sketch the proof, referring to [3, Lemma 10.4.4, Lemma 10.4.5] or to [11] for more general results and detailed arguments. Let $\boldsymbol{t} = \nabla\phi$, with ϕ convex, and $u = \rho \exp(-V)$. Defining $\mu^0 = \rho\gamma$, $\mu^1 = \eta\gamma$ and μ^t as in (2.8), taking (2.9) into account it suffices to bound from below $\frac{d}{dt^+}\mathcal{H}(\mu^t|\gamma)$ at $t = 0$ with $\int \langle \nabla\rho, \boldsymbol{t} - \boldsymbol{i} \rangle \, d\gamma$. Now, a computation based on the area formula (that provides an explicit expression for the density of μ^t with respect to $\mathscr{L}^k$, see also the proof of Lemma A.1 in the Appendix) gives

$$\frac{d}{dt^+}\mathcal{H}(\mu^t|\gamma)\bigg|_{t=0} = -\int_{\mathbb{R}^k} \text{trace} \, (\nabla^2\phi - I)u(x) \, dx + \int_{\mathbb{R}^k} \langle \nabla V, \boldsymbol{t} - \boldsymbol{i} \rangle \, \rho \, d\gamma,$$

where $\nabla^2\phi$ is the Alexandrov pointwise second derivative of ϕ and I is the identity matrix. By the convexity of ϕ we can bound the matrix of absolutely continuous measures $\nabla^2\phi \, \mathscr{L}^k$ from above with the distributional derivative of $\nabla\phi = \boldsymbol{t}$ (which, in general, is a measure, by the convexity of ϕ) to obtain

$$\frac{d}{dt^+}\mathcal{H}(\mu^t|\gamma)\bigg|_{t=0} \geq -\langle \nabla \cdot (\boldsymbol{t} - \boldsymbol{i}), u \rangle + \int_{\mathbb{R}^k} \langle \nabla V, \boldsymbol{t} - \boldsymbol{i} \rangle \, \rho \, d\gamma.$$

Finally, we integrate by parts to obtain (3.18): although u is not compactly supported, this formal integration can be justified by approximation of u with $u\psi_R$, where $\psi_R \in C_c^\infty(\mathbb{R}^k)$, $0 \le \psi_R \le 1$, $\|\nabla\psi_R\| \le 1$, $\psi_R \uparrow 1$ and $\nabla\psi_R \to 0$ as $R \to +\infty$. $\qquad\Box$

Proposition 3.4 *For all $\mu_0 \in \mathscr{P}_2(\mathbb{R}^k)$ the solution (μ_t) of the FP equation (3.4), characterized in Proposition 3.1, satisfies the family of variational evolution inequalities*

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mu_t, \nu) + \mathcal{H}(\mu_t|\gamma) \le \mathcal{H}(\nu|\gamma) \qquad (3.19)$$

in the sense of distributions in $]0, +\infty[$, for all $\nu \in \mathscr{P}_2(\mathbb{R}^k)$.

Proof First of all, we notice that the variational evolution inequalities (3.19) are stable with respect to pointwise convergence in $\mathscr{P}_2(\mathbb{R}^k)$: indeed, if $\mu_t^n \to \mu_t$ for all t, then $\frac{d}{dt}W_2^2(\mu_t^n, \nu) \to \frac{d}{dt}W_2^2(\mu_t, \nu)$ in the sense of distributions, and the lower semicontinuity of $\mathcal{H}(\cdot|\gamma)$ allows to pass to the limit as $n \to \infty$.

Thanks to this remark and to the stability properties of solutions to FP equations, we need only to show the property when $\mu_0 = \rho_0\gamma$ with $\rho_0 \in C_b^2(\mathbb{R}^k)$, $\rho_0 \ge \rho_{\min} > 0$. Then, we know from Proposition 3.2 that $\mu_t = \rho_t\gamma$, with ρ_t smooth, $\rho_t \ge \rho_{\min}$. In addition, writing $u_t = \rho_t\exp(-V)$, since $u_0 = \rho_0\exp(-V)$ belongs to $C_b^2(\mathbb{R}^k)$ as well, standard parabolic regularity theory for the FP equation whose drift is bounded, together will all its derivatives, gives $\partial_t u$, ∇u, $\nabla^2 u \in C_b([0, T] \times \mathbb{R}^k)$ for all $T > 0$.

We are interested in getting pointwise bounds for the velocity field $\boldsymbol{v}_t := -(\nabla\rho_t)/\rho_t$; it appears in a natural way in this problem because, by (3.4), μ_t solve the classical *continuity equation*

$$\frac{d}{dt}\mu_t + \nabla \cdot (\boldsymbol{v}_t\mu_t) = 0 \quad \text{in }]0, +\infty[\times\mathbb{R}^k,$$

describing the evolution of a time-dependent mass distribution μ_t under the action of a velocity field $\boldsymbol{v}_t$. Since $\rho_t = u_t\exp(V)$, and $\exp(V)$ is not bounded above, we can not use directly the C_b^2 bounds on u_t. However, we can use the fact that ρ_t solve the linear PDE $\partial_t\rho_t = \Delta\rho_t - \langle\nabla\rho_t, \nabla V\rangle$, to obtain

$$\partial_t\nabla\rho = \Delta\nabla\rho - \nabla^2 V\nabla\rho - \nabla^2\rho\nabla V.$$

Taking the scalar product with $\nabla\rho$ and using the identity $\Delta\|\nabla\rho\|^2 = 2\langle\Delta\nabla\rho, \nabla\rho\rangle + 2\|\nabla^2\rho\|^2$, we can neglect the nonnegative terms $\|\nabla^2\rho\|^2$ and $\langle\nabla^2 V\nabla\rho, \nabla\rho\rangle$ to get

$$\partial_t\|\nabla\rho_t\|^2 \le \Delta\|\nabla\rho_t\|^2 - \langle\nabla V, \nabla\|\nabla\rho_t\|^2\rangle.$$

The classical maximum principle can now be applied, thanks to the fact that $\|\nabla\rho_t\|^2$ grows at most exponentially [23, p. 216] to obtain $\|\nabla\rho_t\|^2 \le \|\nabla\rho_0\|^2$ for all $t \ge 0$. By the uniform lower bound on ρ_t, $\|\boldsymbol{v}_t\|$ is uniformly bounded.

Now, let us show first that $t \mapsto \mu_t$ is a Lipschitz map in $[0, +\infty[$ with values in $\mathscr{P}_2(\mathbb{R}^k)$. Let $0 \le t_1 \le t_2 < +\infty$; the smoothness and the just proven boundedness of $\boldsymbol{v}_t$ ensure

the existence of a unique flow $Y(t, x) : [0, +\infty[\times \mathbb{R}^k \to \mathbb{R}^k$ associated to $\boldsymbol{v}_t$, i.e. $Y(0, x) = x$ and $\frac{d}{dt} Y(t, x) = \boldsymbol{v}_t(Y(t, x))$ in $[0, +\infty[$. Then, the method of characteristics (see e.g. [3, Proposition 8.1.8]) for solutions to the continuity equation shows that μ_t is given by $Y(t, \cdot)_{\#}\mu_0$ for all $t \in [0, +\infty[$. Therefore we can use the coupling $(Y(t_2, \cdot), Y(t_1, \cdot))_{\#}\mu_0$ to estimate $W_2(\mu_{t_1}, \mu_{t_2})$ as follows:

$$
W_2^2(\mu_{t_1}, \mu_{t_2}) \leq \int_{\mathbb{R}^k} \|Y(t_2, x) - Y(t_1, x)\|^2 \, d\mu_0
$$

$$
= \int_{\mathbb{R}^k} \| \int_{t_1}^{t_2} \boldsymbol{v}_t(Y(t, x)) \, dt \|^2 \, d\mu_0 \leq C(t_2 - t_1)^2.
$$

This proves that $t \mapsto \mu_t$ is Lipschitz in $[0, +\infty[$.

To conclude the proof, it suffices to check (3.19) at any differentiability point $\bar{t} \in [0, +\infty[$ of the map $t \mapsto W_2^2(\mu_t, \nu)$. Let $\boldsymbol{t}$ be the optimal transport map between $\mu_{\bar{t}}$ and ν, and let $h > 0$; since $\Sigma = (Y(\bar{t} + h, \cdot), \boldsymbol{t} \circ Y(\bar{t}, \cdot))_{\#}\mu_0$ is a coupling between $\mu_{\bar{t}+h}$ and ν, we can estimate (using the identity $\|a\|^2 - \|b\|^2 = \langle a + b, a - b \rangle$):

$$
\frac{W_2^2(\mu_{\bar{t}+h}, \nu) - W_2^2(\mu_{\bar{t}}, \nu)}{h}
$$

$$
\leq \frac{1}{h} \int_{\mathbb{R}^k} \|\boldsymbol{t}(Y(\bar{t}, x)) - Y(\bar{t} + h, x)\|^2 \, d\mu_0 - \frac{1}{h} \int_{\mathbb{R}^k} \|\boldsymbol{t}(y) - y\|^2 \, d\mu_{\bar{t}}
$$

$$
= \frac{1}{h} \int_{\mathbb{R}^k} \left[\|\boldsymbol{t}(Y(\bar{t}, x)) - Y(\bar{t} + h, x)\|^2 - \|\boldsymbol{t}(Y(\bar{t}, x)) - Y(\bar{t}, x)\|^2 \right] d\mu_0
$$

$$
= -\int_{\mathbb{R}^k} \langle 2\boldsymbol{t}(Y(\bar{t}, x)) - Y(\bar{t} + h, x) - Y(\bar{t}, x), \frac{1}{h} \left(Y(\bar{t} + h, x) - Y(\bar{t}, x) \right) \rangle \, d\mu_0
$$

$$
\longrightarrow -2 \int_{\mathbb{R}^k} \langle \boldsymbol{t}(Y(\bar{t}, x)) - Y(\bar{t}, x), \boldsymbol{v}_{\bar{t}}(Y(\bar{t}, x)) \rangle \, d\mu_0 = -2 \int_{\mathbb{R}^k} \langle \boldsymbol{t}(y) - y, \boldsymbol{v}_{\bar{t}}(y) \rangle \, d\mu_{\bar{t}},
$$

as $h \downarrow 0$, by dominated convergence. From the energy inequality (3.18) we obtain (3.19). □

Starting from (3.4), we have derived a new relation (3.19) satisfied by solutions to FP equations, at least when $H = \mathbb{R}^k$ and V is smooth and convex, with ∇V bounded and Lipschitz. The idea of [3] is to consider (3.19) as the *definition* of a differential equation in a space of probability measures endowed with the Wasserstein distance even when either V is less regular or H is infinite-dimensional: more precisely, $(\mu_t)_{t \geq 0}$ is said to solve the *gradient flow* of the functional $\mathcal{H}(\cdot | \gamma)$; the choice of δ_x as initial measure is the one that provides the link with the laws ν_t^x of the solution $X_t(x)$ of the SDE (3.1).

Notice that (3.19) is defined only in terms of the Wasserstein distance and the relative entropy, namely of objects which make perfectly sense on an arbitrary Hilbert space H. Motivated by Proposition 3.4, we set the following:

Definition 3.5 (Gradient flows)
Let $F : \mathscr{P}_2(H) \mapsto [0, +\infty]$ and set $D(F) = \{F < +\infty\}$. We say that a continuous map $\mu_t :]0, +\infty[\mapsto \mathscr{P}_2(H)$ is a gradient flow of F if

$$\frac{d}{dt} \frac{1}{2} W_2^2(\mu_t, \nu) + F(\mu_t) \leq F(\nu) \tag{3.20}$$

in the sense of distributions in $]0, +\infty[$, for all $\nu \in D(F)$.
 We say that μ_t starts from μ_0 if $\mu_t \to \mu_0$ in $\mathscr{P}_2(H)$ as $t \downarrow 0$.

The terminology "gradient flow" can be justified, by appealing to Otto's formal differential calculus on $\mathscr{P}_2(\mathbb{R}^k)$; since this calculus will not play a significant role in our paper we will not discuss this issue, and refer to [3,24,36] for much more heuristics on this subject. Here we just point out that existence of gradient flows can be obtained (see Sect. 5) by the analogue in the Wasserstein setting of the Euler scheme for the approximation of gradient flows $\dot{x}(t) = -\nabla F(x(t))$: namely, given a time step $\tau > 0$, we build a sequence (x_k) by minimizing

$$y \mapsto \frac{1}{2\tau} \|y - x_k\|^2 + F(y)$$

recursively (i.e., given x_k, we choose x_{k+1} among the minimizers of the variational problem above). Looking at the discrete Euler equation, $(x_{k+1} - x_k)/\tau = -\nabla F(x_{k+1})$, it is clear that $x_k \sim x(k\tau)$.
 Notice also that (3.20) implies that $\mu_t \equiv \mu_0 \in \mathscr{P}_2(H)$ is a constant gradient flow if and only if μ_0 minimizes F; in the case $F(\mu) = \mathcal{H}(\mu|\gamma)$, since $t \mapsto t \ln t$ is strictly convex, the unique minimizer of F is $\mu = \gamma$. So, from the gradient flow viewpoint, we easily see that the unique invariant measure in $\mathscr{P}_2(H)$ is γ if $\gamma \in \mathscr{P}_2(H)$: see Theorem 1.6 and its proof in Sect. 8.
 In the next sections we are going to adopt the "gradient flow" point of view, and prove that the results of Theorems 1.2, 1.3 and 1.5 are relatively easy consequences of this approach.

4 Implicit Euler scheme

In this section we construct a discrete approximation of the gradient flow. Such construction is based on the following convexity property, which is a stronger version of the one given in Proposition 2.3.

Definition 4.1 We say that $F : \mathscr{P}_2(H) \to [0, +\infty]$ is strongly displacement convex if for any $\bar{\mu}, \nu_0, \nu_1 \in \mathscr{P}_2(H)$ there exists a continuous curve $\nu : [0, 1] \mapsto \mathscr{P}_2(H)$

such that $\nu_{|t=i} = \nu_i$, $i = 0, 1$, such that:

$$\begin{cases} W_2^2(\nu_t, \bar{\mu}) \leq (1-t)W_2^2(\nu_0, \bar{\mu}) + t\,W_2^2(\nu_1, \bar{\mu}) - t(1-t)W_2^2(\nu_0, \nu_1) \\ F(\nu_t) \leq (1-t)F(\nu_0) + t\,F(\nu_1) \end{cases} \quad \forall t \in [0, 1]. \tag{4.1}$$

In this paper we consider only the case $F(\mu) := \mathcal{H}(\mu|\gamma)$, where γ is a log-concave probability measure on H. In fact the following results are true for much more general classes of strongly displacement convex functionals $F : \mathscr{P}_2(H) \to (-\infty, +\infty]$ with essentially the same proof (under suitable lower semicontinuity and coercivity assumptions). However, at one point we shall take advantage of a particular feature of the relative entropy functional $\mathcal{H}(\cdot|\gamma)$, namely the entropy inequality (4.4), in order to simplify the proof. For more general cases, see Chap. 4 in [3].

In order to build gradient flows, we use an implicit Euler scheme, at least when $\bar{\mu} \in D(F)$; then, suitable Cauchy-type estimates provide existence up to initial data in $\overline{D(F)}$, as in the Hilbertian theory. The scheme can be described as follows: given a time step $\tau > 0$, we define a "discrete" solution μ_τ^k setting $\mu_\tau^0 = \bar{\mu}$ and, given μ_τ^k, we choose μ_τ^{k+1} as the unique minimizer of

$$\nu \mapsto \mathcal{H}(\nu|\gamma) + \frac{1}{2\tau}W_2^2(\nu, \mu_\tau^k). \tag{4.2}$$

The fact that this is possible is a consequence of the following

Proposition 4.2 *Let* $\tau > 0$. *For all* $\mu \in \mathscr{P}_2(H)$ *there exists a unique* $\mu_\tau \in \mathscr{P}_2(H)$ *such that*:

$$\mathcal{H}(\mu_\tau|\gamma) + \frac{1}{2\tau}W_2^2(\mu_\tau, \mu) \leq \mathcal{H}(\nu|\gamma) + \frac{1}{2\tau}W_2^2(\nu, \mu) \quad \forall \nu \in \mathscr{P}_2(H). \tag{4.3}$$

The existence part of this result is rather standard and relies on tightness and lower semicontinuity arguments. The uniqueness statement is based on the strong displacement convexity of the relative entropy functional, proved in Proposition 4.3.

