Analysis of a gradient flow scheme

April 12, 2011

We consider here the case study that we saw in class, namely the gradient flow of the functional

$$F(\rho) = \int_\Omega \rho \ln \rho + \int_\Omega V d\rho,$$

where V is a Lipschitz function on the compact domain Ω, starting from a given measure $\rho_0 \in \mathcal{P}(\Omega)$ such that $F(\rho_0) < +\infty$.

We stress from the beginning that the first term of the functional is defined as

$$J(\rho) := \begin{cases} \int_\Omega \rho(x) \ln \rho(x) dx & \text{if } \rho << \mathcal{L}^d, \\ +\infty & \text{otherwise,} \end{cases}$$

where we identify the measure ρ with its density, when it is absolutely continuous.

We will use the following property.

Lemma 0.1. *If a sequence ρ_n satifsfies $J(\rho_n) \le C$ and ρ_n weakly converges to ρ as measures, then $\rho << \mathcal{L}^d$ and ρ_n also weakly converges to ρ in L^1 (in the duality with L^∞).*

Moreover, J is l.s.c. on $\mathcal{P}(\Omega)$ w.r.t. the weak convergence of probabilities.

This lemma is an exercise based on the concept of equi-integrability (for the first part) and on the equality $x \ln x = \sup_y xy - e^{y-1}$. We will not prove it.

The computation of the variational derivative $\delta F/\delta\rho$ gives the the PDE that should be satisfied by a gradient flow of F, i.e.

$$\partial_t \rho - \nabla \cdot \left(\rho \nabla \frac{\delta F}{\delta \rho}\right) = 0 \;\Rightarrow\; \partial_t \rho - \Delta \rho - \nabla \cdot (\rho \nabla V) = 0.$$

The previous lemma allows to establish the following

Proposition 0.2. *The functional F has a minimum over $\mathcal{P}(\Omega)$. In particular F is bounded from below. Moreover, for each $\tau > 0$ the following sequences of optimization problems recursively defined is well-posed*

$$\rho^\tau(k+1) \in \operatorname{argmin}_\rho F(\rho) + \frac{W_2^2(\rho, \rho^\tau(k))}{2\tau}, \tag{0.1}$$

which means that there is a minimizer at every step.

Proof. Just apply the direct method, noticing that $\mathcal{P}(\Omega)$ is compact for the weak convergence, which is the same as the convergence for the W_2 distance. And for this convergence J is l.s.c. and the other terms are continuous. □

Optimality conditions at each time step We first need the following lemma.

Lemma 0.3. *Let ν be a fixed probability measure; consider $\mu \in \mathcal{P}(\Omega)$ such that in the optimal transport from μ to ν for the cost $\frac{1}{2}|x-y|^2$ the Kantorovitch potential is unique (up to additive constants). This means that there is only one pair $(\bar\phi, \bar\psi)$ solving*

$$\max \int \phi d\mu + \psi d\nu \, : \, \phi(x_0) = 0, \, \phi(x) + \psi(y) \le \frac{1}{2}|x-y|^2,$$

where x_0 is any fixed point in Ω, so as to get rid of additive constants.

Then, for ay $\mu_1 \in \mathcal{P}(\Omega)$, setting $\mu_\varepsilon := (1-\varepsilon)\mu + \varepsilon\mu_1$, we have

$$\frac{d}{d\varepsilon}\left(\frac{1}{2}W_2^2(\mu_\varepsilon, \nu)\right)_{|\varepsilon=0} = \int \bar\phi d(\mu_1 - \mu).$$

