

Totally convex functions, L^2 -Optimal transport for laws of random measures, and solution to the Monge problem

Alessandro Pinzi ^{*}, Giuseppe Savaré [†]

September 2, 2025

Abstract

We study the Optimal Transport problem for laws of random measures in the Kantorovich-Wasserstein space $\mathcal{P}_2(\mathcal{P}_2(\mathbb{H}))$, associated with a Hilbert space \mathbb{H} (with finite or infinite dimension) and for the corresponding quadratic cost induced by the squared Wasserstein metric in $\mathcal{P}_2(\mathbb{H})$.

Despite the lack of smoothness of the cost, the fact that the space $\mathcal{P}_2(\mathbb{H})$ is not Hilbertian, and the curvature distortion induced by the underlying Wasserstein metric, we will show how to recover at the level of random measures in $\mathcal{P}_2(\mathcal{P}_2(\mathbb{H}))$ the same deep and powerful results linking Euclidean Optimal Transport problems in $\mathcal{P}_2(\mathbb{H})$ and convex analysis.

Our approach relies on the notion of totally convex functionals, on their total subdifferentials, and their Lagrangian liftings in the space square integrable \mathbb{H} -valued maps $L^2(\mathbb{Q}, \mathbb{M}; \mathbb{H})$.

With these tools, we identify a natural class of regular measures in $\mathcal{P}_2(\mathcal{P}_2(\mathbb{H}))$ for which the Monge formulation of the OT problem has a unique solution and we will show that this class includes relevant examples of measures with full support in $\mathcal{P}_2(\mathbb{H})$ arising from the push-forward transformation of nondegenerate Gaussian measures in $L^2(\mathbb{Q}, \mathbb{M}; \mathbb{H})$.

Contents

1	Introduction	2
2	Notation and preliminary results	6
2.1	The L^2 -Wasserstein space	8
2.2	Regularity of measures and Monge formulation in $\mathcal{P}_2(\mathbb{H})$	9
2.3	Lagrangian representations.	11
2.4	Lagrangian representation of the laws of random measures of $\mathcal{P}_2(\mathbb{H})$	12
3	Totally convex functionals	14
3.1	A Kantorovich version of the Legendre-Fenchel transform in $\mathcal{P}_2(\mathbb{H})$.	15
3.2	c -concave functions and their c -super differentials	18
4	Multivalued probability fields	20
4.1	Totally monotone and cyclically monotone MPVF	20
4.2	Marginal projections of MPVF	24
5	L^2-Random Optimal Transport	25
5.1	Random couplings and couplings of random measures	26
5.2	L^2 -Optimal Transport for laws of random measures	26
5.3	The structure of minimal geodesics in $\mathcal{P}_2(\mathbb{H})$	28
5.4	Lifting (laws of) random measures of $\mathcal{P}_2(\mathbb{H})$ to measures on Lagrangian maps in $\mathcal{P}_2(\mathcal{H})$.	31

^{*}Bocconi University, Department of Decision Sciences, via Roentgen 1, 20136 Milano (Italy). Email: alessandro.pinzi@phd.unibocconi.it

[†]Bocconi University, Department of Decision Sciences and BIDSa, via Roentgen 1, 20136 Milano (Italy). Email: giuseppe.savare@unibocconi.it

6 Random Gaussian-null sets and strict Monge formulation of OT via nonlocal totally cyclically monotone fields	33
6.1 The strict Monge formulation	34
6.2 Regular and super-regular measures in $\mathcal{P}_2(\mathbb{H})$ and solution to the Monge problem	37
6.3 Examples of super-regular measures for finite dimensional \mathbb{H}	41

1 Introduction

One of the most elegant and fascinating aspects of Optimal Transport Theory [RR98; Vil09] for the classical quadratic cost $c(x, y) := \frac{1}{2}|x - y|^2$ in \mathbb{R}^d is its link with convex analysis, which has been thoroughly exploited by Knott-Smith, Rachev-Rüschendorf, and Brenier [KS84; RR90; Bre91] (see also [CM89]). If μ, ν belong to the space $\mathcal{P}_2(\mathbb{R}^d)$ of probability measures with finite quadratic moment

$$m_2^2(\mu) := \int_{\mathbb{R}^d} |x|^2 d\mu(x), \quad (1.1)$$

is in fact possible to prove that a coupling $\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ with marginals μ, ν (we say that $\gamma \in \Gamma(\mu, \nu)$) is optimal for the L^2 -Kantorovich-Wasserstein metric

$$w_2^2(\mu, \nu) := \min \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma : \gamma \in \Gamma(\mu, \nu) \right\} = \min \left\{ \mathbb{E}[|X - Y|^2] : X \sim \mu, Y \sim \nu \right\} \quad (1.2)$$

if and only if its support $S := \text{supp}(\gamma) \subset \mathbb{R}^d \times \mathbb{R}^d$ is cyclically monotone, i.e.

$$\sum_{n=1}^N \langle y_n, x_n - x_{\sigma(n)} \rangle \geq 0 \quad \text{for every } N \in \mathbb{N}, \text{ every choice of } (x_n, y_n) \in S \quad (1.3)$$

and every permutation $\sigma \in S_N$.