Proof For notational convenience we denote $F(\mu) := \mathcal{H}(\mu|\gamma)$, for all $\mu \in \mathscr{P}_2(H)$.
Existence. Let $\nu_k = f_k\gamma \in \mathscr{P}_2(H)$ be such that:

$$\lim_{k\to\infty} \left[\mathcal{H}(\nu_k|\gamma) + \frac{1}{2\tau}W_2^2(\nu_k, \mu) \right] = \inf_{\nu \in \mathscr{P}_2(H)} \left\{ \mathcal{H}(\nu|\gamma) + \frac{1}{2\tau}W_2^2(\nu, \mu) \right\} < +\infty.$$

In particular we have that (ν_k) is bounded in $\mathscr{P}_2(H)$ and $\limsup_k \mathcal{H}(\nu_k|\gamma) < \infty$. By using first the inequality $t \ln t \geq -e^{-1}$ and then Jensen inequality we get

$$\frac{1}{e}\gamma(H \setminus E) + \mathcal{H}(\nu_k|\gamma) \geq \int_E f_k \ln f_k \, d\gamma \geq \nu_k(E) \ln \frac{\nu_k(E)}{\gamma(E)} \quad \forall E \in \mathscr{B}(H), \tag{4.4}$$

so that $\gamma(E) \to 0$ implies $\sup_k \nu_k(E) \to 0$. It follows that (ν_k) is tight in H, so that we can extract a subsequence, that we can still denote by (ν_k), converging weakly to some $\mu_\tau \in \mathscr{P}_2(H)$.

If we prove that both $\mathcal{H}(\cdot|\gamma)$ and $W_2^2(\cdot, \mu)$ are lower semicontinuous with respect to weak convergence, then we have that μ_τ realizes the minimum in (4.3): indeed by lower semicontinuity:

$$\mathcal{H}(\mu_\tau|\gamma) + \frac{1}{2\tau} W_2^2(\mu_\tau, \mu) \leq \liminf_{k \to \infty} \left[\mathcal{H}(\nu_k|\gamma) + \frac{1}{2\tau} W_2^2(\nu_k, \mu) \right]$$
$$= \inf_{\nu \in \mathscr{P}_2(H)} \left\{ \mathcal{H}(\nu|\gamma) + \frac{1}{2\tau} W_2^2(\nu, \mu) \right\}.$$

A nice representation of the relative entropy functional is provided by the *duality formula* (see for instance Lemma 9.4.4 of [3]):

$$\mathcal{H}(\mu|\gamma) = \sup \left\{ \int_H S \, d\mu - \int_H e^{S-1} \, d\gamma : S \in C_b(H) \right\}. \tag{4.5}$$

This formula immediately implies that $\mathcal{H}(\cdot|\gamma)$ is sequentially lower semicontinuous with respect to the weak convergence.

Let now $\Sigma_k \in \Gamma_o(\nu_k, \mu)$ and assume with no loss of generality that $W_2(\nu_k, \mu)$ converges to some limit; since (ν_k) is tight in H, (Σ_k) is tight in $H \times H$ and we can assume that (Σ_k) converges weakly to $\Sigma \in \mathscr{P}(H \times H)$. Obviously $\Sigma \in \Gamma(\nu, \mu)$ and the weak convergence of Σ_n gives

$$\int_{H \times H} \|x - y\|^2 \, d\Sigma \leq \liminf_{k \to \infty} \int_{H \times H} \|x - y\|^2 \, d\Sigma_k. \tag{4.6}$$

Bounding $W_2^2(\mu, \nu)$ from above with $\int \|x - y\|^2 \, d\Sigma$ we obtain the lower semicontinuity property. A similar argument also proves the *joint* lower semicontinuity of $(\mu, \nu) \mapsto W_2(\mu, \nu)$ (we will use this fact at the end of the proof of Proposition 4.3).

Uniqueness. Suppose that $\tilde{\mu}_\tau \neq \mu_\tau$ realize the minimum in (4.3), denoted by m. Let ν_t be a curve between $\tilde{\mu}_\tau$ and μ_τ given by Proposition 4.3 below, with the choice $\bar{\mu} = \mu$. Then we obtain:

$$\mathcal{H}(\nu_{1/2}|\gamma) + \frac{1}{2\tau} W_2^2(\nu_{1/2}, \mu) \leq m - \frac{1}{4} W_2^2(\tilde{\mu}_\tau, \mu_\tau) < m,$$

which is a contradiction. $\qquad \square$

Proposition 4.3 *Let $\gamma \in \mathscr{P}(H)$ be log-concave. Then the functional $\mathcal{H}(\cdot|\gamma):$ $\mathscr{P}_2(H) \to [0, +\infty]$ is strongly displacement convex.*

Proof As shown in [3], it is often enough to build the interpolating curves only for a dense subset $\mathscr{D}$ of measures $\bar{\mu}$. In the case of the relative entropy functional (but also

for more general classes of functionals, see [3]) the set $\mathscr{D}$ is made by finite convex combinations of non-degenerate Gaussian measures; as $\overline{\mathscr{D}}$ is easily seen to contain finite convex combinations of Dirac masses, $\mathscr{D}$ is dense in $\mathscr{P}_2(H)$. Moreover, any measure in $\mathscr{D}$ vanishes on the class $\mathscr{B}_G(H)$ of Gaussian null sets; hence, for any $\bar{\mu} \in \mathscr{D}$ and $\nu_0, \nu_1 \in \mathscr{P}_2(H)$ we can find optimal transport maps $\boldsymbol{r}_i$ between $\bar{\mu}$ and $\nu_i, i = 0, 1$. We define

$$\nu_t := ((1-t)\boldsymbol{r}_0 + t\boldsymbol{r}_1)_\# \mu. \tag{4.7}$$

Let us first check the Lipschitz continuity of $t \mapsto \nu_t \in \mathscr{P}_2(H)$: for $s, t \in [0, 1]$, the coupling

$$\Sigma_{st} := ((1-s)\boldsymbol{r}_0 + s\boldsymbol{r}_1, (1-t)\boldsymbol{r}_0 + t\boldsymbol{r}_1)_\# \bar{\mu}$$

belongs to $\Gamma(\nu_s, \nu_t)$, so that

$$W_2^2(\nu_s, \nu_t) \leq \int_{H \times H} \|x - y\|^2 \, d\Sigma_{st} = |t - s|^2 \int_H \|\boldsymbol{r}_1 - \boldsymbol{r}_0\|^2 \, d\bar{\mu}$$
$$\leq 2|t - s|^2 \left(W_2^2(\nu_1, \bar{\mu}) + W_2^2(\nu_0, \bar{\mu}) \right).$$

Now, let us check the convexity of $t \mapsto W_2^2(\nu_t, \bar{\mu})$:

$$W_2^2(\nu_t, \bar{\mu}) \leq \int_H \|(1-t)(\boldsymbol{r}_0 - \boldsymbol{i}) + t(\boldsymbol{r}_1 - \boldsymbol{i})\|^2 \, d\bar{\mu}$$
$$= (1-t) \int_H \|\boldsymbol{r}_0 - \boldsymbol{i}\|^2 \, d\bar{\mu} + t \int_H \|\boldsymbol{r}_1 - \boldsymbol{i}\|^2 \, d\bar{\mu}$$
$$- t(1-t) \int_H \|\boldsymbol{r}_0 - \boldsymbol{r}_1\|^2 \, d\bar{\mu}$$
$$\leq (1-t) W_2^2(\nu_0, \bar{\mu}) + t W_2^2(\nu_1, \bar{\mu}) - t(1-t) W_2^2(\nu_0, \nu_1).$$

In the last inequality we used the fact that $(\boldsymbol{r}_0, \boldsymbol{r}_1)_\# \bar{\mu} \in \Gamma(\nu_0, \nu_1)$. For the convexity of $t \mapsto \mathcal{H}(\nu_t | \gamma)$, achieved through a finite-dimensional approximation, we refer to [3, Theorem 9.4.11].

Having built the interpolating curves when $\bar{\mu} \in \mathscr{D}$, in the general case, we can approximate any $\bar{\mu} \in \mathscr{P}_2(H)$ by measures $\bar{\mu}^n \in \mathscr{D}$; notice that the interpolating curves $t \mapsto \nu_t^n$ between ν_0 and ν_1 are equi-Lipschitz and, for t fixed, the same tightness argument used in the existence part of Proposition 4.2 shows that (ν_t^n) is tight. Therefore, thanks to a diagonal argument, we can assume that $\nu_t^n \to \nu_t$ weakly for all $t \in [0, 1] \cap \mathbb{Q}$, with $t \mapsto \nu_t$ Lipschitz in $[0, 1] \cap \mathbb{Q}$. Passing to the limit as $n \to \infty$ in the convexity inequalities relative to ν_t^n, and using the weak lower semicontinuity

of $\mathcal{H}(\cdot|\gamma)$ and $W_2^2(\cdot, \cdot)$ we get

$$\begin{cases} W_2^2(\nu_t, \bar{\mu}) \leq (1-t)W_2^2(\nu_0, \bar{\mu}) + tW_2^2(\nu_1, \bar{\mu}) - t(1-t)W_2^2(\nu_0, \nu_1) \\ \mathcal{H}(\nu_t|\gamma) \leq (1-t)\mathcal{H}(\nu_0|\gamma) + t\mathcal{H}(\nu_1|\gamma) \end{cases} \tag{4.8}$$

for all $t \in [0, 1] \cap \mathbb{Q}$. By a density argument, based on the completeness of $\mathscr{P}_2(H)$, we can obtain a Lipschitz curve ν_t defined in the whole of $[0, 1]$, still retaining the inequalities above. $\square$

We prove now an important estimate which plays a key role in the sequel, see the proof of Theorem 5.1 below.

Proposition 4.4 *Let* $F : \mathscr{P}_2(H) \to [0, +\infty]$ *be strongly displacement convex, let* $\bar{\mu} \in \mathscr{P}_2(H)$ *and let* μ_τ *be a minimizer of*

$$\mu \mapsto F(\mu) + \frac{1}{2\tau}W_2^2(\mu, \bar{\mu}).$$

Then

$$W_2^2(\mu_\tau, \nu) - W_2^2(\bar{\mu}, \nu) \leq 2\tau\,[F(\nu) - F(\mu_\tau)], \qquad \forall\,\nu \in D(F). \tag{4.9}$$

Proof Let $\nu_0 = \mu_\tau$, $\nu_1 = \nu$ and consider the interpolating curve $\nu_t : [0, 1] \to \mathscr{P}_2(H)$ along which (4.1) holds. The minimality of μ_τ and (4.1) give

$$F(\mu_\tau) + \frac{1}{2\tau}\,W_2^2(\mu_\tau, \bar{\mu}) \leq F(\nu_t) + \frac{1}{2\tau}\,W_2^2(\nu_t, \bar{\mu})$$

$$\leq (1-t)\left[F(\mu_\tau) + \frac{1}{2\tau}\,W_2^2(\mu_\tau, \bar{\mu})\right]$$

$$+ t\left[F(\nu) + \frac{1}{2\tau}\,W_2^2(\nu, \bar{\mu})\right] - t(1-t)\frac{1}{2\tau}\,W_2^2(\mu_\tau, \nu).$$

Subtracting $F(\mu_\tau) + W_2^2(\mu_\tau, \bar{\mu})/2\tau$ from the left hand side of the first inequality and from the right hand side of the second inequality, and dividing by $t > 0$ we obtain:

$$F(\nu) - F(\mu_\tau) \geq \frac{1}{2\tau}\,W_2^2(\mu_\tau, \bar{\mu}) - \frac{1}{2\tau}\,W_2^2(\nu, \bar{\mu}) + \frac{1-t}{2\tau}\,W_2^2(\mu_\tau, \nu)$$

$$\geq \frac{1}{2\tau}\left[(1-t)\,W_2^2(\mu_\tau, \nu) - W_2^2(\nu, \bar{\mu})\right].$$

Letting $t \downarrow 0$ we have

$$W_2^2(\mu_\tau, \nu) - W_2^2(\nu, \bar{\mu}) + W_2^2(\mu_\tau, \bar{\mu}) \leq 2\tau\,(F(\nu) - F(\mu_\tau)), \tag{4.10}$$

which yields (4.9) by neglecting the nonnegative term $W_2^2(\mu_\tau, \bar{\mu})$. $\square$

5 Existence and uniqueness of gradient flows

In this section we prove existence and uniqueness of gradient flows and convergence of the approximations (4.2). Again the results of this section hold for more general classes of strongly displacement convex functionals, but we are only interested here in the case $F(\cdot) = \mathcal{H}(\cdot|\gamma)$, where we consider a fixed log-concave probability measure γ on H.

We go back to the sequence $(\mu_\tau^k)_k$ defined recursively by (4.2) with $\mu_\tau^0 = \overline{\mu} \in \mathscr{P}_2(H)$, the existence of $(\mu_\tau^k)_k$ being granted by Proposition 4.2. We shall denote the "discrete" semigroup induced by μ_τ^k by $\mathscr{S}_\tau \overline{\mu}(t)$, precisely

$$\mathscr{S}_\tau \overline{\mu}(t) := \mu_\tau^{k+1} \qquad \forall t \in (k\tau, (k+1)\tau]. \tag{5.1}$$

Theorem 5.1 (Existence and uniqueness of gradient flows) *For any $\overline{\mu} \in \mathscr{P}_2(K)$ there exists a unique gradient flow starting from $\overline{\mu}$. The induced semigroup $\mathscr{S}\overline{\mu}(t)$ satisfies*

$$W_2(\mathscr{S}\overline{\mu}(t), \mathscr{S}\overline{\mu}(s)) \le \sqrt{2F(\overline{\mu})}\sqrt{|t-s|}, \quad t, s \ge 0, \ \overline{\mu} \in \mathscr{P}_2(K) \tag{5.2}$$

and the following properties:

(i) (Uniform discrete approximation) $W_2(\mathscr{S}\overline{\mu}(t), \mathscr{S}_\tau \overline{\mu}(t)) \le C\sqrt{\tau F(\overline{\mu})}$ *if* $\overline{\mu} \in D(F)$*, with* $C = 2(2\sqrt{2}+1)$;

(ii) (Contractivity) $W_2(\mathscr{S}\overline{\mu}(t), \mathscr{S}\overline{\nu}(t)) \le W_2(\overline{\mu}, \overline{\nu})$;

(iii) (Regularizing effect) $F(\mathscr{S}\overline{\mu}(t)) \le \inf_{\nu \in D(F)} \frac{1}{2t} W_2^2(\overline{\mu}, \nu) + F(\nu) < +\infty$ *for all* $t > 0$, $\overline{\mu} \in \overline{D(F)}$.

Proof We first sketch the proof of uniqueness of gradient flows, referring to [3, Corollary 4.3.3] for all technical details: if $\mu^1(t)$, $\mu^2(t)$ are gradient flows starting from $\overline{\mu}$, setting $\nu = \mu^1(t)$ into

$$\frac{d}{dt} \frac{1}{2} W_2^2(\mu^2(t), \nu) \le F(\nu) - F\left(\mu^2(t)\right)$$

and $\nu = \mu^2(t)$ into

$$\frac{d}{dt} \frac{1}{2} W_2^2(\mu^1(t), \nu) \le F(\nu) - F(\mu^1(t)),$$

one obtains that $\frac{d}{dt} W_2^2(\mu^1(t), \mu^2(t)) \le 0$, whence the identity of μ^1 and μ^2 follows.

In order to show existence of gradient flows, we consider first the case when $\overline{\mu} \in D(F)$. Notice that μ_τ^{k+1} satisfies

$$F(\mu_\tau^{k+1}) + \frac{1}{2\tau} W_2^2(\mu_\tau^{k+1}, \mu_\tau^k) \le F(\nu) + \frac{1}{2\tau} W_2^2(\nu, \mu_\tau^k) \qquad \forall \nu \in D(F) \tag{5.3}$$

and, choosing in particular $\nu = \mu_\tau^k$, we obtain that $F\left(\mu_\tau^{k+1}\right) \leq F\left(\mu_\tau^k\right)$ and

$$W_2(\mu_\tau^{k+1}, \mu_\tau^k) \leq \sqrt{2\tau \left[F(\mu_\tau^k) - F\left(\mu_\tau^{k+1}\right)\right]}. \tag{5.4}$$

This inequality easily leads to the discrete $C^{1/2}$ estimate

$$W_2(\mathscr{S}_\tau\bar{\mu}(t), \mathscr{S}_\tau\bar{\mu}(s)) \leq \sqrt{2F(\bar{\mu})}\sqrt{|t - s + \tau|}. \tag{5.5}$$

Moreover, a crucial role is played by the formula

$$W_2^2(\mathscr{S}_\tau\bar{\mu}((k+1)\tau), \nu) - W_2^2(\mathscr{S}_\tau\bar{\mu}(k\tau), \nu) \leq 2\tau[F(\nu) - F(\mathscr{S}_\tau\bar{\mu}((k+1)\tau))] \tag{5.6}$$

for all $\nu \in D(F)$, which follows from Proposition 4.4.

