DENSITY IN NONLOCAL NEWTON-SOBOLEV SPACES

1. Introduction

After going through the article [1] I realized that one of the main tools is their lower semicontinuity and compactness result for a sequence of discrete curves to a rectifiable curve. However, due to our nonlocal nature, we can't just copy-paste their technique, since as I mentioned minimizing curves do not necessarily converge to something rectifiable. After playing around a bit and thinking about possible modifications, I devised a strategy involving compactness and lower semicontinuity arguments for cadlag curves with finite variation instead.

Even better, it seems that we are allowed to only consider cadlag curves with pure jumps, which means we can forgo several parts of their paper, and moreover, we can treat more general metrics.

The downside is that it seems to be only possible to prove Lipschitz density in this way for functions whose discrete upper gradient is a 'cadlag upper gradient' for 'almost every' cadlag curve (with 'almost every' depending on the jump kernel and the p-power of the energy). But to verify that the discrete upper gradient is a cadlag upper gradient one either needs continuity, or know that it can be approximated in a suitable way by Lipschitz functions. Which, aside from continuous functions, feels a bit circular :) (However, when V is discrete, this would settle the question.)

A similar game holds for metric measure spaces, in the differences between Newton-Sobolev functions (aka upper gradient along paths, in which case [1] ensures density for Lipschitz functions), Hajlasz Sobolev spaces (for which density of Lipschitz functions holds) and various forms of Cheeger energies (as relaxed versions of integrals of local slopes). The full equivalence usually only holds when gradient flow techniques are used, as in Ambrosio's work. Note however, in the fractional setting, that fractional Sobolev and fractional Hajlasz-Sobolev spaces no longer coincide.

Now, before I'll sketch out the ideas, I will go over several assumptions and conjectures that I need.

2. Assumptions and conjectures

For now, let us make our life simpler, and later think about how the assumptions can be relaxed.

Assumption 2.1. *We take $V = [0,1]$, $E = V \times V$, $D = \{(x,x) | x \in V\}$, $E' = E \setminus D$, and $d(x,y) = |x-y|$. Let $\pi \in \mathcal{M}^+(V)$, $\kappa(x, \mathrm{d}y)$ a singular jump kernel satisfying the detailed balance condition with respect to π, $a(x,y) \in B_b(E)$ a measurable symmetric function, and set*

$$\nu(\mathrm{d}x, \mathrm{d}y) := a(x,y)\pi(\mathrm{d}x)\kappa(x, \mathrm{d}y) \tag{2.1}$$

Note that ν is symmetric. With slight abuse of notation, we will also denote $\nu \in \mathcal{M}_{\mathrm{loc}}(E')$ for its restriction to E', which is a Radon measure, and we further assume

$$\int_{E'} d^2(x,y)\nu(\mathrm{d}x, \mathrm{d}y) < \infty. \tag{2.2}$$

We consider a nonnegative continuous measurable function $f \in B_b(V)$ satisfying

$$\int_{E'} |\overline{\nabla} f|^2 \nu(\mathrm{d}x, \mathrm{d}y) < \infty. \tag{2.3}$$

Remark 2.2. Coming back to our setting for the chain rule, I'm assuming that we can find local $L^1(\pi)$-continuity of u_t via estimates on the action or the Fisher information (or maybe time regularization, with care around the endpoints, is enough). For a smooth entropy function ϕ (with ϕ' bounded and smooth), we would then have absolute continuity in time for the regularized entropy functionals,

$$S(\rho_t) - S(\rho_s) = \int_s^t \int_V \phi'(u_r)\, \partial_t u_r \pi(\mathrm{d}x)\, \mathrm{d}r.$$

and, at points of differentiability,

$$\partial_t S(\rho_t) = \int_V \phi'(u_t)\, \partial_t u_r \pi(\mathrm{d}x)$$

If we now can show via a Lipschitz density approach that for any path $(f_t)_{t\in[0,T]}$ such that

$$\int_0^T \Psi^*(\overline{\nabla} f_t)\sqrt{u(x)u(y)}\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y)\,\mathrm{d}t < \infty,$$

we can find bounded Lipschitz approximating functions f_t^n that converge π-a.e. to f and such that above integrals converge, we are done! There is a nontrivial argument in going from static to dynamic case, but still I will for now only consider the static setting. ■

Remark 2.3. Recall that we have an example where Lipschitz density fails, namely the one that I mentioned previously, where $V = [-1,1]$, $\kappa = \kappa_1 + \kappa_2$ and κ_1, κ_2 are sufficiently singular kernels on $[-1,0]$ and $[0,1]$. In this case the closure of Lipschitz functions under the energy are continous functions, while arbitrary functions with finite energy can have discontinuities.