The dual formulation of (1.2) and the fact that every cyclically monotone subset of $\mathbb{R}^d \times \mathbb{R}^d$ is contained in the graph of the convex subdifferential $\partial\varphi$ of a convex and lower semicontinuous function $\varphi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$ implies that we can find optimal conjugate Kantorovich potentials φ, φ^* such that for all optimal couplings γ_{opt} solving (1.2) the support of γ_{opt} is contained in the graph of $\partial\varphi$ and

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} x \cdot y d\gamma_{\text{opt}}(x, y) = \int_{\mathbb{R}^d} \varphi(x) d\mu(x) + \int_{\mathbb{R}^d} \varphi^*(y) d\nu(y). \quad (1.4)$$

The simple but crucial link between the first integral in (1.4) and the Optimal Transport problem (1.2) is guaranteed by the specific property of the Euclidean norm in \mathbb{R}^d , for which

$$w_2^2(\mu, \nu) = m_2^2(\mu) + m_2^2(\nu) - 2[\mu, \nu], \quad [\mu, \nu] := \max_{\gamma \in \Gamma(\mu, \nu)} \int x \cdot y d\gamma, \quad (1.5)$$

so that the minimum problem (1.2) and the maximum problem (1.5) defining $[\mu, \nu]$ share the same class of optimizers.

Since the subdifferential of a convex function φ is a singleton at every differentiability point of φ and the set of non-Gateux-differentiability points of a convex Lipschitz function can be covered by a countable union of d.c. hypersurfaces [Zaj79] (a result which holds even in infinite dimensional Hilbert spaces), one can prove that when $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ does not give mass to d.c. hypersurfaces (we say that $\mu \in \mathcal{P}_2^r(\mathbb{R}^d)$ is *regular*) there exists a unique optimal coupling γ_{opt} which is moreover concentrated on the graph of a Borel map f , thus satisfying $f_{\#}\mu = \nu$. The class of regular measures $\mathcal{P}_2^r(\mathbb{R}^d)$ coincide with the class of atomless measures when $d = 1$ and contains all the measures absolutely continuous with respect to the d -dimensional Lebesgue measure \mathcal{L}^d for every dimension d .

This remarkable combination of analytic and geometric arguments yields the solution of the Monge formulation of (1.2)-(1.5), i.e. the existence of an optimal transport map f such that

$$f_{\#}\mu = \nu, \quad w_2^2(\mu, \nu) = \int_{\mathbb{R}^d} |f(x) - x|^2 d\mu(x), \quad [\mu, \nu] = \int_{\mathbb{R}^d} f(x) \cdot x d\mu(x). \quad (1.6)$$

The convex theory for random measures

Whenever (X, d_X) is a (complete, separable) metric space, the construction of the L^2 -Kantorovich-Wasserstein metric can be naturally extended to $\mathcal{P}_2(X)$, the space of probability measures on X with finite quadratic moment; the resulting (squared) metric

$$W_{2,d_X}^2(\mu, \nu) := \min \left\{ \int_{X \times X} d_X^2(x, y) d\gamma : \gamma \in \Gamma(\mu, \nu) \right\} = \min \left\{ \mathbb{E}[d_X^2(X, Y)] : X \sim \mu, Y \sim \nu \right\} \quad (1.7)$$

inherits relevant geometric properties from the underlying space X and the problem still admits a dual formulation involving Kantorovich potentials.

Since $(\mathcal{P}_2(\mathbb{R}^d), w_2)$ is a complete and separable metric space, a nice example of applications of the metric perspective is provided by the possibility to lift Optimal Transport problems at the level of the laws of the so-called random measures, i.e. probability measures in the space $\mathfrak{P}_2(\mathbb{R}^d) := \mathcal{P}_2(\mathcal{P}_2(\mathbb{R}^d))$ endowed with the Kantorovich-Wasserstein metric $W_2 := W_{2,w_2}$

$$W_2^2(\mathbf{M}, \mathbf{N}) := \min \left\{ \int_{\mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d)} w_2^2(\mu, \nu) d\Pi(\mu, \nu) : \Pi \in \Gamma(\mathbf{M}, \mathbf{N}) \right\} \quad (1.8)$$

$$= \min \left\{ \mathbb{E}[w_2^2(M, N)] : M \sim \mathbf{M}, N \sim \mathbf{N} \right\}, \quad \mathbf{M}, \mathbf{N} \in \mathfrak{P}_2(\mathbb{R}^d). \quad (1.9)$$

A similar class of problems, starting however from a smooth and compact Riemannian manifold instead of \mathbb{R}^d , have recently been studied by [EP25]. Here we want to focus on the Euclidean case (also including infinite dimensional separable Hilbert spaces), which shows distinguished and remarkable features and has recently attracted a lot of attention in view of many interesting applications [BVK25; FHS23; Acc+25; PS25; CL24].

It is well known that in general metric spaces (X, d_X) (as, in particular, $(\mathcal{P}_2(\mathbb{R}^d), w_2)$) the link between Optimal Transport problems and convex analysis is typically lost, mainly due to the possible lack of a linear structure in X ; even when $X = \mathbb{R}^d$ but the metric cost is induced by a non-Hilbertian norm $\|\cdot\|$, convexity and Legendre duality of Kantorovich potentials do not hold, since the optimality condition and the structure of optimal transport maps involve the *nonlinear* differential associated with $\frac{1}{2}\|\cdot\|^2$.

Even if $(\mathcal{P}_2(\mathbb{R}^d), w_2)$ is a genuine metric space which is positively curved in the sense of Alexandrov [AGS08], the aim of the present paper is to show that

a large part of the convexity landscape of the Euclidean case remarkably holds also for the Optimal Transport problem (1.8) for laws of random measures in $\mathfrak{P}_2(\mathbb{R}^d)$, if we use the appropriate notion of total displacement convexity in $\mathcal{P}_2(\mathbb{R}^d)$,

i.e. convexity along interpolation curves induced by arbitrary couplings.