We start proving the estimate

$$W_2^2(\mathscr{S}_\tau\bar{\mu}(t), \mathscr{S}_{\frac{\tau}{2}}\bar{\nu}(t)) - W_2^2(\bar{\mu}, \bar{\nu}) \leq 2\tau F(\bar{\nu}) \tag{5.7}$$

for all $\tau > 0$ and all times t that are integer multiples of τ. To this aim, from (5.6) we obtain the inequalities

$$W_2^2(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}\left(\tfrac{\tau}{2}\right), \theta) - W_2^2(\bar{\nu}, \theta) \quad \leq \tau\left[F(\theta) - F\left(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}\left(\tfrac{\tau}{2}\right)\right)\right], \tag{5.8}$$

$$W_2^2(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}(\tau), \theta) - W_2^2(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}\left(\tfrac{\tau}{2}\right), \theta) \leq \tau\left[F(\theta) - F(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}(\tau))\right], \tag{5.9}$$

for all $\theta \in D(F)$, whose sum gives

$$W_2^2(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}(\tau), \theta) - W_2^2(\bar{\nu}, \theta) \leq \tau\left[2F(\theta) - F\left(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}\left(\tfrac{\tau}{2}\right)\right) - F(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}(\tau))\right] \tag{5.10}$$

for all $\theta \in D(F)$. Still from (5.6) we get

$$W_2^2(\mathscr{S}_\tau\bar{\mu}(\tau), \theta) - W_2^2(\bar{\mu}, \theta) \leq 2\tau\left[F(\theta) - F(\mathscr{S}_\tau\bar{\mu}(\tau))\right] \quad \forall\theta \in D(F). \tag{5.11}$$

Setting $\theta = \mathscr{S}_\tau\bar{\mu}(\tau)$ in (5.10) and $\theta = \bar{\nu}$ in (5.11), we can add the resulting inequalities to obtain

$$W_2^2(\mathscr{S}_\tau\bar{\mu}(\tau), \mathscr{S}_{\frac{\tau}{2}}\bar{\nu}(\tau)) - W_2^2(\mathscr{S}_\tau\bar{\mu}(0), \mathscr{S}_{\frac{\tau}{2}}\bar{\nu}(0))$$
$$\leq \tau\left(2F(\bar{\nu}) - F\left(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}\left(\tfrac{\tau}{2}\right)\right) - F(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}(\tau))\right)$$
$$\leq 2\tau\left(F(\bar{\nu}) - F(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}(\tau))\right). \tag{5.12}$$

Notice that (5.12) corresponds to (5.7) with $t = \tau$; by adding the inequalities analogous to (5.12) between consecutive times $m\tau$, $(m+1)\tau$, for $m = 0, \ldots, N-1$, we obtain

$$W_2^2(\mathscr{S}_\tau\bar{\mu}(N\tau), \mathscr{S}_{\frac{\tau}{2}}\bar{\nu}(N\tau)) - W_2^2(\bar{\mu}, \bar{\nu}) \leq 2\tau\left(F(\bar{\nu}) - F(\mathscr{S}_{\frac{\tau}{2}}\bar{\nu}(N\tau))\right), \tag{5.13}$$

that yields (5.7) because F is nonnegative. Now, from (5.7) with $\bar{\mu} = \bar{\nu}$ we get

$$W_2(\mathscr{S}_{\frac{\tau}{2^m}}\bar{\mu}(t), \mathscr{S}_{\frac{\tau}{2^{m+1}}}\bar{\mu}(t)) \leq 2^{-m/2}\sqrt{2\tau F(\bar{\mu})}$$

for all t that are integer multiples of $\tau/2^m$, so that

$$W_2(\mathscr{S}_{\frac{\tau}{2^m}}\bar{\mu}(t), \mathscr{S}_{\frac{\tau}{2^n}}\bar{\mu}(t)) \leq \sum_{i=m}^{n-1} 2^{-i/2}\sqrt{2\tau F(\bar{\mu})} \tag{5.14}$$

for all $n > m \geq j$ and all t that is an integer multiple of $\tau/2^j$. For any such t (and therefore on a dense set of times) the sequence $(\mathscr{S}_{\frac{\tau}{2^n}}\bar{\mu}(t))$ has the Cauchy property and converges in $\mathscr{P}_2(H)$ to some limit, that we shall denote by $\mathscr{S}_{(\tau)}\bar{\mu}(t)$.

Using the discrete $C^{1/2}$ estimate (5.5) we obtain convergence for all times, as well as the uniform Hölder continuity (5.2) of $t \mapsto \mathscr{S}_{(\tau)}\bar{\mu}(t)$.

We prove now that $(\mathscr{S}_{(\tau)}\bar{\mu}(t))_{t\geq 0}$ is a gradient flow starting from $\bar{\mu}$. Indeed, we can read (5.6) as follows:

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mathscr{S}_\tau\bar{\mu}(t), \nu) \leq \tau \sum_{i=1}^{\infty}[F(\nu) - F(\mathscr{S}_\tau\bar{\mu}(i\tau))]\delta_{\frac{i}{\tau}}$$

for all $\nu \in D(F)$, in the sense of distributions. Passing to the limit as $n \to \infty$ in the previous inequality with τ replaced by $\tau/2^n$, the lower semicontinuity of F gives

$$\frac{d}{dt}\frac{1}{2}W_2^2(\mathscr{S}_{(\tau)}\bar{\mu}(t), \nu) \leq [F(\nu) - F(\mathscr{S}_{(\tau)}\bar{\mu}(t))] \quad \forall \nu \in D(F)$$

in the sense of distributions. This proves that $\mathscr{S}_{(\tau)}\bar{\mu}(t)$ is a gradient flow starting from $\bar{\mu}$, and since we proved that gradient flows are uniquely determined by the initial condition, from now on we shall denote $\mathscr{S}\bar{\mu}(t) = \mathscr{S}_{(\tau)}\bar{\mu}(t)$.

Proof of (i) Passing to the limit as $n \to \infty$ in (5.14), with $m = j = 0$, we obtain that $W_2(\mathscr{S}\bar{\mu}(t), \mathscr{S}_\tau\bar{\mu}(t))$ can be estimated with $2(\sqrt{2}+1)\sqrt{\tau F(\bar{\mu})}$ when t/τ is an integer. From (5.5), (5.2) and the triangle inequality we obtain (i).

Proof of (ii) *when* $\bar{\mu} \in D(F)$ It suffices to pass to the limit as $\tau \downarrow 0$ in (5.7).

Proof of (iii) *when* $\bar{\mu} \in D(F)$ By adding the inequalities

$$W_2^2(\mathscr{S}_\tau\bar{\mu}((i+1)\tau), \nu) - W_2^2(\mathscr{S}_\tau\bar{\mu}(i\tau), \nu) \leq 2\tau[F(\nu) - F(\mathscr{S}_\tau\bar{\mu}((i+1)\tau))]$$
$$\leq 2\tau[F(\nu) - F(\mathscr{S}_\tau\bar{\mu}(N\tau))]$$

for $i = 0, \ldots, N-1$ we get

$$W_2^2(\mathscr{S}_\tau\bar{\mu}(N\tau), \nu) - W_2^2(\bar{\mu}, \nu) \leq 2N\tau[F(\nu) - F(\mathscr{S}_\tau\bar{\mu}(N\tau))].$$

Replacing now τ by $\tau/2^m$ in this inequality, and defining N as the integer part of $2^m t/\tau$ (so that $N\tau/2^m \to t$), we can let $m \to \infty$ to obtain (iii), neglecting the term $W_2^2(\mathscr{S}_\tau\bar{\mu}(N\tau), \nu)$.

In order to prove (ii) and (iii) when $\bar{\mu} \in \overline{D(F)}$ we use a density argument. Indeed, let $\bar{\mu}_n \in D(F)$ be converging to $\bar{\mu} \in \overline{D(F)}$ in $\mathscr{P}_2(H)$: by (ii) we obtain that $\mathscr{S}\bar{\mu}_n(t)$ is a Cauchy sequence for all $t \geq 0$, and therefore converges to some limit, that we shall denote by $\mathscr{S}\bar{\mu}(t)$. It is not difficult to prove by approximation that $\mathscr{S}\bar{\mu}(t)$ is a gradient flow, and it remains to show that it starts from $\bar{\mu}$. We have indeed $W_2(\mathscr{S}\bar{\mu}(t), \mathscr{S}\bar{\mu}_n(t)) \leq W_2(\bar{\mu}_n, \bar{\mu})$, so that

$$\limsup_{t\downarrow 0} W_2(\mathscr{S}\bar{\mu}(t), \bar{\mu}) \leq 2W_2(\bar{\mu}_n, \bar{\mu}) + \limsup_{t\downarrow 0} W_2(\mathscr{S}\bar{\mu}_n(t), \bar{\mu}_n) = 2W_2(\bar{\mu}_n, \bar{\mu}).$$

Letting $n \to \infty$ we obtain that $\mathscr{S}\bar{\mu}(t) \to \bar{\mu}$ as $t \downarrow 0$. $\qquad\square$

6 Γ-convergence and stability properties

In this section we consider a sequence (γ_n) of log-concave probability measures on H weakly converging to γ and a sequence of Hilbertian norms on $H_n = H^0(\gamma_n)$ satisfying Assumption 1.4. We are going to prove that the gradient flows associated with $\mathcal{H}(\cdot|\gamma_n)$ with respect to W_{2,H^n} converge to the gradient flow associated with $\mathcal{H}(\cdot|\gamma)$ with respect to W_2, where the notation W_{2,H^n} has been introduced in (2.2).

This result is natural in view of Theorem 5.1, since the discrete approximating flow $\mathscr{S}_\tau^n\bar{\mu}^n(\cdot)$ of $\mathcal{H}(\cdot|\gamma_n)$ are defined only in terms of γ_n and W_{2,H^n}. However, the same result is much less obvious in view of the connection with the Fokker–Planck equation (3.4) and the associated stochastic process $(X_t)_{t\geq 0}$: see Sects. 7 and 8.

The main result of this section is the following:

Theorem 6.1 (Stability of gradient flows) *Suppose that $(\gamma_n) \subset \mathscr{P}(H)$ is a sequence of log-concave probability measures converging weakly to $\gamma \in \mathscr{P}(H)$ and that Assumption 1.4 holds. Let $\bar{\mu}^n \in \mathscr{P}_2(A_n)$ and let $(\mu_t^n)_{t\geq 0}$ be the gradient flows associated with $\mathcal{H}(\cdot|\gamma_n)$ in $\mathscr{P}_2(A_n)$ with respect to W_{2,H^n}.*
If $\bar{\mu}_n$ converge to $\bar{\mu} \in \mathscr{P}_2(A)$ in $\mathscr{P}_2(H)$ then $\mu_t^n \to \mu_t$ in $\mathscr{P}_2(H)$ for every $t \in [0, +\infty)$, where $(\mu_t)_{t\geq 0}$ is the gradient flow associated with $\mathcal{H}(\cdot|\gamma)$ in $\mathscr{P}_2(A)$ with respect to W_2.

The crucial property in the proof of this stability result is the Γ-convergence of the functionals $\mathcal{H}(\cdot|\gamma_n)$ to $\mathcal{H}(\cdot|\gamma)$. The concept of Γ-convergence is due to De Giorgi and is a classical tool of Calculus of Variations.

Lemma 6.2 (Convergence of entropy functionals) *If $\gamma_n \in \mathscr{P}(H)$ converge weakly to $\gamma \in \mathscr{P}(H)$ then $\mathcal{H}(\cdot|\gamma_n) : \mathscr{P}_2(H) \to [0, +\infty]$ Γ-converge to $\mathcal{H}(\cdot|\gamma) : \mathscr{P}_2(H) \to [0, +\infty]$, i.e.*

(i) *for any sequence $(\mu_n) \subset \mathscr{P}_2(H)$ converging <u>weakly</u> to $\mu \in \mathscr{P}_2(H)$, we have*

$$\liminf_{n\to\infty} \mathcal{H}(\mu_n|\gamma_n) \geq \mathcal{H}(\mu|\gamma); \qquad (6.1)$$

(ii) *for any $\mu \in \mathscr{P}_2(H)$ there exist $\mu_n \in \mathscr{P}_2(H)$ converging to μ <u>in $\mathscr{P}_2(H)$</u> such that*

$$\limsup_{n\to\infty} \mathcal{H}(\mu_n|\gamma_n) \leq \mathcal{H}(\mu|\gamma). \tag{6.2}$$

Proof The "liminf" inequality (i) in the definition of Γ-convergence follows directly from the duality formula (4.5): if $\mu_n \to \mu$ weakly, for all bounded continuous $S : H \to \mathbb{R}$ we have

$$\int_H S\,d\mu - \int_H e^{S-1}\,d\gamma = \lim_{n\to\infty}\left[\int_H S\,d\mu_n - \int_H e^{S-1}\,d\gamma_n\right] \leq \liminf_{n\to\infty} \mathcal{H}(\mu_n|\gamma_n).$$

Taking the supremum in the left hand side the lim inf inequality is achieved.

In order to show the lim sup inequality we first notice that, by diagonal arguments, we need only to show it for a dense subset $\mathscr{R} \subset D(\mathcal{H}(\cdot|\gamma))$; here density should be understood in the sense that for any $\nu \in D(\mathcal{H}(\cdot|\gamma))$ there exist $\nu_n \in \mathscr{R}$ converging to ν in $\mathscr{P}_2(H)$ with $\mathcal{H}(\nu_n|\gamma) \to \mathcal{H}(\nu|\gamma)$. Let us check that

$$\mathscr{R} := \left\{ e^{-\varepsilon\|\cdot\|_H^2} f\gamma \in \mathscr{P}_2(H) : \ f \in C_b(H), \ f \geq 0, \ \varepsilon > 0 \right\}$$

has these properties: indeed, in this case, given $\mu = g\gamma \in \mathscr{R}$ with $g(x) = e^{-\varepsilon\|x\|_H^2} f(x)$, we can simply take $\mu_n = Z_n^{-1} g\gamma_n$, with $Z_n := \int_H g\,d\gamma_n$, to achieve the lim sup inequality. The "density in energy" of $\mathscr{R}$ in the sense described above can be achieved as follows: first, using the density of $C_b(H)$ in $L^1(\gamma)$ and the dominated convergence theorem, we see that any $\mu = \rho\gamma \in \mathscr{P}_2(H)$ with $\rho \in L^\infty(\gamma)$ can be approximated by elements of $\mathscr{R}$. A truncation argument then gives that any $\mu \in D(\mathcal{H}(\cdot|\gamma))$ can be approximated. $\square$

In order to clarify the structure of the proof of Theorem 6.1, it is useful to introduce the following concept: we say that $\mu_n \in \mathscr{P}(H_n)$ converge with moments to $\mu \in \mathscr{P}(H)$ if $\mu_n \to \mu$ weakly in $\mathscr{P}(H)$ and $\int_{H_n} \|x\|_{H_n}^2\,d\mu_n \to \int_H \|x\|_H^2\,d\mu$. Notice that for any open set $A \subset H$ we can use (2.6) to obtain

$$\liminf_{n\to\infty} \int_A \|x\|_{H_n}^2\,d\mu_n = \liminf_{n\to\infty} \int_A \|\pi_n(x)\|_{H_n}^2\,d\mu_n \geq \int_A \|x\|_H^2\,d\mu \tag{6.3}$$

whenever $\mu_n \to \mu$ weakly in $\mathscr{P}(H)$. Therefore, in the proof of convergence with moments, only the lim sup needs to be proved. The following result is well-known.