Proof. Let us try to estimate the ratio $(\frac{1}{2}W_2^2(\mu_v e, \nu) - \frac{1}{2}W_2^2(\mu, \nu))/\varepsilon$. First, by using that $(\bar\phi, \bar\psi)$ is optimal for μ but not for μ_ε, we have

$$\frac{\frac{1}{2}W_2^2(\mu_\varepsilon, \nu) - \frac{1}{2}W_2^2(\mu, \nu)}{\varepsilon} \ge \int \bar\phi d\mu_\varepsilon + \int \bar\psi d\nu - \int \bar\phi d\mu - \int \bar\psi d\nu \varepsilon = \int \bar\phi d(\mu_1 - \mu),$$

which gives a lower bound. This means that $\liminf_{\varepsilon\to 0}(\frac{1}{2}W_2^2(\mu_v e, \nu) - \frac{1}{2}W_2^2(\mu, \nu))/\varepsilon \ge \int \bar\phi d(\mu_1 - \mu)$.

To look at the lim sup, first fix a sequence of values of ε_k such that $\lim_k(\frac{1}{2}W_2^2(\mu_{ve_k}, \nu) - \frac{1}{2}W_2^2(\mu, \nu))/\varepsilon_k = \limsup_{\varepsilon\to 0}(\frac{1}{2}W_2^2(\mu_v e, \nu) - \frac{1}{2}W_2^2(\mu, \nu))/\varepsilon$. Then we can estimate the same ratio using the optimality of a pair (ϕ_k, ψ_k), Kantorovitch potentials in the transport from μ_{ε_k} to ν.

$$\frac{\frac{1}{2}W_2^2(\mu_{\varepsilon_k}, \nu) - \frac{1}{2}W_2^2(\mu, \nu)}{\varepsilon_k} \le \int \phi_k d\mu_{\varepsilon_k} + \int \psi_k d\nu - \int \phi_k d\mu - \int \psi_k d\nu \varepsilon_k = \int \phi_k d(\mu_1 - \mu). \quad (0.2)$$

The problem in this estimate is that we need to pass to the limit in k. To do this, first notice that the families of functions $(\phi_k)_k$ and $(\psi_k)_k$ are both equicontinuous, since the expressions

$$\phi_k(x) = \inf_y \frac{1}{2}|x-y|^2 - \psi_k(y) \text{ and } \psi_k(y) = \inf_x \frac{1}{2}|x-y|^2 - \phi_k(x) \quad (0.3)$$

allow to give Lipschitz bounds (both ϕ_k and ψ_k are hence Lipschitz continuous with constant $\mathrm{diam}(\Omega)$). Moreover, the condition $\phi_k(x_0) = 0$ gives a bound on ϕ_k and (0.3) turns it into a bound on ψ_k as well. Hence, Ascoli's Theorem allows to pass to the limit up to a subsequence (not relabeled). This gives $(\phi_k, \psi_k) \to (\phi, \psi)$ uniformly, and it is easy to check that (ϕ, ψ) must be optimal in the duality formula for the transport between μ and ν. Actually, from

$$\frac{1}{2}W_2^2(\mu_{\varepsilon_k}, \nu) = \int \phi_k d\mu_{\varepsilon_k} + \int \psi_k d\nu$$

it is easy to pas t the limit and get

$$\frac{1}{2}W_2^2(\mu, \nu) = \int \phi d\mu + \int \psi d\nu,$$

which implies $\phi = \bar\phi$ and $\psi = \bar\psi$ by uniqueness. Finally, passing to the limit in (0.2) we get also $\limsup_{\varepsilon\to 0}(W_2^2(\mu_\varepsilon, \nu) - W_2^2(\mu, \nu))/\varepsilon \le \int \bar\phi d(\mu_1 - \mu)$. □

The previous result is useful when one can guarantee uniqueness of the Kantorovitch potential, which is the case in the following proposition.