The part where the strategy that I will sketch out below fails is in the fact that for f the Heaviside function with $f(0) = 1$ (or any of the conventions) the discrete upper gradient $|\overline{\nabla} f|$ is not a cadlag upper gradient. Loosely translated, for this example, this means I can find $x, y \in V$ and a countable sequence of points x_i with $x_0 = x$ and x_n converging to y as $n \to \infty$ such that

$$|\overline{\nabla} f|(x,y) > \sum_{i=0}^{\infty} |\overline{\nabla} f|(x_i, x_{i+1}).$$

For example, take $x = -1, y = 0$, $x_n = x_0 + \sum_{i=1}^n 2^{-i}$. Then $|\overline{\nabla} f|(x_i, x_{i+1}) = 0$ for all i while $|\overline{\nabla} f|(x,y) = 1$. If instead f was continuous I expect the discrete gradient to be a cadlag upper gradient, which would imply that the closure of Lipschitz functions in energy consists of continuous functions (which coincides with what we already know).

The dependence on the kernel might not be clear here, but as you will see later the definition of cadlag upper gradient is for ν-almost every cadlag curve, which in the case of sequences of points translates to only considering sequences such that

$$\sum_{i=0}^{\infty} h(x_i, x_{i+1}) < \infty,$$

for some $L^2(\nu)$ function h. In the case that the kernel is not singular enough, the measure ν is finite and we can simply take $h = 1$, which restricts us to only considering *finite* number of points, and for any finite sequence of points with $x_0 = x$, $x_n = y$ we have

$$|\overline{\nabla} f|(x,y) \le \sum_{i=0}^{N-1} |\overline{\nabla} f|(x_i, x_{i+1}).$$

Thus $|\overline{\nabla} f|$ is in this case a cadlad upper gradient for ν-almost every curve, and Lipschitz functions are now dense in the space of functions with bounded energy. ■

Things might work for a general compact metric space, and the compactness or even properness might be relaxed via similar strategies as in the paper.

Below are some statements that might be in a book somewhere, or easy to prove, but for now I'll list them separately.

Conjecture 2.4. *Throughout, we'll assume the following statements are true:*

(i) Let Y be a locally compact Polish space, g lower semicontinous and coercive, and $\{g_n\}_{n\in\mathbb{N}}$ a sequence of continuous functions such that $g_n \uparrow g$. Then for any sequence of Radon measures μ_n converging vaguely to μ, we have

$$\int_Y g \, \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_Y g_n \, \mathrm{d}\mu_n \tag{2.4}$$

(ii) Let Y be a locally compact Polish space, g coercive, and $\{g_n\}_{n\in\mathbb{N}}$ a sequence of functions that $g_n \uparrow g$. Then any sequence of finite Borel measures μ_n with

$$\liminf_{n\to\infty} \int_Y g_n \, \mathrm{d}\mu_n < \infty, \tag{2.5}$$

is tight, and in particular there exists a narrowly converging subsequence.

(iii) If a sequence of cadlag functions $X^n \in \mathbb{D} := D([0,1], V)$ with uniformly bounded variation converges to X, then their jump measures J^n converge vaguely on $[0,T] \times E'$ to the jump measure J of X. Here

$$J(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y) := \sum_i \delta_{t_i, X_{t_i^-}, X_{t_i}}(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y), \tag{2.6}$$

where t_i are the jumping times of X.

The first statement might be done via a modification of Dini's theorem or arguments from [1, Lemma 2.19]. The second statement is clearly true for truncations, but for the general case, I'm not sure. For the third statement, it feels like I should be true and I can not find counterexamples, but we should hit the books or ask an expert.

3. Strategy

Let me state the following definition.

Definition 3.1. *We consider a property to hold for ν-a.e. cadlag curve if there exists a nonnegative function $h \in L^2(\nu)$ such that the property holds for every curve X with*

$$\sum_i h(X_{t_i-}, X_{t_i}) < \infty,$$

where t_i are the jump times of the (pure jump) cadlag curve X.

Moreover, a function $g \in L^2(\nu)$ is a cadlag upper gradient for f if for ν-a.e. cadlag curve

$$|\overline{\nabla} f|(X_0, X_1) \leq \sum_i g(X_{t_i-}, X_{t_i}).$$

I'm interested in the following statement.

Conjecture 3.2. *Suppose that Assumption 2.1 holds. Then for every $f \in L^2(\pi)$ with $|\overline{\nabla} f| \in L^2(\nu)$ such that $|\overline{\nabla} f|$ is a cadlag upper gradient of f, there exists a sequence of Lipschitz functions converging in energy.*

I will now sketch a possible procedure to prove the above. We will follow a similar path as in the article, finding suitable approximations g_ε for our discrete upper gradient $g = |\overline{\nabla} f|$ and using this function to obtain f_n via a minimization procedure. To show that f_n converges a.e. to f is the hard part, and goes through cadlag functions.