Lemma 6.3 *Convergence with moments is equivalent to convergence in $\mathscr{P}_2(H)$.*

Proof If $\mu_n \to \mu$ with moments, (6.3) with $A = \{\|x\|_H < R\}$ gives

$$
\lim_{R \to \infty} \limsup_{n \to \infty} \int\limits_{\{\|x\|_H \geq R\}} \|x\|_H^2 \, d\mu_n \leq \kappa^2 \lim_{R \to \infty} \limsup_{n \to \infty} \int\limits_{\{\|x\|_H \geq R\}} \|\pi_n(x)\|_{H_n}^2 \, d\mu_n
$$

$$
\leq \kappa^2 \lim_{R \to \infty} \int\limits_{\{\|x\|_H \geq R\}} \|x\|_H^2 \, d\mu = 0.
$$

We obtain the convergence in $\mathscr{P}_2(H)$ from (2.4). Conversely, if $\mu_n \to \mu$ weakly, (2.6) gives

$$
\lim_{n \to \infty} \int\limits_{H_n} \|\pi_n(x)\|_{H_n}^2 \wedge R^2 \, d\mu_n = \int\limits_H \|x\|_H^2 \wedge R^2 \, d\mu \qquad \forall R > 0. \qquad (6.4)
$$

If $\mu_n \to \mu$ in $\mathscr{P}_2(H)$, we can use (2.4) and (1.10) to obtain $\limsup_n \int_{\{\|x\|_{H_n} \geq R\}} \|x\|_{H_n}^2 \, d\mu_n \to 0$ as $R \to \infty$, and if we combine this information with (6.4) we obtain the convergence with moments. $\qquad\square$

Lemma 6.4 *Assume that $\mu_n, \nu_n \in \mathscr{P}(H_n)$, that $\Sigma_n \in \Gamma(\mu_n, \nu_n)$ is converging to $\Sigma \in \Gamma(\mu, \nu)$ weakly and that $\mu_n \to \mu$ with moments, while $\int_{H_n} \|y\|_{H_n}^2 \, d\nu_n$ is bounded. Then*

$$
\lim_{n \to \infty} \int\limits_{H_n \times H_n} \langle x, y \rangle_{H_n} \, d\Sigma_n = \int\limits_{H \times H} \langle x, y \rangle_H \, d\Sigma.
$$

Proof We prove the lim inf inequality only, the proof of the other one being similar. With the notation of (1.11), we have $\langle \pi_n(x), \pi_n(y) \rangle_{H_n} \to \langle x, y \rangle_H$ as $n \to \infty$ for all $x, y \in H$. For all $\varepsilon > 0$ the functions

$$
\frac{1}{2\varepsilon} \|\pi_n(x)\|_{H_n}^2 + \frac{\varepsilon}{2} \|\pi_n(y)\|_{H_n}^2 + 2\langle \pi_n(x), \pi_n(y) \rangle_{H_n}
$$

are nonnegative, and these functions are equi-continuous in $H \times H$ by (1.10). Therefore (2.6), thanks to the convergence assumption on μ_n, gives

$$
\liminf_{n \to \infty} \int\limits_{H \times H} \frac{\varepsilon}{2} \|\pi_n(y)\|_{H_n}^2 + 2\langle \pi_n(x), \pi_n(y) \rangle_{H_n} \, d\Sigma_n \geq \int\limits_{H \times H} \frac{\varepsilon}{2} \|y\|_H^2 + 2\langle x, y \rangle_H \, d\Sigma.
$$

Using the boundedness assumption on (ν_n) we can obtain the lim inf inequality letting $\varepsilon \downarrow 0$. $\qquad\square$

In the proof of Theorem 6.1 we need some continuity/lower semicontinuity properties of the Wasserstein distance.

Lemma 6.5 *Let $\mu_n, \nu_n \in \mathscr{P}_2(H_n)$ be such that $\mu_n \to \mu \in \mathscr{P}_2(H)$, $\nu_n \to \nu \in \mathscr{P}_2(H)$ weakly in $\mathscr{P}(H)$. Then:*

   (i) $W_2(\mu, \nu) \leq \liminf\limits_{n\to\infty} W_{2,H_n}(\mu_n, \nu_n)$;

   (ii) *if $\mu_n \to \mu$ and $\nu_n \to \nu$ in $\mathscr{P}_2(H)$, then $W_{2,H_n}(\mu_n, \nu_n) \to W_2(\mu, \nu)$.*

Proof (i) Without loss of generality, we can assume that the lim inf is a limit. Let $\Sigma_n \in \Gamma_{H_n,o}(\mu_n, \nu_n)$. Notice that tightness of (μ_n) and (ν_n) in H implies tightness of (Σ_n) in $H \times H$. Let $\Sigma \in \Gamma(\mu, \nu)$ be a weak limit point of (Σ_n), which obviously belongs to $\Gamma(\mu, \nu)$. Then, taking into account the equi-continuity in $H \times H$ of the maps $\|\pi_n(x - y)\|_{H_n}$, ensured by (1.10), by (2.6) we get:

$$W_2^2(\mu, \nu) \leq \int\limits_{H \times H} \|y - x\|_H^2 \, d\Sigma \leq \liminf_{n\to\infty} \int\limits_{H \times H} \|\pi_n(x - y)\|_{H_n}^2 \, d\Sigma_n$$

$$= \liminf_{n\to\infty} \int\limits_{H_n \times H_n} \|y - x\|_{H_n}^2 \, d\Sigma_n = \liminf_{n\to\infty} W_{2,H_n}^2(\mu_n, \nu_n).$$

(ii) We choose optimal couplings Σ_n between μ_n and ν_n, relative to H_n, and prove that any weak limit Σ (which exists, possibly passing to subsequences) is optimal. The same truncation argument used in Lemma 6.3 to show that convergence in $\mathscr{P}_2(H)$ implies convergence with moments shows that

$$\lim_{n\to\infty} \int\limits_{H_n \times H_n} \|y - x\|_{H_n}^2 \, d\Sigma_n = \int\limits_{H \times H} \|y - x\|_H^2 \, d\Sigma.$$

In order to prove the optimality of Σ we recall that $\Sigma \in \Gamma(\lambda, \nu)$ is an optimal coupling (relative to the cost $c(x, y) = \|x - y\|_H^2$) if and only if for any $\ell \in \mathbb{N}$, any $(x_i, y_i)_{i=1,...,\ell}$ in the support of Σ and any permutation σ of $\{1, ..., \ell\}$ the following inequality holds:

$$\sum_{i=1}^{\ell} \|x_i - y_{\sigma(i)}\|_H^2 \geq \sum_{i=1}^{\ell} \|x_i - y_i\|_H^2, \tag{6.5}$$

see for instance [3, Theorem 6.1.4]. Since Σ_n is optimal, a similar inequality holds with $\|\cdot\|_{H_n}$ instead of $\|\cdot\|_H$ for all $(x_i^n, y_i^n)_{i=1,...,\ell}$ in the support of Σ_n. Since Σ_n converge to Σ weakly, for any $(x_i, y_i)_{i=1,...,\ell}$ in the support of Σ there exist $(x_i^n, y_i^n)_{i=1,...,\ell}$ in the support of Σ_n with $(x_i^n, y_i^n) \to (x_i, y_i)$ in $H \times H$. Then (6.5) follows taking limits as $n \to \infty$ and using the fact that $z_n \in H_n$ and $\|z_n - z\|_H \to 0$ implies $\|z_n\|_{H_n} \to \|z\|_H$. □

We can now prove Theorem 6.1. With no loss of generality we can assume (possibly making translations) that $A(\gamma_n) = H_n$.

Proof Set $F_n(\cdot) := \mathcal{H}(\cdot | \gamma_n)$. We consider the case when $F_n(\bar{\mu}^n)$ is bounded first. In this case, property (i) in Theorem 5.1 and (1.10) ensure the uniform (in time,

and with respect to n) estimate $W_2^2(\mu_t^n, \mathscr{S}_\tau^n \bar{\mu}^n(t)) \leq C\tau$. Here $\mathscr{S}_\tau^n \bar{\mu}^n(t)$ is the discrete approximation (5.1) of the gradient flow, obtained by the recursive minimization scheme (4.3): i.e. we define recursively $\mu_\tau^{n,0} := \bar{\mu}^n$, $\mu_\tau^{n,k+1}$ is the unique minimizer of

$$\mathscr{P}_2(H_n) \ni \nu \mapsto F_n(\nu) + \frac{1}{2\tau} W_{2,H_n}^2(\nu, \mu_\tau^{n,k}),$$

and we define $\mathscr{S}_\tau^n \bar{\mu}^n(t) := \mu_\tau^{n,k+1}$ for all $t \in (k\tau, (k+1)\tau]$. Therefore, taking also Lemma 6.3 into account, in this case it suffices to show that, with $\tau > 0$ *fixed*, the convergence with moments is preserved by the minimization scheme. So, let us assume that μ_n converge to μ with moments and $F_n(\mu_n)$ is bounded; we consider the minimizers ν_n of the problems

$$\mathscr{P}_2(H_n) \ni \sigma \mapsto F_n(\sigma) + \frac{1}{2\tau} W_{2,H_n}^2(\sigma, \mu_n),$$

and show that they converge with moments to the minimizer ν of the problem

$$\mathscr{P}_2(H) \ni \sigma \mapsto F(\sigma) + \frac{1}{2\tau} W_2^2(\sigma, \mu). \tag{6.6}$$

Notice first we can use $\lambda = \mu_n$ in the inequality

$$F_n(\nu_n) + \frac{1}{2\tau} W_{2,H_n}^2(\nu_n, \mu_n) \leq F_n(\lambda) + \frac{1}{2\tau} W_{2,H_n}^2(\lambda, \mu_n) \tag{6.7}$$

to obtain that both $\int_{H_n} \|y\|_{H_n}^2 \, d\nu_n$ and $F_n(\nu_n)$ are bounded. Since (γ_n) is tight and $\mathcal{H}(\nu_n|\gamma_n)$ is bounded, then (ν_n) is tight as well, by the entropy inequality (4.4). Therefore (ν_n) has limit points with respect to the weak convergence. We will prove that any limit point is a minimizer of (6.6), so that it must be ν.

Let $\nu' = \lim_k \nu_{n(k)}$ in the weak convergence, let $\lambda \in \mathscr{P}_2(H)$ and let λ_k be converging to λ in $\mathscr{P}_2(H)$, with $\limsup_k F_{n(k)}(\lambda_k) \leq F(\lambda)$, whose existence is ensured by condition (ii) in the definition of Γ-convergence. Setting $\lambda = \lambda_k$, $n = n(k)$ in (6.7), and using also condition (i) in the definition of Γ-convergence to bound $F_{n(k)}(\nu_{n(k)})$ from below, we get from (i) and (ii) of Lemma 6.5

$$\begin{aligned}
F(\nu') + \frac{1}{2\tau} W_2^2(\nu', \mu) &\leq \limsup_{k\to\infty} \left[F_{n(k)}(\nu_{n(k)}) + \frac{1}{2\tau} W_2^2(\nu_{n(k)}, \mu_{n(k)}) \right] \\
&\leq \limsup_{k\to\infty} \left[F_{n(k)}(\lambda_k) + \frac{1}{2\tau} W_2^2(\lambda_k, \mu_{n(k)}) \right] \\
&\leq F(\lambda) + \frac{1}{2\tau} W_2^2(\lambda, \nu).
\end{aligned} \tag{6.8}$$

As λ is arbitrary, this proves that ν' is a minimizer, therefore $\nu' = \nu$.

Now, setting $\lambda = \mu$ in (6.8), we obtain that all inequalities must be equalities, so that $\lim_k W_2^2(\mu_{n(k)}, \nu_{n(k)}) = W_2^2(\mu, \nu)$. Indeed, if $\limsup_k (a_k + b_k) \leq a + b$, $\liminf_k a_k \geq a$ and $\liminf_k b_k \geq b$, then $\lim_k a_k = a$ and $\lim_k b_k = b$.

We shall denote in the sequel by Σ_n optimal couplings between μ_n and ν_n. Let $\Sigma \in \Gamma(\mu, \nu)$ a limit point in the weak convergence of Σ_n, and assume just for notational simplicity that the whole sequence Σ_n weakly converges to Σ. By (2.6) we get

$$\int\limits_{H \times H} \|x - y\|_H^2 \, d\Sigma \leq \liminf_{n \to \infty} \int\limits_{H \times H} \|x - y\|_{H_n}^2 \, d\Sigma_n = W_2^2(\mu, \nu),$$

therefore $\Sigma \in \Gamma_o(\mu, \nu)$. We can now apply Lemma 6.4 to obtain that $\int_{H_n \times H_n} \langle x, y \rangle_{H_n} \, d\Sigma_n \to \int_{H \times H} \langle x, y \rangle \, d\Sigma$; from the identity

$$W_2^2(\mu, \nu) = \int\limits_H \|x\|_H^2 \, d\mu + \int\limits_H \|y\|_H^2 \, d\nu - 2 \int\limits_{H \times H} \langle x, y \rangle_H \, d\Sigma,$$

and from the analogous one with the Hilbert spaces H_n we obtain that ν_n converge with moments to ν.

In the general case when $F_n(\bar{\mu}_n)$ is not bounded we can find, for any $\varepsilon > 0$, $\bar{\nu} \in D(F)$ with $W_2(\bar{\mu}, \bar{\nu}) < \varepsilon$. By the definition of Γ-convergence we can also find $\bar{\nu}^n$ converging to $\bar{\nu}$ in $\mathscr{P}_2(H)$ with $\limsup_n F_n(\bar{\nu}_n) \leq F(\bar{\nu})$. For n large enough we still have $W_{2,H_n}(\bar{\mu}^n, \bar{\nu}^n) < \varepsilon$, so that the contracting property of gradient flows (see Theorem 5.1 (ii)) gives

$$W_2(\mathscr{S}\bar{\mu}^n(t), \mathscr{S}\bar{\nu}^n(t)) + W_2(\mathscr{S}\bar{\mu}(t), \mathscr{S}\bar{\nu}(t))$$
$$\leq \kappa W_{2,H_n}(\mathscr{S}\bar{\mu}^n(t), \mathscr{S}\bar{\nu}^n(t)) + W_2(\mathscr{S}\bar{\mu}(t), \mathscr{S}\bar{\nu}(t)) < (\kappa + 1)\varepsilon, \quad \forall t \geq 0.$$

By applying the local uniform convergence property to $\bar{\nu}^n$ we get

$$\limsup_{n \to \infty} \sup_{t \in [0,T]} W_2(\mathscr{S}\bar{\mu}^n(t), \mathscr{S}\bar{\mu}(t)) \leq (\kappa + 1)\varepsilon \quad \forall T > 0.$$

$\square$

7 Wasserstein semigroup and Dirichlet forms

In this section we establish a general link between the Wassertein semigroups and the semigroups arising from natural "gradient" Dirichlet forms, extending Proposition 3.2 to the general case of a log-concave measure γ in H. We denote by K the support of γ (a closed convex set, coinciding with $\overline{\{V < +\infty\}}$ when $H = \mathbb{R}^k$ and $\gamma = \exp(-V)\mathscr{L}^k$) and, without a real loss of generality, we consider the case when

$$A(\gamma) = H^0(\gamma) = H. \tag{7.1}$$

We consider, recalling (1.4), the bilinear form

$$\mathcal{E}_\gamma(u, v) := \int_H \langle \nabla u, \nabla v \rangle \, d\gamma, \qquad u, v \in C_b^1(H). \tag{7.2}$$

Accordingly, we define the induced scalar product and norm on $C_b^1(H)$:

$$\mathcal{E}_{\gamma,1}(u, v) := \int_H u \, v \, d\gamma + \mathcal{E}_\gamma(u, v), \qquad \|u\|_{\mathcal{E}_{\gamma,1}} := \sqrt{\mathcal{E}_{\gamma,1}(u, u)}. \tag{7.3}$$

We start proving that $\mathcal{E}_\gamma$ is closable. We recall that closability means the following: for all sequences $(u_n) \subset C_b^1(H)$ which are Cauchy with respect to $\| \cdot \|_{\mathcal{E}_{\gamma,1}}$ and such that $u_n \to 0$ in $L^2(\gamma)$, we have $\|u_n\|_{\mathcal{E}_{\gamma,1}} \to 0$. This is equivalent to saying that the operator $\nabla : C_b^1(H) \mapsto L^2(\gamma; H)$ is closable in $L^2(\gamma)$.

The following theorem states closability of $(\mathcal{E}_\gamma, C_b^1(H))$ in $L^2(\gamma)$ for any log-concave γ, a rather unusual level of generality for such kind of result. Indeed, our technique of proof seems quite innovative and uses a variational characterization of $\mathcal{E}_\gamma$ in terms of optimal transport and relative entropy, see also Lemma A.1 below. Previously known results for log-concave measures needed much stronger assumptions: see [1].

Theorem 7.1 (Closability) *The bilinear form $(\mathcal{E}_\gamma, C_b^1(H))$ is closable in $L^2(\gamma)$.*

Proof Let us denote by $\mathrm{Cyl}(H)$ the subspace of $C_b^1(H)$ made by cylindrical functions; by a simple density argument we can assume that the sequence (u_n) is contained in $\mathrm{Cyl}(H)$. We claim that closability follows by the lower semicontinuity of $v \mapsto \mathcal{E}_\gamma(v, v)$ on $\mathrm{Cyl}(H)$: indeed, if this lower semicontinuity property holds, we can pass to the limit as $m \to \infty$ in the inequality $\mathcal{E}_\gamma(u_n - u_m, u_n - u_m) < \varepsilon$, for $n, m \geq n(\varepsilon)$, to obtain $\mathcal{E}_\gamma(u_n, u_n) < \varepsilon$ for $n \geq n(\varepsilon)$, i.e. $\|u_n\|_{\mathcal{E}_{\gamma,1}} \to 0$.

So, let $(v_n) \subset \mathrm{Cyl}(H)$ be converging in $L^2(\gamma)$ to $v \in \mathrm{Cyl}(H)$ and let us prove that the inequality $\liminf_n \mathcal{E}_\gamma(v_n, v_n) \geq \mathcal{E}_\gamma(v, v)$ holds.