Proposition 0.4. *If at least one of the measures μ or ν is equivalent to the Lebesgue measure (i.e. it is absolutely continuous and has non-zero density), then the Kantorovitch potential in the transport from μ to ν for the cost $\frac{1}{2}|x-y|^2$ is unique up to additive constants.*

Proof. Suppose for simplicity that μ is a.c. with positive density. It is sufficient to notice that the optimal transport map T sending μ onto ν is unique and that we have $T(x) = x - \nabla\phi(x)$ $\mu-$a.e. (this comes from the general formula $T(x) = x - \nabla h^*(\nabla\phi(x))$ but in this case $h(z) = \frac{1}{2}|z|^2$ and ∇h^* is the identity). Yet, in this case $\mu-$a.e. means Lebesgue-a.e. and this allows to know $\nabla\phi$ a.e. ϕ being Lipschitz, on a connected domain, this implies uniqueness up to additive constants for ϕ. □

Let us come back to Problem (0.1).

Lemma 0.5. *Any minimizer $\bar\rho$ in* (0.1) *must satisfy $\bar\rho > 0$ a.e.*

Proof. Consider the measure $\tilde\rho$ with constant positive density c in Ω (i.e. the density equals $|\Omega|^{-1}$). Let us define ρ_ε as $(1-\varepsilon)\bar\rho + \varepsilon\tilde\rho$ an compare $\bar\rho$ to ρ_ε.

We write

$$J(\bar\rho) - J(\rho_\varepsilon) \leq \int V d\rho_\varepsilon - \int V d\bar\rho + \frac{W_2^2(\rho_\varepsilon, \rho_k)}{2\tau} - \frac{W_2^2(\bar\rho, \rho(k))}{2\tau}.$$

The Wasserstein term in the right hand side may be estimated by convexity, which gives

$$\frac{W_2^2(\rho_\varepsilon, \rho_k)}{2\tau} \leq (1-\varepsilon)\frac{W_2^2(\bar\rho, \rho_k)}{2\tau} + \varepsilon\frac{W_2^2(\bar\rho, \rho_k)}{2\tau}.$$

This shows that the right hand side is estimated by $C\varepsilon$ and we get

$$\int f(\bar\rho) - f(\rho_\varepsilon) \leq C\varepsilon$$

where we use $f(t) = t\ln t$. Since f is convex we write, on the set $A = \{x \in \Omega : \bar\rho(x) > 0\}$, $f(\bar\rho(x)) - f(\rho_\varepsilon(x)) \geq (\bar\rho(x) - \rho_\varepsilon(x))f'(\rho_\varepsilon(x)) = \varepsilon(\bar\rho(x) - \tilde\rho(x))(1 + \ln\rho_\varepsilon(x))$. On the set $B = \{x \in \Omega : \bar\rho(x) = 0\}$ we simply write $f(\bar\rho(x)) - f(\rho_\varepsilon(x)) = -\varepsilon c\ln(\varepsilon c)$. This allows to write

$$-\varepsilon c\ln(\varepsilon c)|B| + \varepsilon\int(\bar\rho(x) - c)(1 + \ln\rho_\varepsilon(x))dx \leq C\varepsilon.$$

Dividing by ε and letting $\varepsilon \to 0$ provides a contradiction, unless $|B| = 0$. □

We can now compute the first variation and give optimality conditions on the optimal $\rho^\tau(k+1)$.

Proposition 0.6. *The optimal measure $\rho^\tau(k+1)$ in* (0.1) *satisfies*

$$\ln(\rho^\tau(k+1)) + V + \frac{\bar\phi}{\tau} = \textit{constant a.e.}$$

where $\bar{\phi}$ is the (unique) Kantorovitch potential from $\rho^\tau(k+1)$ to $\rho^\tau(k)$. In particular, it T_k^τ is the optimal transport from $\rho^\tau(k+1)$ to $\rho^\tau(k)$, then it satisfies

$$v^\tau(k) := \frac{id - T_k^\tau}{\tau} = -\nabla\left(\ln(\rho(k+1)) + V\right) \;\; a.e. \tag{0.4}$$

Proof. Take the optimal measure $\bar{\rho} := \rho^\tau(k+1)$ and compute variations with respect to perturbations of the form $\rho_\varepsilon := (1-\varepsilon)\bar{\rho} + \varepsilon\tilde{\rho}$, where $\tilde{\rho}$ is any other probability measure, with L^∞ density (so as to ensure every integrability condition). This means choosing a perturbation $\chi = \tilde{\rho} - \bar{\rho}$, which guarantees that, for $\varepsilon > 0$, the measure ρ_ε is actually a probability over Ω.