Such a nice and somehow unexpected result relies on two important properties. First of all, as for (1.5), the lifted Wasserstein metric W_2 given by (1.8) still satisfies a similar identity

$$W_2^2(\mathbf{M}_1, \mathbf{M}_2) = M_2^2(\mathbf{M}_1) + M_2^2(\mathbf{M}_2) - 2\llbracket \mathbf{M}_1, \mathbf{M}_2 \rrbracket, \quad (1.10)$$

$$M_2^2(\mathbf{M}) = \int m_2^2(\mu) d\mathbf{M}(\mu), \quad \llbracket \mathbf{M}_1, \mathbf{M}_2 \rrbracket = \max_{\Pi \in \Gamma(\mathbf{M}_1, \mathbf{M}_2)} \int [\mu_1, \mu_2] d\Pi(\mathbf{M}_1, \mathbf{M}_2),$$

so that we can study the equivalent formulation in terms of the maximization of the function $(\mu_1, \mu_2) \mapsto [\mu_1, \mu_2]$.

Even if $[\cdot, \cdot]$ is not bilinear, it exhibits many analogies with a scalar product, in particular along displacement interpolation of measures, i.e. curves obtained by general couplings $\boldsymbol{\mu} \in \mathcal{P}(R^d \times \mathbb{R}^d)$ via the dynamic push forward

$$\mu_t = (\pi_t^{1 \rightarrow 2})_{\#} \boldsymbol{\mu}, \quad \pi_t^{1 \rightarrow 2}(x_1, x_2) := (1-t)x_1 + tx_2, \quad t \in [0, 1], \quad \boldsymbol{\mu} \in \Gamma(\mu_0, \mu_1). \quad (1.11)$$

When μ is an optimal coupling between μ_0 and μ_1 for the Wasserstein metric (1.2) the curve $t \mapsto \mu_t$ is in fact a minimal, constant speed geodesic in $\mathcal{P}_2(\mathbb{R}^d)$ (which we call *optimal displacement interpolation*) and plays a crucial geometric role in the Optimal Transport setting, since the pioneering paper of McCann [McC97]. In particular a function $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\infty\}$ is called geodesically (or displacement) convex if for every $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ there exists a geodesic $(\mu_t)_{t \in [0,1]}$ connecting μ_0 to μ_1 such that $t \mapsto \phi(\mu_t)$ is convex in $[0, 1]$.

A more restricted class of functions are in fact convex along *any* displacement interpolation, induced by arbitrary couplings $\mu \in \Gamma(\mu_0, \mu_1)$ as in (1.11), thus satisfying

$$\phi((\pi_t^{1 \rightarrow 2})\# \mu) \leq (1-t)\phi(\mu_0) + t\phi(\mu_1) \quad \text{for every } \mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d), \mu \in \Gamma(\mu_0, \mu_1). \quad (1.12)$$

Such a class of *totally displacement convex* functionals, thoroughly studied in [CSS23a], enjoys better properties. Starting from the fact that (see also the inspiring notes by Brenier [Bre20])

$$\mu \mapsto [\mu, \nu] \quad \text{is totally convex for every } \nu \in \mathcal{P}_2(\mathbb{R}^d), \quad (1.13)$$

we can introduce the Kantorovich-Legendre-Fenchel transform

$$\phi^*(\nu) := \sup_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} [\nu, \mu] - \phi(\mu) \quad (1.14)$$

and prove that proper, totally convex, and lower semicontinuous functions are characterized by the identity $\phi = \phi^{**}$ as for convex functions in Euclidean spaces (see Section 3.1).

The Rockafellar type transformation (1.14) is a typical technique in Optimal Transport (where it is applied to the concave version of the potentials and it is called *c-transform*). What distinguishes the Euclidean and the current random-Euclidean setting is the possibility to characterize c-transforms and self-biconjugate functions in a simple way using convexity.

This remarkable property allows us to retrace the same strategy as in the finite-dimensional Euclidean theory and has relevant applications:

1. it provides an intrinsic characterization for the optimal Kantorovich potentials associated with the Wasserstein metric (1.8) (Section 3.2): they coincide with the class of totally convex functionals;
2. it allows for a deeper understanding of the Wasserstein (total) subdifferential of ϕ [AGS08; CSS23a], interpreted as a *Multivalued Probability Vector Field* (Section 4);
3. it clarifies the structure of optimal couplings and minimal geodesics (Sections 5.2, 5.3)
4. it suggests a general lifting strategy to the L^2 -space of Lagrangian maps, importing a Hilbertian perspective for regularity of laws of random measures (Section 6.2) which plays a crucial role in proving uniqueness for solution to the Monge formulation.

Let us briefly summarize the main points: combining Kantorovich duality and total displacement convexity, we will prove that for every $M, N \in \mathcal{P}_2(\mathbb{R}^d)$ there exists a pair of conjugate totally convex function $\phi, \phi^* : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{\infty\}$ such that

$$\int \phi(\mu) dM(\mu) + \int \phi^*(\nu) dN(\nu) = \llbracket M, N \rrbracket. \quad (1.15)$$

We will then show that there is a natural correspondence between optimal couplings $\Pi \in \Gamma_o(M, N)$ minimizing (1.8) and random coupling laws $P \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ which can be used to characterize W_2 as

$$W_2^2(M, N) = \min \left\{ \int \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma(x, y) dP(\gamma) : (\pi_{\#}^1)\#P = M, (\pi_{\#}^2)\#P = N \right\}. \quad (1.16)$$

Random couplings are optimal if and only if their support is contained in the so-called total subdifferential $\partial_t \phi$ of the optimal potential ϕ in (1.15), a closed subset of $\mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$. Optimality can also be characterized by total cyclical monotonicity and in particular implies that optimal random couplings are concentrated on the subset $\mathcal{P}_{2,o}(\mathbb{R}^d \times \mathbb{R}^d)$ of the usual optimal couplings in $\mathbb{R}^d \times \mathbb{R}^d$.