We recall that γ belongs to $\mathscr{P}_2(H)$, by (1.3). We can assume, possibly adding and multiplying by constants, that $m := \inf v > 0$ and $\int v^2 \, d\gamma = 1$. By a simple truncation argument we can also assume that $\inf v_n \geq m/2$, $\sup v_n \leq \sup v + 1$, and set $w_n = v_n/\|v_n\|_2$; obviously $w_n \to v$ in $L^2(\gamma)$ and, as a consequence, $w_n^2 \gamma \to v^2 \gamma$ weakly. By Lemma A.1 we get

$$\mathcal{H}(\mu|\gamma) \geq \mathcal{H}(w_n^2 \gamma | \gamma) - 2\sqrt{\mathcal{E}_\gamma(w_n, w_n)} \, W_2(\mu, w_n^2 \gamma) \qquad \forall \mu \in \mathscr{P}_2(H).$$

The uniform upper bound on w_n ensures, taking (2.4) into account, that $w_n^2 \gamma \to v^2 \gamma$ in $\mathscr{P}_2(H)$. Passing to the limit as $n \to \infty$, the lower semicontinuity of the relative Entropy gives

$$\mathcal{H}(\mu|\gamma) \geq \mathcal{H}(v^2 \gamma | \gamma) - 2 \liminf_{n \to \infty} \sqrt{\mathcal{E}_\gamma(w_n, w_n)} \, W_2(\mu, v^2 \gamma) \qquad \forall \mu \in \mathscr{P}_2(H).$$

By applying Lemma A.1 again we get $\liminf_n \mathcal{E}_\gamma(w_n, w_n) \geq \mathcal{E}_\gamma(v, v)$, and from the definition of w_n we see that the same inequality holds if we replace w_n with v_n. □

Being $\mathcal{E}_\gamma$ closable, we shall denote by $D(\mathcal{E}_\gamma)$ its domain (i.e. the closure of $C_b^1(H)$ with respect to the norm $\| \cdot \|_{\mathcal{E}_{\gamma,1}}$), which obviously can be identified with a subset of $L^2(\gamma)$, and keep the notation $\mathcal{E}_\gamma$ for the extension of $\mathcal{E}_\gamma$ to $D(\mathcal{E}_\gamma) \times D(\mathcal{E}_\gamma)$. In the next lemma we show that $D(\mathcal{E}_\gamma)$ contains $\mathrm{Lip}_b(K)$ and some useful representation formulas for the extension.

Recall that a *finite signed measure* is an $\mathbb{R}$-valued set function defined on Borel sets that can be written as the difference of two positive finite measures; by Hahn decomposition, any such measure μ can be uniquely written as $\mu = \mu^+ - \mu^-$, with $\mu^\pm$ nonnegative and $\mu^+ \perp \mu^-$. The *total variation* $|\mu|$ is the finite measure defined by $\mu^+ + \mu^-$.

Lemma 7.2 ($\mathcal{E}_\gamma$ *is a Dirichlet form*) $\mathcal{E}_\gamma$ *is a Dirichlet form,* $\mathrm{Lip}_b(K) \subset D(\mathcal{E}_\gamma)$ *and*

$$\sqrt{\mathcal{E}_\gamma(u, u)} \leq [u]_{\mathrm{Lip(K)}} \qquad \forall u \in \mathrm{Lip}_b(K). \tag{7.4}$$

Moreover, the following properties hold:

(i) *if H is finite-dimensional, $h \in H$ and $\ell_h(x) = \langle h, x \rangle$, there exists a finite signed measure Σ_h in H supported on K such that*

$$\mathcal{E}_\gamma(u, \ell_h) = \int_H u \, d\Sigma_h \qquad \forall u \in \mathrm{Lip}_b(K); \tag{7.5}$$

(ii) *if $\pi : H \to L$ is a finite-dimensional orthogonal projection, then*

$$\mathcal{E}_\gamma(u \circ \pi, v \circ \pi) = \mathcal{E}_{\pi_\# \gamma}(u, v) \qquad \forall u, v \in \mathrm{Lip}_b(L). \tag{7.6}$$

Proof Let $u \in \mathrm{Lip}_b(K)$ and let $\tilde{u}$ be a bounded Lipschitz extension of u to the whole of H. Combining finite-dimensional approximation and smoothing, we can easily find a sequence $(u_n) \subset C_b^1(H)$ converging to $\tilde{u}$ pointwise and with $[u_n]_{\mathrm{Lip}(H)}$ bounded. It follows that $u_n \to \tilde{u}$ in $L^2(\gamma)$ and, possibly extracting a subsequence, $\nabla u_n \to U$ weakly in $L^2(\gamma; H)$. Then, a sequence (g_n) of convex combinations of u_n still converges to $\tilde{u}$ in $L^2(\gamma; H)$ and is Cauchy with respect to $\| \cdot \|_{\mathcal{E}_{\gamma,1}}$. It follows that $u \in D(\mathcal{E}_\gamma)$. A similar argument proves (7.4) and the fact that $\mathcal{E}_\gamma(\phi(u), \phi(u))$ is less than $\mathcal{E}_\gamma(u, u)$ whenever $u \in D(\mathcal{E}_\gamma)$ and $\phi : \mathbb{R} \to \mathbb{R}$ is 1-Lipschitz. This last property shows that $\mathcal{E}_\gamma$ is a Dirichlet form.

Now, let $H = \mathbb{R}^k$, so that $\gamma = \exp(-V)\mathcal{L}^k$, and let us prove (i). By the closability of $\mathcal{E}_\gamma$, we need only to prove that

$$\int_{\mathbb{R}^k} \langle \nabla u, h \rangle \, d\gamma = \int_{\mathbb{R}^k} u \, d\Sigma_h \qquad \forall u \in C_b^1(\mathbb{R}^k) \tag{7.7}$$

for some finite signed measure Σ_h. The existence of such a measure Σ_h (obvious in the case when ∇V is Lipschitz, as an integration by parts gives $\Sigma_h = -\langle \nabla V, h \rangle \gamma$) is ensured by Proposition A.2.

Finally, notice that (7.6) trivially holds by the definitions of $\mathcal{E}_\gamma$ and $\mathcal{E}_{\pi_\# \gamma}$ when $u \in C_b^1(L)$, because $u \circ \pi \in C_b^1(H)$. By approximation the equality extends to the case $u, v \in \mathrm{Lip}_b(L)$. □

By the previous lemma, there exists a unique contraction semigroup P_t in $L^2(\gamma)$ associated to $\mathcal{E}_\gamma$. We are now going to compare it with the Wasserstein semigroup $\mathscr{S}\mu(t)$ of Theorem 5.1, and we shall denote in the sequel $v_t^x := \mathscr{S}\delta_x(t)$.

Theorem 7.3 *The semigroup P_t is regularizing from $L^\infty(\gamma)$ to $C_b(K)$, and the identity*

$$P_t f(x) := \int_H f \, dv_t^x, \quad t > 0, \ f \in L^\infty(\gamma), \ x \in K \tag{7.8}$$

provides a continuous version of $P_t f$. In addition, P_t acts on $\mathrm{Lip}_b(K)$:

$$[P_t f]_{\mathrm{Lip}(K)} \le [f]_{\mathrm{Lip}(K)} \quad t \ge 0, \ \forall f \in \mathrm{Lip}_b(K). \tag{7.9}$$

Moreover, for any $\mu \in \mathscr{P}_2(K)$, we have the identity

$$\mathscr{S}\mu(t) = \int v_t^x \, d\mu(x), \quad \forall t \ge 0. \tag{7.10}$$

Proof Assuming (7.8), let us first show why it provides a continuous version of P_t: if $x_n \to x$, and we denote by ρ_n the densities of $v_t^{x_n}$ with respect to γ, whose existence is ensured by the estimate Theorem 5.1(iii), the contracting property of the semigroup gives that $\rho_n \gamma \to \rho \gamma$ weakly, where ρ is the density of v_t^x with respect to γ. On the other hand, the same estimate shows that $\mathcal{H}(\rho_n \gamma | \gamma)$ are uniformly bounded, therefore ρ_n are equi-integrable in $L^1(\gamma)$ and weakly converge in $L^1(\gamma)$ to ρ. This proves that the right hand side in (7.8) is continuous. Finally, (7.9) is a direct consequence of (7.8) and Theorem 6.1(iii): indeed, choosing $\Sigma \in \Gamma_o(v_t^x, v_t^y)$, we get

$$|P_f(x) - P_t f(y)| = \left| \int_H f \, dv_t^x - \int_H f \, dv_t^y \right| = \left| \int_H (f(u) - f(v)) \, d\Sigma(u, v) \right|$$

$$\le [f]_{\mathrm{Lip}(K)} \int_H \|u - v\| \, d\Sigma(u, v) \le [f]_{\mathrm{Lip}(K)} W_2(v_t^x, v_t^y)$$

$$\le [f]_{\mathrm{Lip}(K)} \|x - y\|.$$

Step 1: the general finite-dimensional case $H = \mathbb{R}^k$. It suffices to show that the class of convex V's for which the equivalence (7.8) holds for the probability measure $\gamma = Z^{-1} \exp(-V)\mathscr{L}^k$, is closed under monotone convergence.

Indeed, if V is smooth with ∇V and all its derivatives bounded, then we know from Sect. 3 that the Wasserstein semigroup coincides with the FP semigroup; therefore from point 3 of Proposition 3.2 we obtain (7.8), and (7.9) corresponds to (3.16).

By a convolution approximation, we extend the result to all convex Lipschitz functions V with $\int \exp(-V)\,dx < \infty$; indeed, if ρ_n is the density of $\mathcal{N}(0, n^{-1}I)$ with respect to $\mathscr{L}^k$, where I is the identity matrix in $\mathbb{R}^k$, then $V_n := V * \rho_n$ forms an increasing sequence by convexity of V. Eventually we obtain all convex functions V with $\int \exp(-V)\,dx < \infty$ using the fact that they can be represented as the supremum (see for instance [10]) of countably many affine functions ℓ_i, and applying the equivalence to $V_n := \max_{1 \le i \le n} \ell_i$ (notice that for n large enough V_n has at least linear growth at infinity).

So, let us consider a log-concave probability measure $\gamma = \exp(-V)\mathscr{L}^k$ and a sequence $V_n \uparrow V$, with V_n real-valued and V_1 having at least a linear growth at infinity, such that the statement of the theorem holds for all measures $\gamma_n = Z_n^{-1} \exp(-V_n)\mathscr{L}^k$; obviously the normalization constants Z_n converge to 1 and $\gamma_n \to \gamma$ weakly. Notice also that $\operatorname{supp}\gamma_n = \mathbb{R}^k$.

We will also use the fact that both sides in (7.8) are continuous with respect to γ-almost sure and dominated convergence, so we need only to check the identity when $f \in \operatorname{Lip}_b(K)$. We recall that, in general, the semigroup $P_t f$ is related to the resolvent family $R_\lambda f$ by

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} P_t f(x)\,dt \qquad x \in K, \ f \in C_b(K). \tag{7.11}$$

We define the bilinear form

$$\mathcal{E}^n(u, v) := \int_{\mathbb{R}^k} \langle \nabla u, \nabla v \rangle \, d\gamma_n, \qquad u, v \in C_b^1(\mathbb{R}^k).$$

Moreover, we denote by R_λ^n the resolvent family of $\mathcal{E}^n$, again related to the semigroup P_t^n on $C_b(\mathbb{R}^k)$ relative to $\mathcal{E}^n$ by $\int_0^\infty \exp(-\lambda t) P_t^n\,dt$. Using the representation (7.8) of P_t^n, by Theorem 6.1 we know that, for all $f \in C_b(\mathbb{R}^k)$, $R_\lambda^n f$ pointwise converge, on K, to the function $F_\lambda f$ defined by

$$F_\lambda f(x) := \int_0^\infty e^{-\lambda t} \int_H f\, v_t^x\,dt \qquad x \in K.$$

We are going to show that $F_\lambda f$ coincides with $R_\lambda f$, the resolvent family of $\mathcal{E}_\gamma$, for all $\operatorname{Lip}_b(K)$, so that

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} \int_{\mathbb{R}^k} f\, dv_t^x\,dt \qquad x \in K, \ f \in \operatorname{Lip}_b(K).$$

Since, by the injectivity of the Laplace transform, (7.11) uniquely determines the semigroup P_t on $\mathrm{Lip}_b(K)$ (7.8), would be achieved.

So, let $f \in \mathrm{Lip}_b(K)$; possibly replacing f by a Lipschitz extension to the whole of $\mathbb{R}^k$ with the same Lipschitz constant, we can assume that $f \in \mathrm{Lip}_b(\mathbb{R}^k)$ and $[f]_{\mathrm{Lip}(\mathbb{R}^k)} = [f]_{\mathrm{Lip}(K)}$ (indeed, neither $F_\lambda f$ nor $P_t f$ depend on this extension). By applying (7.9) to γ_n one obtains that $\lambda[R_\lambda^n f]_{\mathrm{Lip}(\mathbb{R}^k)} \leq [f]_{\mathrm{Lip}(K)}$, hence $F_\lambda f \in \mathrm{Lip}_b(K)$ and Lemma 7.2 gives $F_\lambda f \in D(\mathcal{E}_\gamma)$. Now, in order to prove that $F_\lambda f$ coincides with $R_\lambda f$, by a density argument it is enough to show that

$$\lambda \int_{\mathbb{R}^k} F_\lambda f \, v \, d\gamma + \mathcal{E}_\gamma(F_\lambda f, v) = \int_{\mathbb{R}^k} f v \, d\gamma \quad \forall v \in C_c^2(\mathbb{R}^k). \tag{7.12}$$

Our strategy is to pass to the limit as $n \to \infty$ in:

$$\int_{\mathbb{R}^k} f v \, d\gamma_n = \lambda \int_{\mathbb{R}^k} R_\lambda^n f \, v \, d\gamma_n + \mathcal{E}^n(R_\lambda^n f, v) \quad \forall v \in C_c^2(\mathbb{R}^k). \tag{7.13}$$

Let $(e_1, \ldots, e_k)$ be the canonical basis of $\mathbb{R}^k$. By applying the integration by parts formula (7.5) with $h = e_i$ and $u = \langle \nabla v, e_i \rangle R_\lambda^n f$, we get

$$\int_{\mathbb{R}^k} f v \, d\gamma_n = \int_{\mathbb{R}^k} R_\lambda^n f(\lambda v - \Delta v) \, d\gamma_n + \sum_{i=1}^k \int_{\mathbb{R}^k} R_\lambda^n f \, \langle \nabla v, e_i \rangle \, d\Sigma_{e_i}^n \tag{7.14}$$

where $\Sigma_{e_i}^n$ are associated to the measure γ_n. The crucial fact is now the following: we can apply Lemma 7.4 to

$$\sigma_n := (\lambda v - \Delta v) \, d\gamma_n + \sum_{i=1}^k \langle \nabla v, e_i \rangle \, d\Sigma_{e_i}^n,$$

$$\sigma_\infty := (\lambda v - \Delta v) \, d\gamma + \sum_{i=1}^k \langle \nabla v, e_i \rangle \, d\Sigma_{e_i},$$

with $\varphi_n(x) := R_\lambda^n f(x)$, $\varphi_\infty(x) = F_\lambda f(x)$ and Σ_{e_i} associated to the measure γ. Indeed, assumptions (i) and (ii) of the Lemma 7.4 are guaranteed by Proposition A.3 in the Appendix, while (iii) and (iv) hold trivially. Therefore, by (7.14) we have:

$$\int_{\mathbb{R}^k} f v \, d\gamma = \lim_{n \to \infty} \int_{\mathbb{R}^k} f v \, d\gamma_n = \int_{\mathbb{R}^k} F_\lambda f \, (\lambda v - \Delta v) \, d\gamma + \sum_{i=1}^k \int_{\mathbb{R}^k} F_\lambda f \, \langle \nabla v, e_i \rangle \, d\Sigma_{e_i}.$$

Again, by the integration by parts formula (7.5) shows that the last expression is equal to the right hand side in (7.12). This proves that $F_\lambda f = R_\lambda f$ on K for all $f \in \mathrm{Lip}_b(K)$.

Notice now that (7.10) holds for smooth V by (3.3) and Proposition 3.4. By approximation, using the stability result of Theorem 6.1, we obtain the general case.