We now compute the first variation and, due to optimality, we have

$$0 \le \frac{d}{d\varepsilon}\left(F(\bar{\rho} + \varepsilon\chi) + \frac{1}{\tau}\frac{W_2^2(\bar{\rho} + \varepsilon\chi, \rho^\tau(k))}{2}\right)_{|\varepsilon=0} = \int \left(\frac{\delta F}{\delta\rho}(\bar{\rho}) + \frac{\bar{\phi}}{\tau}\right) d\chi.$$

If we set for a while $\psi = \frac{\delta F}{\delta\rho}(\bar{\rho}) + \frac{\bar{\phi}}{\tau}$ we would have

$$\int \psi \, d\chi \ge 0 \quad \text{i.e.} \int \psi \, d\tilde{\rho} \ge \int \psi \, d\bar{\rho} \quad \text{for all } \tilde{\rho} \in L^\infty(\Omega). \tag{0.5}$$

Set $l = ess \inf \psi$: on the one hand, the right hand side in (0.5) is larger than l, on the other hand, choosing $\tilde{\rho}$ concentrated on a set $\{\psi < l + \varepsilon\}$ (which has positive measure), we can get the left hand side smaller than $l + \varepsilon$. Hence, we finally get

$$l = \int \psi \, d\bar{\rho} \;\text{ and }\; \psi \ge l \quad \bar{\rho} - a.e.$$

This gives $\psi = l$a.e. w.r.t. $\bar{\rho}$, but since we know $\bar{\rho} > 0$ a.e., this gives that ψ is constant a.e.

This gives the first part of the thesis if we replace $\frac{\delta F}{\delta\rho}$ with $f'(\rho) + V = \ln(\rho) + 1 + V$. In particular it implies that $\rho(k+1)$ is Lipschitz continuous, since it holds

$$\rho^\tau(k+1)(x) = \exp\left(C - V(x) - \frac{\bar{\phi}(x)}{\tau}\right).$$

Then, one differentiates and gets the equality

$$\nabla\bar{\phi}(x) = \frac{id - T_k^\tau}{\tau} = -\nabla\left(\ln(\rho^\tau(k+1)) + V\right) \text{ a.e.}$$

and this allows to conclude. □

Interpolation between time steps and uniform estimates. Let us collect some other tools

Proposition 0.7. *For any $\tau > 0$, the sequence of minimizers satisfies*

$$\sum_k \frac{W_2^2(\rho^\tau(k+1), \rho^\tau(k))}{\tau} \le C = 2(F(\rho_0) - \inf F),$$

where the constant C is finite and independent of ρ.

Proof. This is obtained by comparing the optimizer $\rho^\tau(k+1)$ to the previous measure ρ^τ_k. We get

$$F(\rho^\tau(k+1)) + \frac{W_2^2(\rho^\tau(k+1),\rho^\tau(k))}{2\tau} \le F(\rho^\tau_k),$$

which implies

$$\sum_k \frac{W_2^2(\rho^\tau(k+1),\rho^\tau(k))}{\tau} \le \sum_k 2(F(\rho^\tau_k) - F(\rho^\tau(k+1))),$$

and this last sum is telescopic and gives the thesis. □

Let us define two interpolations between the measures ρ^τ_k.