As for the classic subdifferentials of convex functions, among all the elements of the total subdifferential $\partial_t \phi$ there is a minimal distinguished one (called the minimal section and denoted by $\partial_t^\circ \phi$) which can be represented by a nonlocal deterministic field $\mathbf{f}^\circ : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d$:

$$\gamma \in \partial_t^\circ \phi(\mu) \quad \Leftrightarrow \quad \gamma = (\mathbf{i} \times \mathbf{f}^\circ(\cdot, \mu))\# \mu. \quad (1.17)$$

Total cyclical monotonicity can be more easily understood in the case of \mathbf{f}° , where it reads as

$$\sum_{n=1}^N \int \langle \mathbf{f}^\circ(x_n, \mu_n), x_n - x_{\sigma(n)} \rangle d\boldsymbol{\mu}(x_1, \dots, x_N) \geq 0 \quad (1.18)$$

for every $\mu_i \in D(\partial_t \phi)$, $\boldsymbol{\mu} \in \Gamma(\mu_1, \mu_2, \dots, \mu_N)$, $\sigma \in S_N = \text{Sym}(\{1, \dots, N\})$.

As a byproduct, when $\partial_t \phi[\mu]$ reduces to a singleton for M-a.e. $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the OT problem (1.8) has a unique solution Π_o which is also concentrated on a Monge map $\mathcal{F} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathcal{P}_2(\mathbb{R}^d)$, $\Pi = (\text{Id} \times \mathcal{F})\# \mathbb{M}$. Moreover, \mathcal{F} can be expressed as a push-forward via \mathbf{f}° :

$$\mathcal{F}(\mu) = \mathbf{f}^\circ(\cdot, \mu)\# \mu, \quad \mathbb{W}_2^2(\mathbb{M}, \mathbb{N}) = \int \mathbb{w}_2^2(\mu, \mathcal{F}(\mu)) d\mathbb{M}(\mu) = \int \left(\int |\mathbf{f}^\circ(x, \mu) - x|^2 d\mu(x) \right) d\mathbb{M}(\mu). \quad (1.19)$$

It turns out that \mathbf{f}° solves the strict Monge formulation of (1.8)

$$\inf \left\{ \int \left(\int |\mathbf{f}(x, \mu) - x|^2 d\mu(x) \right) d\mathbb{M}(\mu) : \mathbf{f} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d, \mathbb{N} = \int \delta_{\mathbf{f}(\cdot, \mu)\# \mu} d\mathbb{M}(\mu) \right\}. \quad (1.20)$$

Lagrangian parametrizations, Hilbertian liftings, and laws of Gaussian-generated random measures

The above discussion raises the crucial question to find general condition on \mathbb{M} ensuring that $\partial_t \phi$ is concentrated on a singleton M-a.e. A similar problem has also been considered in [EP25] by assuming suitable regularity properties on the Dirichlet form associated with \mathbb{M} in $\mathcal{P}_2(M)$, in particular the Rademacher property studied in [Del20].

Here we adopt a different perspective, inspired by the crucial fact that Optimal Kantorovich potentials are totally convex on $\mathcal{P}_2(\mathbb{R}^d)$, i.e. convex along arbitrary couplings as in (1.12). It is then possible to apply a natural Lions-Lagrangian lifting technique that has been systematically studied in [CSS23a] in the context of convex analysis and evolution problems (but see also the relevant notions of L-convexity, L-monotonicity [Car13; CD18], Fréchet differentiability [GT19], and the discussion of [CSS23a, Remark 5.4]).

The main idea is to introduce a standard Borel space (Q, \mathcal{F}_Q) endowed with a nonatomic probability measure \mathbb{M} which can represent every measure of $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ as the law $X_\# \mathbb{M}$ of a map X in the Hilbert space $\mathcal{H} := L^2(Q, \mathbb{M}; \mathbb{R}^d)$. The law map $\iota = \iota_{\mathbb{M}} : X \mapsto X_\# \mathbb{M}$ is a 1-Lipschitz surjection from \mathcal{H} to $\mathcal{P}_2(\mathbb{R}^d)$ and total convexity of ϕ in $\mathcal{P}_2(\mathbb{R}^d)$ is equivalent to the usual convexity of $\hat{\phi} := \phi \circ \iota$ in the Hilbert space \mathcal{H} . Such a correspondence also holds at the levels of the respective subdifferentials, so that the measures μ where $\partial_t \phi[\mu]$ reduces to a singleton correspond to maps $X \in \mathcal{H}$ where $\hat{\phi}$ is Gateaux-differentiable.

Since every measure \mathbb{M} on $\mathcal{P}_2(\mathbb{R}^d)$ can be obtained as the push-forward $\iota_\# \mathbb{m}$ of a measure \mathbb{m} on \mathcal{H} , it is tempting to lift the problem of M-a.e. differentiability in $\mathcal{P}_2(\mathbb{R}^d)$ to the problem of \mathbb{m} -a.e. differentiability of $\hat{\phi}$ in \mathcal{H} , for which many powerful result are known. In particular, d.c. hypersurfaces in a Hilbert space are Gaussian null [Aro76; Phe78; Bog84; Csö99]), i.e. are negligible w.r.t. every nondegenerate Gaussian measure.