Step 2: from the finite-dimensional to the infinite-dimensional case. We fix a complete orthonormal system $\{e_i\}_{i \geq 1}$ in H and we set $H_n := \text{span}\{e_1, \ldots, e_n\}$, denoting as usual by $\pi_n : H \to H_n$ the finite-dimensional projections. Setting $\gamma_n = (\pi_n)_{\#}\gamma$, it is immediate to check that γ_n is log-concave in H_n and that $H^0(\gamma_n) = H_n$ (if not, we would get that $H^0(\gamma)$ is contained in a proper subspace of H, contradicting (7.1)). We set:

$$\mathcal{E}^n(u, v) := \int_{H_n} \langle \nabla u, \nabla v \rangle \, d\gamma_n \qquad u, v \in C_b^1(H_n).$$

We should rather write $\nabla_{H_n} u$ for $u \in C_b^1(H_n)$, but since the scalar product of H_n is induced by H there is no ambiguity in writing $\nabla u : H_n \mapsto H_n$. We denote by $(R_\lambda)_{\lambda > 0}, P_t$ [respectively, $(R_\lambda^n)_{\lambda > 0}, P_t^n$] the resolvent family and the semigroup of $\mathcal{E}_\gamma$ (resp. $\mathcal{E}^n$). We also know, by the previous step, that $R_\lambda^n f$ is representable on $C_b(K_n)$ by $\int_0^\infty \exp(-\lambda t) \int_{H_n} f \, d\nu_t^{n,x} \, dt$: here K_n denotes the support of γ_n and $\nu_t^{n,x}$ the associated Wasserstein semigroup in $\mathscr{P}_2(K_n)$. Since $\gamma_n = (\pi_n)_{\#}\gamma$, we have $\pi_n(x) \in K_n$ for all $x \in K$. As a consequence, by Theorem 6.1 we obtain that

$$\lim_{n \to \infty} R_\lambda^n f(\pi_n(x)) = \lim_{n \to \infty} \int_0^\infty e^{-\lambda t} \int_H f \, d\nu_t^{n,\pi_n(x)}$$

$$dt = \int_0^\infty e^{-\lambda t} \int_H f \, d\nu_t^x \, dt \qquad \forall \, x \in K$$

for all $f \in C_b(H)$. We shall denote, as in Step 1, by $F_\lambda f$ the right-hand side. Our strategy will be, again, to show that $R_\lambda f(x) = F_\lambda f(x)$. We assume first that $f = g \circ \pi_k$ is cylindrical function, with $g \in \text{Lip}_b(H_k)$; by applying (7.9) to γ_n one obtains that $\lambda[(R_\lambda^n f) \circ \pi_n]_{\text{Lip}(K)} \leq [f]_{\text{Lip}(K_n)} \leq [g]_{\text{Lip}(H_k)}$, hence $F_\lambda f \in \text{Lip}_b(K)$ and Lemma 7.2 gives $F_\lambda f \in D(\mathcal{E}_\gamma)$.

Now, let $v = u \circ \pi_l$ and $u \in C_b^1(H_l)$; for $n \geq \max\{k, l\}$, taking into account (7.6) and the identities $f = f \circ \pi_n$, $v = v \circ \pi_n$, we have

$$\int_H vf \, d\gamma = \int_{H_n} vf \, d\gamma_n = \lambda \int_{H_n} v R_\lambda^n f \, d\gamma_n + \mathcal{E}^n(v, R_\lambda^n f)$$

$$= \lambda \int_H v(R_\lambda^n f) \circ \pi_n \, d\gamma + \mathcal{E}_\gamma(v, (R_\lambda^n f) \circ \pi_n). \qquad (7.15)$$

Now, $(R_\lambda^n f) \circ \pi_n$ converge to $F_\lambda f$ in $L^2(\gamma)$ and is bounded with respect to the norm $\| \cdot \|_{\mathcal{E}_{\gamma,1}}$, by the uniform Lipschitz bound. Therefore, by the closability of $\mathcal{E}_\gamma$,

$(R_\lambda^n f) \circ \pi_n \to F_\lambda f$ in the weak topology of $D(\mathcal{E})$. Thus, we can passing to the limit as $n \to \infty$ in (7.15) to obtain

$$\int_H v f \, d\gamma = \lambda \int_H v F_\lambda f \, d\gamma + \mathcal{E}_\gamma(v, F_\lambda f) \qquad \forall v = u \circ \pi_l, \ u \in C_b^1(H_l).$$

By the $L^2(\gamma)$ density of C_b^1 cylindrical functions v, we obtain $R_\lambda f(x) = F_\lambda f$ for all Lipschitz and bounded cylindrical functions f. As a consequence, (7.8) holds for this class of functions. Since both sides in (7.8) are continuous with respect to γ-almost sure and dominated convergence, again a density argument shows that the equality (7.8) extends to all $f \in L^\infty(\gamma)$.

By the previous step, we know that (7.10) holds for the finite-dimensional case. By approximation, using the stability result of Theorem 6.1, the contractivity of gradient flows of Theorem 5.1 and Lemma 7.4 below, we obtain the general case. $\qquad \square$

In the proof of Theorem 7.3 we also used the following result.

Lemma 7.4 *Let σ_n, σ_∞ be finite signed measures on H and let φ_n, $\varphi_\infty : H \to \mathbb{R}$ satisfy*:

(i) $\sup_n |\sigma_n|(H) < +\infty$ *and*

$$\lim_{n \to \infty} \int \varphi \, d\sigma_n = \int \varphi \, d\sigma_\infty \qquad \forall \varphi \in C_b(H); \tag{7.16}$$

(ii) *there exist compacts sets $J_m \subset H$ such that $\sup_n |\sigma_n|(H \setminus J_m) \to 0$ as $m \to \infty$;*

(iii) *$\{\varphi_n\}_{n \in \mathbb{N} \cup \{\infty\}}$ is equi-bounded and equi-continuous;*

(iv) *φ_n converge pointwise to φ_∞ on $\operatorname{supp} \sigma_\infty$.*

Then $\lim_n \int_H \varphi_n \, d\sigma_n = \int_H \varphi_\infty \, d\sigma_\infty$.

Proof Without loss of generality we can assume that $\lim_n \int_H \varphi_n \, d\sigma_n$ exists (so that we can freely extract subsequences) and $|\varphi_n| \leq 1$, $|\varphi_\infty| \leq 1$. Let us fix m and assume, possibly extracting a subsequence, that $\varphi_n \to \psi_m$ uniformly on J_m as $n \to \infty$; obviously $\varphi_\infty = \psi_m$ on $J_m \cap \operatorname{supp} \sigma_\infty$. We extend ψ_m continuously to the whole of H with $|\psi_m| \leq 1$. Then:

$$\left| \int_H \varphi_n \, d\sigma_n - \int_H \varphi_\infty \, d\sigma_\infty \right| \leq \left| \int_H \varphi_n \, d\sigma_n - \int_H \psi_m \, d\sigma_n \right| + \left| \int_H \psi_m \, d\sigma_n - \int_H \psi_m \, d\sigma_\infty \right|$$

$$+ \left| \int_H \psi_m \, d\sigma_\infty - \int_H \varphi_\infty \, d\sigma_\infty \right|.$$

The first term in the right hand side can be estimated, splitting the integration on J_m and on $H \setminus J_m$, with $\sup_{J_m} |\varphi_n - \psi_m| |\sigma_n|(H) + 2|\sigma_n|(H \setminus J_m)$. The second term tends to 0 as $n \to \infty$ by our first assumption, while the third one can be estimated with

$2|\sigma_\infty|(H \setminus J_m)$. Therefore, taking first the limsup as $n \to \infty$ and then letting $m \to \infty$ we have the thesis. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ □

Remark 7.5 In Lemma 7.4, (ii) is actually a consequence of (i): indeed, an extension of the classical Prohorov theorem holds for signed measures: see [7], Theorem 8.6.2; in our case (see Proposition A.3), in order to prove (i) we actually need to show first (ii), and that is why we prefer to include both assumptions in Lemma 7.4. Moreover, in (i), the condition $\sup_n |\sigma_n|(H) < +\infty$ is implied by (7.16) and the Banach–Steinhaus theorem.

Remark 7.6 (Continuity of P_t) The $\mathscr{P}_2(H)$-continuity of $t \mapsto \mathscr{S}\delta_x(t)$ in $[0, +\infty[$ shows that $P_t f \to f$ pointwise in K as $t \downarrow 0$ for all functions $f \in C(K)$ with at most quadratic growth at infinity, and in particular for $f \in C_b(K)$. Taking (7.9) into account, the convergence is uniform on compact subsets of H if $f \in \text{Lip}_b(K)$; by density, $P_t f \to f$ uniformly on compacts sets as $t \downarrow 0$ for all $f \in UC_b(K)$, the space of bounded *uniformly* continuous functions on H. It is also possible to show the regularizing effect $P_t(C_b(K)) \subset UC_b(K)$ for $t > 0$: indeed, the finite-dimensional smooth systems are easily seen to be Strong Feller (see Sect. 7.1 of [15]), and this property extends to the general case by approximation.

8 The Markov process

In this section we complete the proofs of Theorems 1.2, 1.3, 1.5 and 1.6, proving the existence of a unique Markov family of probability measures $\mathbb{P}_x$ on $K^{[0,+\infty[}$ satisfying

$$P_t f(x) = \mathbb{E}_x[f(X_t)] \qquad \forall x \in K \tag{8.1}$$

for all bounded Borel functions f. The continuity of $\mathbb{P}_x$ will be a consequence of the regularizing properties of the Wasserstein semigroup $\mathscr{S}\mu(t)$ (in particular the continuity of $x \mapsto \mathscr{S}\delta_x(t)$ will play an important role). The regularity property (8.5), instead, is based on general results from [27], that provide a Markov family satisfying a weaker property, and on the continuity of $t \mapsto \mathscr{S}\delta_x(t)$. As in the previous section we will use the notation v_t^x for $\mathscr{S}\delta_x(t)$.

Recall also that the regularizing estimate (iii) in Theorem 5.1 gives $v_t^x \ll \gamma$ for all $x \in K$; by the uniform bound on the relative entropy (which yields equi-integrability of the densities), Dunford–Pettis theorem provides the continuity property

$$x_n \in K, \quad \|x_n - x\| \to 0, \quad \rho_n \gamma = v_t^{x_n}, \quad \rho\gamma = v_t^x$$
$$\implies \rho_n \rightharpoonup \rho \text{ weakly in } L^1(\gamma) \tag{8.2}$$

for all $t > 0$.

Proof of Theorem 1.2 We already proved statement (a) in Lemma 7.2. Let us consider the semigroup P_t induced by $\mathcal{E}_\gamma$, linked to v_t^x by (7.8); the semigroup property of P_t can be read at the level of v_t^x, and gives the Chapman–Kolmogorov equations. Therefore

these measures are the transition probabilities of a time-homogeneous Markov process $\mathbb{P}_x$ in $K^{[0,+\infty[}$. In particular, the Markov property gives the explicit formula

$$
\mathbb{P}_x(\{X_{t_1} \in A_1\} \cap \cdots \cap \{X_{t_{n-1}} \in A_{n-1}\} \cap \{X_{t_n} \in A_n\}) \qquad (8.3)
$$
$$
= \int\limits_{A_1} \cdots \int\limits_{A_{n-1}} \int\limits_{A_n} 1 \, dv^{y_{n-1}}_{t_n-t_{n-1}}(y_n) \, dv^{y_{n-2}}_{t_{n-1}-t_{n-2}}(y_{n-1}) \cdots dv^{y_0}_{t_1-t_0}(y_1)
$$

(with $y_0 = x$, $0 = t_0 \le t_1 < \cdots < t_n < +\infty$, $A_1, \ldots, A_n \in \mathscr{B}(K)$) for these finite-dimensional distributions. The continuity of $x \mapsto \mathbb{P}_x$, namely the continuity of all finite-dimensional distributions, is a direct consequence of (8.2) and (8.3).

Now, let us prove (8.5). In order to apply the general results of [27], we need to emphasize two more properties of $\mathcal{E}_\gamma$. First, $\mathcal{E}_\gamma$ is *tight*: this means that there exists a nondecreasing sequence of compact sets $F_m \subset H$ such that $\mathrm{cap}_\gamma (H \backslash F_m) \to 0$ (cap_γ being the capacity induced by $\mathcal{E}_\gamma$, see [27]). This can be proved using (7.4) and the argument in [27, Proposition IV.4.2]: let $(x_n) \subset H$ be a dense sequence and define

$$
w_n(x) := \min \left\{ 1, \min_{0 \le i \le n} \|x - x_i\| \right\}.
$$

It is immediate to check that $0 \le w_n \le 1$, $w_n \downarrow 0$ in H and $[w_n]_{\mathrm{Lip}(H)} \le 1$. Therefore (w_n) is bounded in the weak topology of $D(\mathcal{E}_\gamma)$ and converges to 0 in the weak topology of $\mathcal{E}_\gamma$. The Banach–Saks theorem ensures the existence of a subsequence (n_k) such that the Cesaro means

$$
v_k := \frac{w_{n_1} + \cdots + w_{n_k}}{k}
$$

converge to 0 strongly in $D(\mathcal{E}_\gamma)$. This implies [27, Proposition III.3.5] that a subsequence $(v_{k(l)})$ of (v_k) converges to 0 quasi-uniformly, i.e. for all integers $m \ge 1$ there exists a closed set $G_m \subset H$ such that $\mathrm{cap}_\gamma (H \backslash G_m) < 1/m$ and $v_{k(l)} \to 0$ uniformly on G_m. As $w_{n(k(l))} \le v_{k(l)}$, if we set $F_m = \cup_{i \le m} G_i$, we have that $w_{n(k(l))} \to 0$ uniformly on F_m for all m and $\mathrm{cap}_\gamma (H \backslash F_m) \le 1/m$. If $\epsilon > 0$ and n is an integer such that $w_n < \epsilon$ on F_m, the definition of w_n implies

$$
F_m \subset \bigcup_{i=1}^{n} B(x_i, \epsilon).
$$

Since ϵ is arbitrary this proves that F_m is totally bounded, hence compact. This completes the proof of the tightness of $\mathcal{E}_\gamma$.

Second, $\mathcal{E}_\gamma$ is *local*, i.e. $\mathcal{E}_\gamma(u, v) = 0$ whenever u, $v \in D(\mathcal{E}_\gamma)$ have compact and disjoint support. This can be easily achieved (see also [27, Lemma V.1.3]) taking sequences (u_n), $(v_n) \subset C^1_b(A(\gamma))$ converging to u and v respectively in the norm $\| \cdot \|_{\gamma,1}$, and modifying them, without affecting the convergence, so that u_n and v_n have disjoint supports. One concludes noticing that $\mathcal{E}_\gamma(f, g) = 0$ whenever f, $g \in C^1_b(A(\gamma))$ have disjoint supports.

These properties imply, according to [27, Theorems IV.3.5, V.1.5] the existence of a Markov family of probability measures $\{\tilde{\mathbb{P}}_x\}_{x \in K}$ on $C([0, +\infty[; K)$ (uniquely determined up to γ-negligible sets), satisfying

$$P_t f(x) = \tilde{\mathbb{E}}_x[f(X_t)] \qquad \text{for } \gamma\text{-a.e. } x \in K \tag{8.4}$$

for all bounded Borel functions f on K (here $\tilde{\mathbb{E}}_x$ is the expectation with respect to $\tilde{\mathbb{P}}_x$). Now, since H is separable we can find a countable family $\mathcal{A}$ of open sets stable under finite intersections which generates $\mathscr{B}(H)$; choosing $f = 1_A$ in (8.4) and (7.8), and taking into account that $\mathcal{A}$ is countable, we can find a γ-negligible set $N \subset K$ such that $\nu_t^x(A) = \mathbb{E}_x[1_A(X_t)]$ for all $A \in \mathcal{A}$, $t \in \mathbb{Q}$ and all $x \in K \setminus N$. As a consequence, ν_x^t is the law of X_t under $\tilde{\mathbb{P}}_x$ for all $x \in K \setminus N$ and all $t \in \mathbb{Q}$. We can now use the continuity of the process X_t and of $t \mapsto \nu_t^x$ to obtain that ν_t^x are the one-time marginals of $\tilde{\mathbb{P}}_x$ for all $x \in K \setminus N$.