With this time-discretized method, we have obtained, for each $\tau > 0$, a sequence $(\rho^\tau(k))_k$. We can use it to build at least two interesting curves in the space of measures:

- first we can define some piecewise constant curves, i.e. $\rho^\tau_t := \rho^\tau(k+1)$ for $t \in]k\tau,(k+1)\tau]$; associated to this curve we also define the velocities $v^\tau_t = v^\tau(k+1)$ for $t \in]k\tau,(k+1)\tau]$, where $v^\tau(k)$ is defined as in (0.4): $v^\tau(k) = (id - T^\tau_k)/\tau$, taking as T^τ_k the optimal transport from $\rho^\tau(k+1)$ to $\rho^\tau(k)$; we also define the momentum variable $E^\tau = \rho^\tau v^\tau$;

- then, we can also consider the densities $\tilde\rho^\tau_t$ that interpolate the discrete values $(\rho^\tau(k))_k$ along geodesics:
$$\tilde\rho^\tau_t = \left(\frac{k\tau - t}{\tau} v^\tau(k) + id\right)_\# \rho^\tau(k), \quad \text{for } t \in](k-1)\tau, k\tau[; \tag{0.6}$$
the velocities $\tilde v^\tau_t$ are defined so that $(\tilde\rho^\tau, \tilde v^\tau)$ satisfy the continuity equation, taking
$$\tilde v^\tau_t = v^\tau_t \circ ((k\tau - t)v^\tau(k) + id)^{-1};$$
as before, we define: $\tilde E_\tau = \tilde\rho^\tau \tilde v^\tau$.

After these definitions we consider some a priori bounds on the curves and the velocities that we defined. We start from some estimates which are standard in the framework of Minimizing Movements (this is the name of the discrete procedure which minimizes the funtional plus a quadratic penalization on the distance, see [1, 4]).

Notice that the velocity (i.e., metric derivative) of $\tilde\rho^\tau$ is constant on each interval $]k\tau,(k+1)\tau[$ and equal to $W_2(\rho^\tau(k+1),\rho^\tau(k))/\tau$. This distance also equals $(\int |id - T^\tau_k|^2 d\rho^\tau(k+1))^{1/2} = \tau||v^\tau_{k+1}||_{L^2(\rho^\tau(k+1))}$, which gives

$$||\tilde v^\tau_t||_{L^2(\tilde\rho^\tau_t)} = |(\tilde\rho^\tau)'|(t) = \frac{W_2(\rho^\tau(k+1),\rho^\tau(k))}{\tau} = ||v^\tau_t||_{L^2(\rho^\tau_t)},$$

where we used the fact that the velocity field $\tilde v^\tau$ has been choses so that its L^2 norm equals the metric derivative of the curve $\tilde\rho^\tau$.

In particular we can obtain

$$\begin{aligned}|E^\tau|([0,T]\times\Omega) &= \int_0^T dt\int_\Omega |v_t^\tau|d\rho_t^\tau = \int_0^T ||v_t^\tau||_{L^1(\rho_t^\tau)} \le \int_0^T ||v_t^\tau||_{L^2(\rho_t^\tau)}\\ &\le T^{1/2}\int_0^T ||v_t^\tau||^2_{L^2(\rho_t^\tau)} = T^{1/2}\sum_k \tau\left(\frac{W_2(\rho^\tau(k+1),\rho^\tau(k))}{\tau}\right)^2 \le C.\end{aligned}$$

The estimate on $\tilde{E}^\tau$ is completely analogous

$$|\tilde{E}^\tau|([0,T]\times\Omega) = \int_0^T dt\int_\Omega |\tilde{v}_t^\tau|d\tilde{\rho}_t^\tau \le T^{1/2}\int_0^T ||\tilde{v}_t^\tau||^2_{L^2(\tilde{\rho}_t^\tau)} = T^{1/2}\sum_k \tau\left(\frac{W_2(\rho^\tau(k+1),\rho^\tau(k))}{\tau}\right)^2 \le C.$$

This gives compactness of E^τ and $\tilde{E}^\tau$ in the space of vector measures on space-time, or the weak convergence. As far as $\tilde{\rho}^\tau$ is concerned, we can obtain more than that. Consider the following estimate, for $t < t$

$$W_2(\tilde{\rho}_t^\tau,\tilde{\rho}_s^\tau) \le \int_s^t |(\tilde{\rho}^\tau)'|(r)dr \le (t-s)^{1/2}\left(\int_s^t |(\tilde{\rho}^\tau)'|(r)^2 dr\right)^{1/2}.$$