We will systematically pursue this direction, showing that the strict Monge problem in $\mathcal{P}_2(\mathbb{R}^d)$ has a unique solution if \mathbf{M} satisfies two conditions:

- (R1) it is concentrated on the set of *regular* measures $\mathcal{P}_2^r(\mathbb{R}^d)$, i.e. $\mu \in \mathcal{P}_2^r(\mathbb{R}^d)$ for \mathbf{M} -a.e. μ ;
- (R2) $\mathbf{M}(B) = 0$ on every Borel set $B \subset \mathcal{P}_2^r(\mathbb{R}^d)$ such that $\iota^{-1}(B)$ is contained in a d.c. hypersurface of the Hilbert space $\mathcal{H} = L^2(Q, \mathbb{M}; \mathbb{R}^d)$.

We call $\mathcal{P}_2^{rr}(\mathbb{R}^d)$ the set of *super-regular measures* satisfying the two conditions above; it is worth noticing that the first condition has also been assumed by [EP25], whereas the second one is strongly related to the lifting procedure. We will show that those conditions are nearly optimal if we look for measures \mathbf{M} for which the usual Monge formulation has at least one solution for every target measure \mathbf{N} . Both conditions are stable if we replace \mathbf{M} by $\mathbf{M}' \ll \mathbf{M}$.

A simple way to construct measures satisfying (R2) is to start from a regular measure $\mathbf{m} \in \mathcal{P}_2^r(\mathcal{H})$ and taking its push forward $\mathbf{M} = \iota_{\#}\mathbf{m}$. We will focus on the relevant case of Laws of Gaussian-Generated Random Measures (LGGRM), i.e. measures in $\mathcal{P}_2(\mathbb{R}^d)$ of the form $\mathbf{G} = \iota_{\#}\mathbf{g}$ where \mathbf{g} is a non-degenerate Gaussian measure in \mathcal{H} [Bog98]. They have full support in $\mathcal{P}_2(\mathbb{R}^d)$ and will satisfy condition (R2). We will show that *every* LGGRM in dimension $d = 1$ is super-regular and we will exhibit a large class of super-regular LGGRM in every dimension. As a byproduct, we obtain that the class of super-regular measures is dense in $\mathcal{P}_2(\mathbb{R}^d)$.

We have developed our analysis in the case of the “Euclidean” 2-Wasserstein metric, since we believe that its distinguished features deserve a separate analysis. Since (finite) dimension play a role only in the final discussion of super-regular measures, we decided to develop our theory in an arbitrary separable Hilbert space \mathcal{H} , even if all the results are new also in finite dimension. In addition, we think that many tools we have introduced in this paper may be useful to study the more general case of the L^p -Wasserstein metric induced by a smooth norm in \mathbb{R}^d . We are also confident that the class of LGGRM measures may reveal interesting features from the viewpoint of the induced Dirichlet form in $\mathcal{P}_2(\mathbb{R}^d)$ (see also [FSS23]), and we plan to address both questions in a forthcoming paper.

Plan of the paper

After a quick recap of the main notions and notation in Section 2, we will deal with totally convex function and their link with optimal Kantorovich potentials in Section 3. The related notions of multivalued probability vector fields and total subdifferentials are developed in Section 4.

Section 5 is devoted to applications of these tools to the Optimal Transport problem in $\mathcal{P}_2(\mathbb{R}^d)$. The (strict) Monge formulation, its solution for super-regular measures, and the discussion of the relevant examples is presented in the last section 6.

Acknowledgments

We wish to thank Anna Korba, Eugenio Regazzini and Luciano Tubaro for stimulating and insightful discussions.

GS has been supported by the MIUR-PRIN 202244A7YL project Gradient Flows and Non-Smooth Geometric Structures with Applications to Optimization and Machine Learning, by the INDAM project E53C23001740001, and by the Institute for Advanced Study of the Technical University of Munich, funded by the German Excellence Initiative.

2 Notation and preliminary results

The following table contains the main notation that we shall use throughout the paper:

\mathbb{H}	separable Hilbert space;
$\mathcal{P}_2(\mathbb{H})$	the space of probability measures on \mathbb{H} with finite quadratic moment, 2.1;
$\mathcal{P}_2^r(\mathbb{H})$	regular measures, Def. 2.3;
$\mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H})$	the set of optimal couplings, 2.1;
w_2	the L^2 -Wasserstein metric on $\mathcal{P}_2(\mathbb{H})$, 2.1;
$[\cdot, \cdot]$	the maximal correlation pairing, Definition 2.2;
μ, ν, \dots	notation for typical measures in $\mathcal{P}_2(\mathbb{H})$;
$\mathfrak{P}_2(\mathbb{H}) = \mathcal{P}_2(\mathcal{P}_2(\mathbb{H}))$	the space of probability measures on $\mathcal{P}_2(\mathbb{H})$ with finite quadratic moment, 2.1;
W_2	the L^2 -Wasserstein metric on $(\mathcal{P}_2(\mathbb{H}), w_2)$, 2.1;
$\mathbb{M}, \mathbb{N}, \dots$	notation for typical laws of random measures in $\mathfrak{P}_2(\mathbb{H})$;
$f_{\#}\mu$	push forward of a measure μ via the map f , 2;
$f_{\#}\mathbb{M} = (f_{\#})_{\#}\mathbb{M}$	iterated push forward of a measure \mathbb{M} via the map f defined in \mathbb{H} ;
π^i	projection on the i -th coordinate in a product space, 2;
$S_N = \text{Sym}(\{1, \dots, N\})$	the symmetric group of permutations of $\{1, \dots, N\}$;
$\Gamma(\mu, \nu), \Gamma_o(\mu, \nu)$	set of (resp. optimal) couplings with marginals μ, ν , 2;
$\mathcal{P}^{\text{det}}(\mathbb{X}^1 \times \mathbb{X}^2)$	set of deterministic couplings, concentrated on the graph of a Borel map, (2.4);
$(\mathbb{Q}, \mathcal{F}_{\mathbb{Q}}, \mathbb{M}), (\Omega, \mathcal{F}, \mathbb{P})$	standard Borel spaces endowed with nonatomic probability measures, 2.3;
$\mathcal{H} = L^2(\mathbb{Q}, \mathcal{F}_{\mathbb{Q}}, \mathbb{M}; \mathbb{H})$	the L^2 space of \mathbb{H} valued maps, 2.3;
ι	the law-pushforward map $\iota(X) := X_{\#}\mathbb{M}$ from \mathcal{H} to $\mathcal{P}_2(\mathbb{H})$, 2.3;
\mathbf{i}	the identity vector field $\mathbf{i}(x) = x$ in \mathbb{H} ;
Id	the identity map in a general set;
\mathbf{m}, \mathbf{g}	typical probability measures in $\mathcal{P}_2(\mathcal{H})$;
$\mathbb{G}(\mathbb{Q})$	the group of measure preserving isomorphisms of $(\mathbb{Q}, \mathcal{F}_{\mathbb{Q}}, \mathbb{M})$, 2.3
$\hat{\phi}$	$= \phi \circ \iota$, the lifting of a function defined in $\mathcal{P}_2(\mathbb{H})$ to \mathcal{H} , 2.3;
ϕ^*	the Kantorovich-Legendre-Fenchel conjugate of a function ϕ , Def. 3.6;
$\mathbf{F}, \hat{\mathbf{F}}$	a typical MPVF in $\mathcal{P}_2(\mathbb{H} \times \mathbb{H})$ and its lifting to $\mathcal{H} \times \mathcal{H}$, 4;
$\partial_t \phi, \partial_t^\circ \phi$	the total subdifferential of ϕ and its minimal section, Def. 4.4;
$\nabla_W \phi$	the nonlocal field associated to $\partial_t^\circ \phi$;
$\text{RG}(\mathbb{M}, \mathbb{N})$	random couplings between two laws in $\mathfrak{P}_2(\mathbb{H})$, 5.1;
$[\cdot, \cdot]$	cost in $\mathfrak{P}_2(\mathbb{H})$ induced by the maximal correlation pairing $[\cdot, \cdot]$, Def. 5.2;
$\mathfrak{P}_2^r(\mathbb{H})$	set of super-regular measures in $\mathfrak{P}_2(\mathbb{H})$, Def. 6.6;

If \mathbb{X} is a Polish topological space (i.e. it is separable and its topology is induced by a complete metric) we denote by $\mathcal{P}(\mathbb{X})$ the space of Borel probability measures on \mathbb{X} endowed with the weak (Polish) topology in duality with continuous and bounded real functions. Given a Borel map between Polish spaces $f : \mathbb{X} \rightarrow \mathbb{Y}$ and $\mu \in \mathcal{P}(\mathbb{X})$ we denote by $f_{\#}\mu$ the image measure $\mu \circ f^{-1}$ given by $f_{\#}\mu(B) := \mu(f^{-1}(B))$ for every Borel subset B of \mathbb{Y} . In the case of a cartesian product of Polish spaces $\mathbf{X} := \prod_{i=1}^N \mathbb{X}^i$ we will denote by $\pi^i : \mathbf{X} \rightarrow \mathbb{X}^i$ the map $\pi^i(x_1, \dots, x_N) := x_i$, and similarly $\pi^{ij}(x_1, \dots, x_N) := (x_i, x_j)$.

A probability kernel is a Borel map $\kappa : \mathbb{X} \ni x \mapsto \kappa_x \in \mathcal{P}(\mathbb{Y})$; if $\mu \in \mathcal{P}(\mathbb{X})$ there exists a unique Borel probability measure $\mu \otimes \kappa = \int_{\mathbb{X} \times \mathbb{Y}} \delta_x \otimes \kappa_x d\mu(x) \in \mathcal{P}(\mathbb{X} \times \mathbb{Y})$ which satisfies

$$\int_{\mathbb{X} \times \mathbb{Y}} f(x, y) d(\mu \otimes \kappa)(x, y) := \int_{\mathbb{X}} \left(\int_{\mathbb{Y}} f(x, y) d\kappa_x(y) \right) d\mu(x) \quad (2.1)$$

for every bounded (or nonnegative) real Borel function f defined in $\mathbb{X} \times \mathbb{Y}$.

Elements of $\mathcal{P}(\mathbb{X} \times \mathbb{Y})$ are also called *couplings* or *transport plans*. The universal disintegration Theorem [Kal17, Corollary 1.26] says that there exists a Borel map $\mathcal{K} : \mathbb{X} \times \mathcal{P}(\mathbb{X} \times \mathbb{Y}) \rightarrow \mathcal{P}(\mathbb{Y})$ such that $\kappa_x := \mathcal{K}(x, \gamma)$ provides a disintegration of γ with respect to its first marginal, i.e.

$$\gamma = \mu \otimes \kappa, \quad \mu := \pi_{\mathbb{X}}^1 \gamma, \quad \kappa_x := \mathcal{K}(x, \gamma) \quad x \in \mathbb{X}, \quad \text{for every } \gamma \in \mathcal{P}(\mathbb{X} \times \mathbb{Y}). \quad (2.2)$$