Choice of g_ε and g_n: Fix any $\varepsilon > 0$, and choose $\tilde{g}_\varepsilon$, $h \in B^+(E')$ and $\sigma > 0$, in such a way that $g_\varepsilon \geq g$ is lower semicontinous, h is lower semicontinuous and coercive with $h \to \infty$ as $d \to 0$, and

$$\int_{E'} |g - g_\varepsilon|^2 \nu(\mathrm{d}x, \mathrm{d}y) < \varepsilon,$$

where

$$g_\varepsilon = \tilde{g}_\varepsilon + \sigma d + hd.$$

The g_ε stems from the Vitali-Carathéodory Theorem, and the latter two follow from the fact that $d^2\nu$ is a finite measure. Moreover, by Baire's theorem, we can find an increasing sequence of continuous functions $\tilde{g}_n$ that converge pointwise to g on E', and we set $g_n := \tilde{g}_n \wedge nd$. Note that g_n as well are an increasing sequence of continuous functions that converge pointwise to g on E'. Finally, choose a compact set K such that the restriction f_K is continuous and $\mu(K^c) < \varepsilon$.

Similar as [1, Lemma 2.9], but now more straightforward in the nonlocal setting, we have by the extended dominated convergence theorem that if $f_\varepsilon \to f$ in $L^2(\pi)$, $|\overline{\nabla} f_\varepsilon| \leq g_\varepsilon$ and g_ε converges to g in $L^2(\nu)$, then f_ε converges to f in energy.

Approximating sequence f_n: For each $n \in \mathbb{N}$ we define

$$f_n(x) = \min\left\{ \inf_{p_0,\dots,p_N} f(p_0) + \sum_{k=0}^{N-1} g_n(x_k, x_{k+1}), \|f\|_\infty \right\} \tag{3.1}$$

where the infimum is taken over all discrete paths $(x_0, \dots, x_N)$ with $x_0 \in K$. Note that

$$|\overline{\nabla} f_n(x,y)| \leq |\overline{\nabla} g_n(x,y)|, \qquad \text{for all } x, y \in V$$

and in particular $f_n \in Lip_b(V)$ for each n.

Pointwise convergence of f_n on K: Note that $f_n \leq f$ on K. Moreover, let x_i^n be the points corresponding to n^{-1}-approximate minimizers of the above functional. Note that for any sequence of $x_0^n \in K$ we can find a converging subsequence in K, and by continuity f_n on K it follows that $f(x_0^n) \to f(x_0)$. Therefore to obtain convergence of f_n on K all that is required is the liminf statement

$$\liminf_{n\to\infty} \sum_{k=0}^{N_n-1} g_n(x_k^n, x_{k+1}^n) \geq |\overline{\nabla} f| \tag{3.2}$$

for $x, y \in K$. We will show compactness and lower semicontinuity for the discrete paths, in essence doing a Γ-convergence result for cadlag curves, resolving the proof.

Compactness and lower semicontinuiy for cadlag curves Assume the limiting sum above is finite for a suitable subsequence. By the bound of g_ε for $\sigma > 0$ the lenghts are bounded, i.e.

$$\limsup_{n\to\infty} L_n < \infty, \qquad L_n := \sum_{n=1}^{N_n-1} d(x_i^n, x_{i+1}^n).$$

Similar as in [1], set interpolating times $t_0^n := 0$, $t_N^n = 1$ and

$$t_i^n := L_n^{-1} \sum_{n=1}^{i-1} d(x_i^n, x_{i+1}^n),$$

but instead of linear interpolations our paths will have jumps. Namely, construct the cadlag curves X^n with jump times t_i^n and values x_i^n, and bounded variation L_n. The cadlag modulus of continuity $\omega^\delta(X^n)$ can shown to vanish uniformly in n as $\delta \to 0$, and thus we can find a converging subsequence and a limiting cadlag path X.

Moreover, by part (iii) of Conjecture 2.4 the jump measures J^n of X^n converge vaguely in $[0,T]\times E'$ to the jump measure J of X. Note that

$$\sum_{k=0}^{N_n-1} g_n(x_k^n, x_{k+1}^n) = \int_{[0,T]\times E'} g_n \,\mathrm{d}J^n, \tag{3.3}$$

and by part (ii) of Conjecture 2.4 and coercivity of h we find that the uniformly finite measures $d(x,y)J^n$ converge to $d(x,y)J$. By lower semicontinuity of variation of cadlag paths we deduce that all the variation of X comes from the jump measure, and hence the limiting path has no continuous part, i.e it is pure jump. The desired statement now follows from the fact that $|\overline{\nabla} f|$ is a cadlag upper gradient, from vague convergence of J^n and (i), and of Conjecture 2.4, since

$$\begin{aligned}
|f(X(1)) - f(X(0))| &\leq \sum_i |\overline{\nabla} f|(X_{t_i-}, X_{t_i}) \\
&= \int_{[0,T]\times E'} |\overline{\nabla} f| \,\mathrm{d}J \\
&\leq \int_{[0,T]\times E'} |\overline{\nabla} g_\varepsilon| \,\mathrm{d}J \\
&\leq \liminf_{n\to\infty} \int_{[0,T]\times E'} |\overline{\nabla} g_n| \,\mathrm{d}J^n.
\end{aligned}$$

References

[1] S. Eriksson-Bique. Density of lipschitz functions in energy. *Calculus of Variations*, 2023.