From the previous computations, we have again

$$\int_0^T |(\tilde{\rho}^\tau)'|(r)^2 dr = \sum_k \tau\left(\frac{W_2(\rho^\tau(k+1),\rho^\tau(k))}{\tau}\right)^2 \le C,$$

and this implies

$$W_2(\tilde{\rho}_t^\tau,\tilde{\rho}_s^\tau) \le C(t-s)^{1/2}, \tag{0.7}$$

which means that the curves $\tilde{\rho}^\tau$ are uniformly Hölder continuous. Since they are defined on $[0,T]$ and valued in $\mathcal{P}(\Omega)$ which is compact, when endowed with the Wasserstein distance, we can apply Ascoli's Theorem. This means that, up to subsequences, we have

$$E^\tau \rightharpoonup E \text{ in } \mathcal{M}([0,T]\times\Omega;\mathbb{R}^d),\ \tilde{E}^\tau \rightharpoonup \tilde{E} \text{ in } \mathcal{M}([0,T]\times\Omega;\mathbb{R}^d);\ \tilde{\rho}^\tau \to \rho \text{ uniformly for the } W_2 \text{ distance.}$$

As far as the curves ρ^τ are concerned, they also converge uniformly to the same curve ρ, since $W_2(\rho_t^\tau,\tilde{\rho}_t^\tau) \le C\sqrt{\tau}$ (a consequence of (0.7), of the fact that $\tilde{\rho}^\tau = \rho^\tau$ on the points of the form $k\tau$ and of the fact that ρ^τ is constant on each interval $]k\tau,(k+1)\tau]$.

Let us now prove that $\tilde{E} = E$.

Lemma 0.8. *Suppose that we have two families of vector measures E^τ and $\tilde{E}^\tau$ such that*

- $\tilde{E}^\tau = \tilde{\rho}^\tau\tilde{v}^\tau$; $E^\tau = \rho^\tau v\tau$;
- $\tilde{v}_t^\tau = v_t^\tau \circ ((k\tau - t)v^\tau(k) + id)^{-1}$; $\tilde{\rho}^\tau = ((k\tau - t)v^\tau(k) + id)_{\#}\,\rho^\tau$;
- $\int\int |v^\tau|^2 d\rho^\tau \le C$ *(with C independent of τ);*

- *$E^\tau \rightharpoonup E$ and $\tilde{E}^\tau \rightharpoonup \tilde{E}$ as $\tau \to 0$*

Then $\tilde{E} = E$.

Proof. It is sufficient to fix a Lipschitz function $f : [0,T] \times \Omega \to \mathbb{R}^d$ and to prove $\int f \cdot dE = \int f \cdot d\tilde{E}$. To do that, we write

$$\int f \cdot d\tilde{E}^\tau = \int_0^T dt \int_\Omega f \cdot \tilde{v}_t^\tau d\tilde{\rho}^\tau = \int_0^T dt \int_\Omega f \circ ((k\tau - t)v^\tau + id) \cdot v_t^\tau d\rho^\tau,$$

which implies

$$\left| \int f \cdot d\tilde{E}^\tau - \int f \cdot dE^\tau \right| \leq \int_0^T dt \int_\Omega |f \circ ((k\tau - t)v^\tau + id) - f| \, |v_t^\tau| d\rho^\tau \leq \mathrm{Lip}(f)\tau \int_0^T \int_\Omega |v_t^\tau|^2 d\rho^\tau \leq C\tau.$$

This estimate proves that the limit of $\int f \cdot d\tilde{E}^\tau$ and $\int f \cdot dE^\tau$ is the same, i.e. $E = \tilde{E}$. □