We prove now path continuity under $\mathbb{P}_x$ for $x \in K \setminus N$, using the property $\nu_t^x \ll \gamma$: we adapt the approach of [16] to our setting. By the Markov property we obtain that all finite-dimensional distributions of $\tilde{\mathbb{P}}_x$ and $\mathbb{P}_x$ coincide; as a consequence, if we denote by $i : \Omega \to K^{[0,+\infty[}$ the (obviously measurable) injection map, $i_\#\tilde{\mathbb{P}}_x = \mathbb{P}_x$. By the Ulam lemma, we can find compacts sets $K_n \subset \Omega$ with $\tilde{\mathbb{P}}_x(K_n) \uparrow 1$; now, if $S \subset [0, +\infty)$ is bounded, countable and $B_S \subset K^{[0,+\infty[}$ is the measurable set defined by

$$B_S := \{\omega \in \Omega : \text{ the restriction of } \omega \text{ to } S \text{ is uniformly continuous}\},$$

from the inclusion $B_S \cap \Omega \supset K_n$ we obtain $\mathbb{P}_x(B_S) \geq \tilde{\mathbb{P}}_x(K_n)$, hence $\mathbb{P}_x(B_S) = 1$. A well known criterion [34, Lemma 2.1.2] then gives that $\mathbb{P}_x^*(\Omega) = 1$. This proves that

$$\mathbb{P}_x^*(\Omega) = 1 \qquad \text{for } \gamma\text{-a.e. } x \in K. \tag{8.5}$$

In order to show the first part of statement (c), fix $x \in K$, and define B_S as above, with $S \subset]0, +\infty[$ satisfying $\varepsilon := \inf S > 0$ and $\sup S < \infty$. Since the law of $X_\varepsilon(x)$ is absolutely continuous with respect to γ, we know from (8.5) that $\mathbb{P}_{X_\varepsilon}(B_{S-\varepsilon}) = 1$ $\mathbb{P}_x$-almost surely. Taking expectations, and using the Markov property, we get $\mathbb{P}_x(B_S) = 1$. Again the same argument in [34, Lemma 2.1.2] shows that $\mathbb{P}_x^*(C(]0, +\infty[; H)) = 1$.

Finally, we use the representation (7.8) and the fact that P_t is self adjoint (due to the fact that $\mathcal{E}_\gamma$ is symmetric) to obtain

$$\int\limits_K \int\limits_K \varphi(x)\psi(y) \, d\nu_t^y(x) \, d\gamma(y) = \int\limits_K \int\limits_K \varphi(x)\psi(y) \, d\nu_t^x(y) \, d\gamma(x)$$

$$\forall \varphi, \ \psi \in L^\infty(\gamma). \tag{8.6}$$

This means that the process $\mathbb{P}_x$ is reversible.

Proof of Theorem 1.3 It is a direct consequence of the estimates in Theorem 5.1 and of the coincidence, proved above, of the law of X_t under $\mathbb{P}_x$ with ν_t^x. $\qquad \square$

Proof of Theorem 1.5 We shall denote by $v_t^{n,x}$, $\mathbb{E}_x^n$ (resp. v_t^x, $\mathbb{E}_x$) the transition probabilities and the expectations relative to $\mathbb{P}_x^n$ (resp. $\mathbb{P}_x$). From Theorem 6.1 we obtain:

$$x_n \in K_n, \ \|x_n - x\| \to 0, \ x \in K \implies v_t^{n,x_n} \to v_t^x \text{ in } \mathscr{P}_2(H). \quad (8.7)$$

We shall prove by induction on m that $\mathbb{E}_{x_n}^n[f(X_{t_1}, \ldots, X_{t_m})] \to \mathbb{E}_x[f(X_{t_1}, \ldots, X_{t_m})]$ for all $f \in C_b(H^m)$. Obviously we can restrict ourselves to $f \in \text{Lip}_b(H^m)$ and the case $m = 1$ corresponds to (8.7). So, let us assume the statement valid for $m \geq 1$ and let us prove it for $m+1$. Let $f \in \text{Lip}_b(H^{m+1})$, and let $\Pi_n : H \to K_n$ be the canonical projection. By the weak convergence of γ_n to γ, we have $\|\Pi_n(y) - y\| \to 0$ for all $y \in K$. As a consequence, the induction assumption gives $\mathbb{E}_{\Pi_n(y)}^n[f(y, X_{t_2}, \ldots, X_{t_{m+1}})] \to \mathbb{E}_y[f(y, X_{t_2}, \ldots, X_{t_{m+1}})]$. Since f is Lipschitz we have also

$$\lim_{n\to\infty} \mathbb{E}_{\Pi_n(y)}^n[f(\Pi_n(y), X_{t_2}, \ldots, X_{t_{m+1}})] = \mathbb{E}_y[f(y, X_{t_2}, \ldots, X_{t_{m+1}})] \quad \forall y \in K. \quad (8.8)$$

Thanks to (8.8) and Lemma 7.4, we can pass to the limit as $n \to \infty$ in the identity

$$\mathbb{E}_x^n[f(X_{t_1}, \ldots, X_{t_{m+1}})] = \int_K \mathbb{E}_y^n[f(y, X_{t_2}, \ldots, X_{t_{m+1}})] \, dv_{t_1}^{n,x_n}(y)$$
$$= \int_H \mathbb{E}_{\Pi_n(y)}^n[f(\Pi_n(y), X_{t_2}, \ldots, X_{t_{m+1}})] \, dv_{t_1}^{n,x_n}(y)$$

to obtain $\mathbb{E}_x^n[f(X_{t_1}, \ldots, X_{t_{m+1}})] \to \int_K \mathbb{E}_y[f(y, X_{t_2}, \ldots, X_{t_{m+1}})] \, dv_{t_1}^x(y) = \mathbb{E}_x[f(X_{t_1}, \ldots, X_{t_{m+1}})]$. This proves statement (a). Statements (b) and (c) follow at once by the tightness Lemma 8.1 below. □

Lemma 8.1 (Tightness) *Let γ_n and γ as in Theorem 1.5, let $x \in K(\gamma)$ and let $x_n \in K(\gamma_n)$ be such that $x_n \to x$. For all $0 < \varepsilon \leq T < +\infty$, $h \in H$, the laws of $(\langle X_t, h\rangle_H, t \in [\varepsilon, T])$ under $\mathbb{P}_{x_n}^n$, $n \in \mathbb{N}$, form a tight sequence in $C([\varepsilon, T])$. Moreover the laws of $(\langle X_t, h\rangle_H, t \in [0, T])$ under $\mathbb{P}_{\gamma_n}^n$, $n \in \mathbb{N}$, form a tight sequence in $C([0, T])$.*

Proof Let $H_n = H^0(\gamma_n)$, $K_n = K(\gamma_n)$ and $\mathbb{P}_{\gamma_n}^n = \int_{H_n} \mathbb{P}_x^n \, d\gamma_n(x)$. For any $h \in H_n$ we have by the Lyons–Zheng decomposition, see e.g. [18, Theorem 5.7.1] that, under $\mathbb{P}_{\gamma_n}^n$,

$$\langle h, X_t - X_0 \rangle_{H_n} = \frac{1}{2} M_t - \frac{1}{2}(N_T - N_{T-t}), \quad \forall t \in [0, T],$$

where M, respectively N, is a $\mathbb{P}_{\gamma_n}^n$-martingale with respect to the natural filtration of $(X_t, \ t \in [0, T])$, respectively of $(X_{T-t}, \ t \in [0, T])$. Moreover the quadratic variations $\langle M \rangle_t$, $\langle N \rangle_t$ are both equal to $t \cdot \|h\|_{H_n}^2$. By the Burkholder–Davis–Gundy inequality we can find, for all $p > 1$, a constant $c_p \in (0, +\infty)$ such that

$$\mathbb{E}_{\gamma_n}^n\left[\left|\langle h, X_t - X_s\rangle_{H_n}\right|^p\right] \leq c_p \|h\|_{H_n}^p |t - s|^{p/2}, \quad t, s \in [0, T]. \quad (8.9)$$

Let us denote by $\Pi_n : H \to H_n$ the duality map satisfying $\langle \Pi_n(h), v \rangle_{H_n} = \langle h, v \rangle_H$ for all $h \in H_n$. Then, choosing $v = \Pi_n(h)$, from (1.10) we get $\|\Pi_n(h)\|_{H_n} \leq \kappa \|h\|_H$, so that (8.9) gives

$$\mathbb{E}_{\gamma_n}^n \left[\left| \langle h, X_t - X_s \rangle_H \right|^p \right] \leq \kappa^p c_p \|h\|_H^p |t - s|^{p/2}, \qquad t, s \in [0, T].$$

Then tightness of the laws of $\langle X_t, h \rangle_H$ under $\mathbb{P}_{\gamma_n}^n$ in $C([0, T])$ follows e.g. by [34, Exercise 2.4.2].

Let $\varepsilon > 0$ and let us prove that $\sup_n \mathcal{H}(\nu_\varepsilon^{n,x_n} | \gamma_n) < +\infty$. Since $x \in K$ there exist $R_\varepsilon > 0$ such that $\gamma(B_{R_\varepsilon}(x)) > 1/2$, so that there exists n_ε such that $\gamma_n(B_{R_\varepsilon}(x_n)) \geq 1/2$ for all $n \geq n_\varepsilon$. Let $\nu_n := \gamma_n(\cdot | B_{R_\varepsilon}(x_n))$; then for $n \geq n_\varepsilon$ we have $W_{2,H_n}^2(\delta_{x_n}, \nu_n) \leq \kappa^2 W_2^2(\delta_{x_n}, \nu_n) \leq 2\kappa^2 R^2$, and since $\mathcal{H}(\nu_n | \gamma_n) = -\ln \gamma_n(B_{R_\varepsilon}(x_n))$, by (iii) of Theorem 5.1 we get

$$\mathcal{H}(\nu_\varepsilon^{x_n} | \gamma_n) \leq \frac{2\kappa^2 R^2}{\varepsilon} + \ln 2 \qquad \forall n \geq n_\varepsilon.$$

Let $\mathbf{P}_n$ (resp. $\overline{\mathbf{P}}_n$) be the law of $(X_t, t \in [\varepsilon, T])$ under $\mathbb{P}_{x_n}^n$ (resp. $\mathbb{P}_{\gamma_n}^n$). Let us prove that $\mathbf{P}_n \ll \overline{\mathbf{P}}_n$ and

$$\frac{d\mathbf{P}_n}{d\overline{\mathbf{P}}_n} = \rho_\varepsilon^n(X_\varepsilon) \qquad \overline{\mathbf{P}}_n\text{-almost surely,} \tag{8.10}$$

where ρ_ε^n is the density of ν_ε^{n,x_n} with respect to γ_n. For any bounded and Borel functional $\Phi : C([0, T - \varepsilon]; H) \mapsto \mathbb{R}$, we have by the Markov property:

$$\mathbb{E}_{x_n}^n(\Phi(X_{\varepsilon+\cdot})) = \mathbb{E}_{x_n}^n(\mathbb{E}_{X_\varepsilon}^n(\Phi)) = \int d\nu_\varepsilon^{n,x_n}(y)\, \mathbb{E}_y^n(\Phi) = \int d\gamma_n(y)\, \rho_\varepsilon^n(y)\, \mathbb{E}_y^n(\Phi)$$

$$= \mathbb{E}_{\gamma_n}^n(\rho_\varepsilon^n(X_0)\, \Phi(X_\cdot)) = \mathbb{E}_{\gamma_n}^n(\rho_\varepsilon^n(X_\varepsilon)\, \Phi(X_{\varepsilon+\cdot})),$$

where in the last equality we use stationarity, and (8.10) is proven.

Let now $h \in H$ and let $\mathbf{P}_n^h$ (resp. $\overline{\mathbf{P}}_n^h$) be the law of $(\langle X_t, h \rangle_H, t \in [\varepsilon, T])$ under $\mathbb{P}_{x_n}^n$ (resp. $\mathbb{P}_{\gamma_n}^n$); since the relative Entropy does not increase under marginals [3, 9.4.5], from (8.10) we get

$$\mathcal{H}(\mathbf{P}_n^h | \overline{\mathbf{P}}_n^h) \leq \mathcal{H}(\mathbf{P}_n | \overline{\mathbf{P}}_n) = \mathcal{H}(\nu_\varepsilon^{x_n} | \gamma_n).$$

It follows that $\sup_n \mathcal{H}(\mathbf{P}_n^h | \overline{\mathbf{P}}_n^h)$ is finite. By applying the entropy inequality (4.4), tightness of $(\overline{\mathbf{P}}_n^h)$ implies tightness of $(\mathbf{P}_n^h)$. □

Proof of Theorem 1.6 Let $\mu \in \mathscr{P}_2(H)$ be an invariant measure for $(P_t)_{t \geq 0}$. Then, by (7.10), $\mathscr{S}\mu(t) \equiv \mu$ is a constant gradient flow of $\mathcal{H}(\cdot | \gamma)$ and therefore, by (3.20), $\mathcal{H}(\mu | \gamma) \leq \mathcal{H}(\nu | \gamma)$ for all $\nu \in \mathscr{P}_2(H)$. Since $t \mapsto t \ln t$ is strictly convex, the unique minimizer of $\mathcal{H}(\cdot | \gamma)$ in $\mathscr{P}_2(H)$ is γ, and therefore $\mu = \gamma$. □

Appendix A: Some properties of log-concave measures

In this appendix we state and prove some useful properties of log-concave measures and of convex functions used throughout the paper.

First of all, for lower semicontinuous convex functions $V : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ (i.e. the typical densities of log-concave measures), we recall that the properties $\int \exp(-V) \, dx < +\infty$, $V(x) \to +\infty$ as $\|x\| \to +\infty$ and $V(x) \to +\infty$ at least linearly as $\|x\| \to +\infty$ are all equivalent: indeed, the equivalence between the second and the third one simply follows by the monotonicity of difference quotients along radial directions, and clearly a linear growth at infinity implies finiteness of the integral. On the other hand, if the integral is finite, a crude growth estimate on V can be obtained as follows: assuming with no loss of generality that $\{V < +\infty\}$ has nonempty interior, we can find a ball B and $M < +\infty$ such that $V \leq M$ on B; then, on the convex cone C_x generated by x and B, we have the inequality $V \leq M + V^+(x)$. Changing signs and taking exponentials we can integrate on C_x to obtain

$$e^{V^+(x)+M} \geq \left(\int_{\mathbb{R}^k} e^{-V(y)} \, dy \right)^{-1} \mathscr{L}^k(C_x) \to +\infty \quad \text{as } |x| \to +\infty.$$

Lemma A.1 (Variational characterization of $\mathcal{E}_\gamma(u, u)$) *Let $\gamma \in \mathscr{P}_2(H)$ be a non-degenerate log-concave measure, and let u be a bounded C^1 cylindrical function, with $\inf u > 0$ and $\int u^2 \, d\gamma = 1$. Then $\sqrt{\mathcal{E}_\gamma(u, u)}$ is the smallest constant S satisfying*

$$\mathcal{H}(\mu|\gamma) \geq \mathcal{H}(u^2\gamma|\gamma) - 2\, S\, W_2(\mu, u^2\gamma), \qquad \forall \mu \in \mathscr{P}_2(H). \tag{A.1}$$

Proof First, we realize that this is essentially a finite-dimensional statement. Indeed, if $\pi : H \to L$ is a finite-dimensional orthogonal projection such that $u = u \circ \pi$, a simple application of Jensen's inequality gives [3, Lemma 9.4.5] $\mathcal{H}(\mu|\gamma) \geq \mathcal{H}(\pi_\#\mu|\pi_\#\gamma)$. Since $W_2(\pi_\#\mu, u^2\pi_\#\gamma) = W_2(\pi_\#\mu, \pi_\#(u^2\gamma)) \leq W_2(\mu, u^2\gamma)$ and $\mathcal{E}_\gamma(u, u) = \mathcal{E}_{\pi_\#\gamma}(u, u)$, we need only to check the analog of (A.1) with γ replaced by $\pi_\#\gamma$ and H replaced by L. So, from now on we shall assume that $H = \mathbb{R}^k$ for some integer k.

In Lemma 3.3 we proved that

$$\mathcal{H}(\mu|\gamma) \geq \mathcal{H}(u^2\gamma|\gamma) - 2\sqrt{\mathcal{E}_\gamma(u, u)}\, W_2(\mu, u^2\gamma), \qquad \forall \mu \in \mathscr{P}_2(\mathbb{R}^k)$$

when $\gamma = \exp(-V)\mathscr{L}^k$ with V smooth, convex, and ∇V and all its derivatives are bounded [it suffices to use the Schwartz inequality to estimate from below the scalar product in (3.18)]. By monotone approximation (see Step 2 in the proof of Theorem 7.3) the same inequality holds for all log-concave γ in $\mathbb{R}^k$.