We will denote by $\mathcal{P}^{\text{det}}(\mathbb{X} \times \mathbb{Y})$ the set of deterministic couplings:

$$\mathcal{P}^{\text{det}}(\mathbb{X} \times \mathbb{Y}) := \left\{ \gamma \in \mathcal{P}(\mathbb{X} \times \mathbb{Y}) : \exists f : \mathbb{X} \rightarrow \mathbb{Y} \text{ Borel, such that } \gamma = (\text{Id} \times f)_{\#}\mu, \mu = \pi_{\mathbb{X}}^1 \gamma \right\}. \quad (2.3)$$

Since disintegrations are uniquely defined almost everywhere with respect to the marginal, we also have

$$\begin{aligned} \gamma \in \mathcal{P}^{\text{det}}(\mathbf{X} \times \mathbf{Y}) \quad \text{if and only if} \quad \mathcal{K}(\cdot, \gamma) \in \mathcal{P}_\delta(\mathbf{Y}) \quad \pi_{\sharp}^1 \gamma\text{-a.e.} \\ \text{where} \quad \mathcal{P}_\delta(\mathbf{Y}) := \left\{ \delta_y : y \in \mathbf{Y} \right\}. \end{aligned} \quad (2.4)$$

It is possible to prove [LS25] that $\mathcal{P}^{\text{det}}(\mathbf{X} \times \mathbf{Y})$ is a G_δ (thus Borel) subset of $\mathcal{P}(\mathbf{X} \times \mathbf{Y})$.

Given $\mu_i \in \mathcal{P}(\mathbf{X}^i)$ we will denote by $\Gamma(\mu_1, \dots, \mu_N)$ the subset of $\mathcal{P}(\mathbf{X})$ whose elements $\boldsymbol{\mu}$ satisfy $\pi_{\sharp}^i(\boldsymbol{\mu}) = \mu_i$, $i = 1, \dots, N$.

We will often use the following consequence of Von Neumann selection Theorem [Sch73, Theorem 13, pag. 127]:

Theorem 2.1. *Let \mathbf{X}, \mathbf{Y} be Polish spaces, $\nu \in \mathcal{P}(\mathbf{Y})$ concentrated on the Borel set $B \subset \mathbf{Y}$ and let $f : \mathbf{X} \rightarrow \mathbf{Y}$ be a Borel map. If $f(\mathbf{X}) \supset B$ then there exists a measure $\mu \in \mathcal{P}(\mathbf{X})$ such that*

$$\mu \text{ is concentrated on } f^{-1}(B), \quad f_{\sharp} \mu = \nu. \quad (2.5)$$

2.1 The L^2 -Wasserstein space

Let $(\mathbf{X}, d_{\mathbf{X}})$ be a complete and separable metric space and let x_o a point of \mathbf{X} . We denote by $\mathcal{P}_2(\mathbf{X})$ the space of Borel probability measures on \mathbf{X} with finite quadratic moment

$$\int_{\mathbf{X}} d_{\mathbf{X}}^2(x, x_o) d\mu(x) < \infty. \quad (2.6)$$

It is easy to check that the definition is independent of the choice of the reference point x_o . The space $\mathcal{P}_2(\mathbf{X})$ can be endowed with the L^2 -Kantorovich-Wasserstein metric $W_{2, d_{\mathbf{X}}}$

$$W_{2, d_{\mathbf{X}}}^2(\mu_1, \mu_2) := \min \left\{ \int d_{\mathbf{X}}^2(x_0, x_1) d\boldsymbol{\mu}(x_0, x_1) : \boldsymbol{\mu} \in \Gamma(\mu_1, \mu_2) \right\}; \quad (2.7)$$

we will denote by $\Gamma_o(\mu_1, \mu_2)$ the (compact, non-empty) subset of $\Gamma(\mu_1, \mu_2)$ where the minimum is attained. We will also denote by $\mathcal{P}_{2, o}(\mathbf{X} \times \mathbf{X})$ the set of couplings $\boldsymbol{\mu} \in \mathcal{P}_2(\mathbf{X} \times \mathbf{X})$ such that $\boldsymbol{\mu} \in \Gamma_o(\pi_{\sharp}^1 \boldsymbol{\mu}, \pi_{\sharp}^2 \boldsymbol{\mu})$.

$(\mathcal{P}_2(\mathbf{X}), W_{2, d_{\mathbf{X}}})$ is a complete and separable metric space as well [AGS08, Chap. 7]. In this paper we will mainly consider three important cases, when $\mathbf{X} = \mathbf{H}$ is a Hilbert space, when $\mathbf{X} = L^2(\mathbf{Q}, \mathbb{M}; \mathbf{H})$ is the space of \mathbf{H} -valued (Bochner) square-integrable Lagrangian maps defined in some probability space (\mathbf{Q}, \mathbb{M}) , and when $\mathbf{X} = \mathcal{P}_2(\mathbf{H})$ is the Wasserstein space itself.

The Hilbertian case $\mathcal{P}_2(\mathbf{H})$ and the maximal correlation pairing. In this paper we will denote by \mathbf{H} a separable Hilbert space with scalar product $\langle \cdot, \cdot \rangle$, norm $|\cdot|$, and induced metric $d_{\mathbf{H}}$. The finite dimensional case $\mathbf{H} = \mathbb{R}^d$ provides an important example covered by the theory.