Relation between ρ and E. We can obtain the following

Proposition 0.9. *The pair (ρ, E) satisfies, in distributional sense*

$$\partial_t \rho + \nabla \cdot E = 0, \quad E = -\nabla \rho - \rho \nabla V.$$

In particular we have found a solution to

$$\begin{cases} \partial_t \rho + \Delta \rho + \nabla \cdot (\rho \nabla V), \\ \rho(0) = \rho_0 & \textit{given.} \end{cases}$$

Proof. First, consider the weak convergence $(\tilde{\rho}^\tau, \tilde{E}^\tau) \rightharpoonup (\rho, E)$ (which is a consequence of $\tilde{E} = E$). Weak convergences let easily any linear condition pass to the limit and the continuity equation $\partial_t \tilde{\rho}^\tau + \nabla \cdot \tilde{E}^\tau = 0$ satisfied in the sense of distributions stays true at the limit (it is enough to test the equations against any C^1 function on $[0,T] \times \Omega$).

Then, use the convergence $(\rho^\tau, E^\tau) \rightharpoonup (\rho, E)$. Actually, using the optimality conditions of Proposition 0.6 and the defintiion of $E^\tau = v^\tau \rho^\tau$, we have, for each $\tau > 0$, $E^\tau = -\nabla \rho^\tau - \rho^\tau \nabla V$. It is not difficult to pass this condition to the limit neither. Take $f \in C^1([0,T] \times \Omega; \mathbb{R}^d)$ and test:

$$\int f \cdot dE^\tau = -\int f \cdot \nabla \rho^\tau - \int f \cdot \nabla V \rho^\tau = \int \nabla \cdot f d\rho^\tau - \int f \cdot \nabla V \rho^\tau.$$

These terms pass to the limit as $\rho^\tau \rightharpoonup \rho$, at least if $V \in C^1$, since all the test functions above are continuous. This would give $\int f \cdot dE = \int \nabla \cdot f d\rho - \int f \cdot \nabla V \rho$, which implies $E = -\nabla \rho - \rho \nabla V$.

To handle the case where V is only Lipschitz continuous, let us notice that for every τ, t we have $F(\rho_t^\tau) \leq F(\rho_0)$. This gives a uniform bound on $J(\rho_t^\tau)$ and Lemma 0.1 turns the weak convergence $\rho_t^\tau \rightharpoonup \rho_t$ as measures into a weak convergence in L^1. Once we have weak convergence in L^1, multiplying times a fixed L^∞ function, i.e. ∇V, preserves the limit. □

Notice that we were quite sloppy about the boundary conditions for the PDE that we got, which are actually Neumann (a consequence of the fact that we can test against any C^1 function, with no need to vanish on the boundary).

Last remark: this proof is not the main proof used in [2, 3] or [5], and the main different point is the use of "vertical" perturbations, i.e. $\rho_\varepsilon := (1-\varepsilon)\bar{\rho} + \varepsilon\tilde{\rho}$ rather than $\rho_\varepsilon := (id + \varepsilon\xi)_\#\bar{\rho}$.

References

[1] L. Ambrosio, Movimenti minimizzanti, *Rend. Accad. Naz. Sci. XL Mem. Mat. Sci. Fis. Natur.* **113** (1995) 191–246.

[2] L. Ambrosio, N. Gigli, G. Savaré, Gradient flows in metric spaces and in the space of probability measures, *Lectures in Mathematics* (ETH Zürich, 2005).

[3] L. Ambrosio, G. Savaré, Gradient flows of probability measures, *Handbook of differential equations, Evolutionary equations* **3**, ed. by C.M. Dafermos and E. Feireisl (Elsevier, 2007).

[4] E. De Giorgi, New problems on minimizing movements, *Boundary Value Problems for PDE and Applications, C. Baiocchi and J. L. Lions eds.* (Masson, 1993) 81–98.

[5] R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker-Planck equation, *SIAM J. Math. Anal.* **29**(1) (1998) 1–17.