It remains to show that $\sqrt{\mathcal{E}_\gamma(u, u)}$ is the smallest constant with this property. In order to prove this fact, we fix $s \in C_c^\infty(\mathbb{R}^k; \mathbb{R}^k)$ with support contained in the interior of $\{V < +\infty\}$, and consider the maps $t_\varepsilon := i + \varepsilon s$ and the measures $\mu_\varepsilon = (t_\varepsilon)_\#(u^2\gamma)$,

so that $W_2^2(\mu_\varepsilon, u^2\gamma) \le \varepsilon^2 \int \|s\|^2 u^2 \, d\gamma$. On the other hand, the area formula gives that the density of μ_ε with respect to $\mathscr{L}^k$ is given by f_ε, where

$$f_\varepsilon := \frac{u^2 \exp(-V)}{|\det \nabla t_\varepsilon|} \circ t_\varepsilon^{-1}$$

(notice that for ε small enough t_ε is a diffeomorphism which leaves $\{V < +\infty\}$ invariant). Since

$$\mathcal{H}(\mu_\varepsilon | \gamma) = \int_{\mathbb{R}^k} f_\varepsilon \ln f_\varepsilon \, dx + \int_{\mathbb{R}^k} f_\varepsilon V \, dx = \int_{\mathbb{R}^k} \ln(f_\varepsilon \circ t_\varepsilon) u^2 \, d\gamma + \int_{\mathbb{R}^k} (V \circ t_\varepsilon) u^2 \, d\gamma$$

$$= \mathcal{H}(u^2\gamma | \gamma) - \varepsilon \int_{\mathbb{R}^k} (\nabla \cdot s) u^2 \exp(-V) - \langle \nabla V, s \rangle u^2 \exp(-V) \, dx + o(\varepsilon)$$

$$= \mathcal{H}(u^2\gamma | \gamma) + 2\varepsilon \int_{\mathbb{R}^k} u \langle \nabla u, s \rangle \, d\gamma + o(\varepsilon)$$

from (A.1) we get

$$\int_{\mathbb{R}^k} u \langle \nabla u, s \rangle \, d\gamma \ge -S \sqrt{\int_{\mathbb{R}^k} \|s\|^2 u^2 \, d\gamma}.$$

Since $u^2\gamma$ is concentrated in the interior of $\{V < +\infty\}$, we can approximate in $L^2(u^2\gamma; \mathbb{R}^k)$ the function $-\nabla \ln u$ with s it follows that $S \ge \sqrt{\mathcal{E}_\gamma(u, u)}$. $\qquad\square$

In the next two propositions, borrowed essentially from [25–40], we show that, for convex functions $U : \mathbb{R}^k \to \mathbb{R}$, the growth at infinity of ∇U is always balanced by the factor e^{-U}; this leads to uniform bounds and tightness estimates for the measures $|\nabla U| e^{-U} \mathscr{L}^k$, under uniform lower bounds on U.

Proposition A.2 *Let $U : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ be convex and lower semicontinuous, with $U(x) \to +\infty$ as $\|x\| \to +\infty$, $\{U < +\infty\}$ having a nonempty interior, and set $\gamma = \exp(-U)\mathscr{L}^k$. Then, for all unit vectors $h \in \mathbb{R}^k$ there exists a unique finite signed measure Σ_h^U in $\mathbb{R}^k$ supported on $\overline{\{U < +\infty\}}$ such that*

$$\int_{\mathbb{R}^k} \frac{\partial u}{\partial h} \, d\gamma = \int_{\mathbb{R}^k} u \, d\Sigma_h \quad \forall u \in C_b^1(\mathbb{R}^k). \tag{A.2}$$

Moreover, we have $|\Sigma_h^U|(\mathbb{R}^k) = 2 \int_{h^\perp} \exp(-\min_{t \in \mathbb{R}} U(y + th)) \, dy$.

Proof Assume first $k = 1$; the function $t \mapsto \exp(-U(t))$ is infinitesimal at infinity, non-decreasing on a half-line $(-\infty, t_0)$ and non-increasing on $(t_0, +\infty)$, where t_0 is any point in the interior of $\{U < +\infty\}$ where U attains its minimum value.

Then $\exp(-U)$ has bounded variation on $\mathbb{R}$ and the total variation of its distributional derivative $\frac{d}{dt}\exp(-U)$ is representable by:

$$\left|\frac{d}{dt}e^{-U}\right| = 1_{(t<t_0)}\frac{d}{dt}e^{-U} - 1_{(t>t_0)}\frac{d}{dt}e^{-U}.$$

It follows that $|\frac{d}{dt}\exp(-U)|(\mathbb{R}) = 2\exp(-U(t_0)) = 2\exp(-\min_{\mathbb{R}} U)$; by definition of distributional derivative, $\Sigma^U = -\frac{d}{dt}\exp(-U)$ fulfils (A.2) when the function u is compactly supported, and a simple density argument gives the general case.

In the case $k > 1$ we denote $U_y(t) := U(y+th)$; since U has at least linear growth at infinity, it is easy to check that $\exp(-\min_{\mathbb{R}} U_y)$ is integrable on $h^\perp$. Now, notice that Fubini's theorem implies the existence of Σ_h^U and its coincidence with the measure $\int_{h^\perp}\Sigma^{U_y}\,dy$, i.e.

$$\int_{\mathbb{R}^k}\frac{\partial u}{\partial h}e^{-U}\,dy\,dt = \int_{h^\perp}\left(\int_{\mathbb{R}}\frac{d}{dt}u(y+th)e^{-U_y(t)}\,dt\right)dy, \quad u \in C_b^1(\mathbb{R}^k).$$

On the other hand, if we denote by A the projection on $h^\perp$ of the interior of the convex set $\{U < +\infty\}$, and by C the projection of $\{U < +\infty\}$, we have that $\{U_y < +\infty\}$ has nonempty interior for all $y \in A$, while U_y is identically equal to $+\infty$ for all $y \in h^\perp \setminus C$; points y in $C \setminus A$ correspond to projections of boundary points of $\{U < +\infty\}$ where h is tangential to the boundary, and the co-area formula gives that this set of points is $\mathscr{L}^{k-1}$-negligible in $h^\perp$. As a consequence, $|\Sigma^{U_y}|(\mathbb{R}) = 2\exp(-\min U_y)$ for $\mathscr{L}^{k-1}$-a.e. $y \in h^\perp$. A general result [2, Corollary 2.29] allows to commute total variation and integral, so that

$$\left|\int_{h^\perp}\Sigma^{U_y}\,dy\right|(\mathbb{R}^k) = \int_{h^\perp}|\Sigma^{U_y}|(\mathbb{R})\,dy = 2\int_{h^\perp}\exp(-\min_{t\in\mathbb{R}} U(y+th))\,dy.$$

$\square$

Proposition A.3 (Continuity and tightness) *Let $V_n : \mathbb{R}^k \to \mathbb{R}\cup\{+\infty\}$ be convex and lower semicontinuous function, with $V_n \uparrow V$ and $\int\exp(-V_1) < +\infty$. Then for all unit vectors $h \in \mathbb{R}^k$ there exist compact sets $J_m \subset \mathbb{R}^k$ such that:*

$$|\Sigma_h^{V_n}|(\mathbb{R}^k \setminus J_m) \le \frac{1}{m} \quad \forall\, n,\, m \ge 1. \tag{A.3}$$

Furthermore, $\Sigma_h^{V_n} \to \Sigma_h^V$ in the duality with $C_b(\mathbb{R}^k)$.

Proof Let $A \in \mathbb{R}$, $B > 0$ be such that $V_1(x) \ge A + B\|x\|$ for all $x \in \mathbb{R}^k$. We set $\Sigma^n := \Sigma_h^{V_n}$, $\Sigma := \Sigma_h^V$. We first notice that $V_n(y+th) \ge V_1(y+th) \ge A + B\|y\|$

for all $y \in h^\perp$. Therefore, taking into account the representation of $|\Sigma_h^U|(\mathbb{R}^k)$ given by the previous proposition, we obtain that $|\Sigma^n|(\mathbb{R}^k)$ is uniformly bounded. On the other hand, since $\exp(-V_n)\mathscr{L}^k$ weakly converge to $\exp(-V)\mathscr{L}^k$ (by the dominated convergence theorem) from (A.2) we infer that $\Sigma^n \to \Sigma$ weakly in the duality with $C_c^1(\mathbb{R}^k)$, and then in the duality with $C_c(\mathbb{R}^k)$.

We will prove that

$$\lim_{n\to\infty} |\Sigma^n|(\mathbb{R}^k) = |\Sigma|(\mathbb{R}^k). \tag{A.4}$$

Before proving (A.4), we show that it implies (A.3): consider a dense sequence (x_j) in $\mathbb{R}^k$ and set, for $p, l \geq 1$ integers, $A_l^p := \cup_{j=1}^l B(x_j, 1/p)$. It is enough to prove that for all p there exists $l = l(p)$ such $|\Sigma^n|(\mathbb{R}^k \setminus A_l^p) \leq 2^{-p}/m$ for all n: indeed, in this case $J_m := \cap_p \overline{A}_{l(p)}^p$ is a compact set such that $|\Sigma^n|(\mathbb{R}^k \setminus J_m) \leq 1/m$ for all $n \geq 1$. If, for some p, we can not find such l, then for all l there exists $n(l)$ such that $|\Sigma^{n(l)}|(A_l^p) \leq |\Sigma_h^{n(l)}|(\mathbb{R}^k) - 2^{-p}/m$. Since $n(l)$ must tend to $+\infty$ as $l \to\to \infty$, and any open ball $B_r(0)$ is contained in A_l^p for l large enough, by the lower semicontinuity of the total variation on open sets (see for instance [2, Proposition 1.62(b)]) we find:

$$|\Sigma|(B_r(0)) \leq \liminf_{l\to\infty} |\Sigma^{n(l)}|(B_r(0)) \leq \liminf_{l\to\infty} |\Sigma^{n(l)}|(A_l^p) \leq |\Sigma|(\mathbb{R}^k) - 2^{-p}/m.$$

Letting $r \uparrow \infty$ we obtain a contradiction. Therefore (A.3) is proven.

In order to prove (A.4), taking again into account the representation of $|\Sigma_h^U|(\mathbb{R}^k)$ given by the previous proposition and the dominated convergence theorem, it suffices to show that, with $y \in h^\perp$ fixed, $\min_{t\in\mathbb{R}} V_n(y+th)$ converges as $n \to \infty$ to $\min_{t\in\mathbb{R}} V(y+th)$. By monotonicity we need only to show that

$$\liminf_n \min_{t\in\mathbb{R}} V_n(y + th) \geq \min_{t\in\mathbb{R}} V(y + th). \tag{A.5}$$

Let $n(k)$ be a subsequence along which the liminf is achieved, let t_k be minimizers of $t \mapsto V_{n(k)}(y + th)$, and assume (possibly extracting one more subsequence) that $t_k \to t$. The lower semicontinuity of $V_{n(p)}$ gives $\lim_k V_{n(k)}(y + t_k h) \geq \liminf_k V_{n(p)}(y + t_k h) \geq V_{n(p)}(y + th)$. Letting $p \to \infty$ we obtain (A.5). Finally, the tightness estimate allows to pass from convergence of Σ^n in the duality with $C_c(\mathbb{R}^k)$ to the convergence in the duality with $C_b(\mathbb{R}^k)$. $\qquad\square$

References

1. Albeverio, S., Kusuoka, S.: Maximality of infinite-dimensional Dirichlet forms and Høegh–Krohn's model of quantum fields. Ideas and methods in quantum and statistical physics (Oslo, 1988), pp. 301–330. Cambridge University Press, Cambridge (1992)
2. Ambrosio, L., Fusco, N., Pallara, D.: Functions of Bounded Variation and Free Discontinuity Problems. Oxford University Press, Oxford (2000)
3. Ambrosio, L., Gigli, N., Savaré, G.: Gradient flows in metric spaces and in the spaces of probability measures. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2005)

4. Ambrosio, L., Savaré, G.: Gradient flows in spaces of probability measures. Handbook of Differential Equations. Evolutionary equations III. North Holland, Amsterdam (2007)
5. Bénilan, P.: Solutions intégrales d'équations d'évolution dans un espace de Banach. C. R. Acad. Sci. Paris Sér. A-B, **274** (1972)
6. Bogachev, V.I.: Gaussian measures. Mathematical Surveys and Monographs, vol. 62. AMS (1998)
7. Bogachev, V.I.: Measure Theory, vol. I, II. Springer, Heidelberg (2007)
8. Borell, C.: Convex measures on locally convex spaces. Ark. Mat. **12**, 239–252 (1974)
9. Borell, C.: Convex set functions in d-space. Period. Math. Hungar. **6**, 111–136 (1975)
10. Brézis, H.: Opérateurs Maximaux Monotones. North-Holland, Amsterdam (1973)
11. Carrillo, J.A., McCann, R., Villani, C.: Contraction in the 2-Wasserstein length space and thermalization of granular media. Arch. Ration. Mech. Anal. **179**, 217–263 (2006)
12. Cépa, E.: Problème de Skorohod multivoque. Ann. Probab. **26**(2), 500–532 (1998)
13. Debussche, A., Zambotti, L.: Conservative stochastic Cahn–Hilliard equation with reflection. Ann. Probab. **35**(5), 1706–1739 (2007)
14. Da Prato, G., Röckner, M.: Singular dissipative stochastic equations in Hilbert spaces. Probab. Theory Relat. Fields **124**(2), 261–303 (2002)
15. Da Prato, G., Zabczyk, J.: *Ergodicity for Infinite Dimensional Systems*, London Mathematical Society Lecture Notes, no. 229. Cambridge University Press, Cambridge (1996)
16. Dohmann, J.M.N.: Feller-type properties and path regularities of Markov processes. Forum Math. **17**, 343–359 (2005)
17. Feyel, D., Ustünel, A.S.: Monge–Kantorovitch measure transportation and Monge–Ampère equation on Wiener space. Probab. Theory Relat. Fields **128**, 347–385 (2004)
18. Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet Forms and Symmetric Markov Processes. Walter de Gruyter, Berlin-New York (1994)
19. Funaki, T.: Stochastic interface models. In: Picard, J. (ed.) Lectures on Probability Theory and Statistics, Ecole d'Eté de Probabilités de Saint-Flour XXXIII-2003, pp. 103–274, Lecture Notes in Mathematics, vol. 1869. Springer, Heidelberg (2005)
20. Funaki, T., Olla, S.: Fluctuations for $\nabla\phi$ interface model on a wall. Stoch. Proc. Appl. **94**, 1–27 (2001)
21. Funaki, T., Spohn, H.: Motion by mean curvature from the Ginzburg-Landau $\nabla\phi$ interface model. Commun. Math. Phys. **185**, 1–36 (1997)
22. Giacomin, G., Olla, S., Spohn, H.: Equilibrium fluctuations for $\nabla\phi$ interface model. Ann. Probab. **29**, 1138–1172 (2001)
23. John, F.: Partial Differential Equations, 4th edn. Springer, Heidelberg (1970)
24. Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker–Planck equation. SIAM J. Math. Anal. **29**, 1–17 (1998)
25. Krugova, E.P.: Differentiability of convex measures. Math. Notes **58**(5–6), 1294–1301 (1995)
26. Lions, P.-L., Sznitman, A.-S.: Stochastic differential equations with reflecting boundary conditions. Commun. Pure Appl. Math. **37**, 511–537 (1984)
27. Ma, Z.M., Röckner, M.: Introduction to the Theory of (Non-Symmetric) Dirichlet Forms. Universitext. Springer, Heidelberg (1992)
28. Nualart, D., Pardoux, E.: White noise driven quasilinear SPDEs with reflection. Prob. Theory Relat. Fields **93**, 77–89 (1992)
29. McCann, R.J.: A convexity principle for interacting gases. Adv. Math. **128**, 153–179 (1997)
30. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion. Springer, Heidelberg (1991)
31. Sheffield, S.: Random Surfaces, Asterisque, No. 304 (2005)
32. Skorohod, A.V.: Stochastic equations for diffusions in a bounded region. Theory Probab. Appl. **6**, 264–274 (1961)
33. Spohn, H.: Interface motion in models with stochastic dynamics. J. Stat. Phys. **71**, 1081–1132 (1993)
34. Stroock, D.W., Varadhan, S.R.S.: Multidimensional Diffusion Processes, 2nd edn. Springer, Heidelberg (1997)
35. Tanaka, H.: Stochastic differential equations with reflecting boundary condition in convex regions. Hiroshima Math. J. **9**, 163–177 (1979)
36. Villani, C.: Topics in optimal transportation. Graduate Studies in Mathematics, vol. 58. AMS (2003)
37. Zambotti, L.: Integration by parts formulae on convex sets of paths and applications to SPDEs with reflection. Probab. Theory Relat. Fields **123**(4), 579–600 (2002)
38. Zambotti, L.: Integration by parts on δ-Bessel Bridges, $\delta > 3$, and related SPDEs. Ann. Probab. **31**(1), 323–348 (2003)

39. Zambotti, L.: Fluctuations for a $\nabla\varphi$ interface model with repulsion from a wall. Probab. Theory Relat. Fields **129**(3), 315–339 (2004)
40. Zambotti, L.: Convergence of approximations of monotone gradient systems. J. Evol. Equ. **6**(4), 601–619 (2006)
41. Zambotti, L.: Fluctuations for a conservative interface model on a wall. ALEA **4**, 167–184 (2008)
42. Zambotti, L.: A conservative evolution of the Brownian excursion. Electron. J. Probab. **13**, 1096–1119 (2008)