The previous construction applied to $(\mathbf{H}, d_{\mathbf{H}})$ yields the space $\mathcal{P}_2(\mathbf{H})$; the quadratic moment of $\mu \in \mathcal{P}_2(\mathbf{H})$ is

$$m_2^2(\mu) := \int_{\mathbf{H}} |x|^2 d\mu(x) = \int_{\mathbf{H}} d_{\mathbf{H}}^2(x, 0) d\mu(x), \quad \text{corresponding to the choice of } x_o := 0. \quad (2.8)$$

To simplify notation, we will simply denote the Wasserstein metric $W_{2, d_{\mathbf{H}}}$ by w_2 :

$$w_2^2(\mu_1, \mu_2) := \min \left\{ \int_{\mathbf{H} \times \mathbf{H}} |x_0 - x_1|^2 d\boldsymbol{\mu}(x_0, x_1) : \boldsymbol{\mu} \in \Gamma(\mu_1, \mu_2) \right\}. \quad (2.9)$$

The Euclidean structure of $d_{\mathbf{H}}$ allows for a useful decomposition of w_2 .

Definition 2.2 (The maximal correlation pairing). *For every $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbf{H})$ we set*

$$[\mu_1, \mu_2] := \max \left\{ \int_{\mathbf{H} \times \mathbf{H}} \langle x_0, x_1 \rangle d\boldsymbol{\mu} : \boldsymbol{\mu} \in \Gamma(\mu_1, \mu_2) \right\}. \quad (2.10)$$

It is clear that $[\cdot, \cdot]$ is finite in $\mathcal{P}_2(\mathbb{H})$ and satisfies

$$|[\mu_1, \mu_2]| \leq m_2(\mu_1) m_2(\mu_2), \quad (2.11)$$

$$w_2^2(\mu_1, \mu_2) = m_2^2(\mu_1) + m_2^2(\mu_2) - 2[\mu_1, \mu_2] \quad \text{for every } \mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{H}). \quad (2.12)$$

In particular, a coupling μ belongs to $\Gamma_o(\mu_1, \mu_2)$ if and only if it attains the maximum in (2.10).

The space $\mathfrak{P}_2(\mathbb{H}) = \mathcal{P}_2(\mathcal{P}_2(\mathbb{H}))$ Since $(\mathcal{P}_2(\mathbb{H}), w_2)$ is a complete and separable metric space, we can iterate the Wasserstein construction and consider the space $\mathcal{P}_2(\mathcal{P}_2(\mathbb{H}))$, which we will also denote by $\mathfrak{P}_2(\mathbb{H})$; its elements (also called laws of random measures) will be denoted by capital letters $\mathbf{M}, \mathbf{N}, \dots$.

Using $\delta_0 \in \mathcal{P}_2(\mathbb{H})$ as a reference measure and observing that $w_2^2(\mu, \delta_0) = m_2^2(\mu)$, the quadratic moment in $\mathfrak{P}_2(\mathbb{H})$ is

$$\mathbf{M}_2^2(\mathbf{M}) := \int_{\mathcal{P}_2(\mathbb{H})} m_2^2(\mu) d\mathbf{M}(\mu) = \int_{\mathcal{P}_2(\mathbb{H})} \int_{\mathbb{H}} |x|^2 d\mu d\mathbf{M}(\mu). \quad (2.13)$$

The corresponding Wasserstein metric W_{2, w_2} will be denoted by W_2 :

$$W_2^2(\mathbf{M}_0, \mathbf{M}_1) := \min \left\{ \int_{\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H})} w_2^2(\mu_1, \mu_2) d\Pi(\mu_1, \mu_2) : \Pi \in \Gamma(\mathbf{M}_0, \mathbf{M}_1) \right\}. \quad (2.14)$$

2.2 Regularity of measures and Monge formulation in $\mathcal{P}_2(\mathbb{H})$

A particularly relevant case when the optimal coupling μ solving (2.9) or (2.10) is unique and deterministic according to (2.3), is related to the regularity of (at least one of) the marginals μ_i . The most refined description of such a regularity is characterized by the vanishing of the marginal μ_i on all the so-called d.c. (or δ -convex or cc) hypersurfaces, i.e. graphs of difference of convex Lipschitz functions in a suitable coordinate system. More precisely, we recall that a subset $S \subset \mathbb{H}$ is a d.c. hypersurface if we can find a decomposition of \mathbb{H} in a orthogonal sum $\mathbb{H} = \mathbb{E} \oplus v\mathbb{R}$, where $v \in \mathbb{H}$ and \mathbb{E} is the hyperplane of \mathbb{H} orthogonal to v , and two Lipschitz convex functions $f, g : \mathbb{E} \rightarrow \mathbb{R}$ such that

$$S = \left\{ x + (f(x) - g(x))v : x \in \mathbb{E} \right\}. \quad (2.15)$$

When $\mathbb{H} = \mathbb{R}$, d.c. hypersurfaces just reduce to a point. We say that

$B \subset \mathbb{H}$ is σ -d.c. hypersurface if it can be covered by a countable union of d.c. hypersurfaces. (2.16)

A larger class of negligible subsets is provided by the so-called Gaussian-null sets introduced by Phelps in [Phe78]: a Borel subset B of \mathbb{H} is Gaussian-null, if $\mathbf{g}(B) = 0$ for every non-degenerate Gaussian measure $\mathbf{g} \in \mathcal{P}(\mathbb{H})$ (see Section 2.4 below). Thanks to [Csö99] the class of Gaussian null Borel sets coincides with the σ -ideals of cube null Borel sets and of Aronsajn null Borel sets [Aro76] (see also [Bog84] and the other comparisons and extensions discussed in [Bog18]). Since every d.c. hypersurface is Aronsajn null, it is immediate to check that σ -d.c. hypersurfaces sets are Gaussian-null.

Definition 2.3 (Atomless, regular, and G-regular measures). *We denote by $\mathcal{P}_2^{al}(\mathbb{H})$ the collection of atomless probability measures in $\mathcal{P}_2(\mathbb{H})$, thus satisfying $\mu(\{x\}) = 0$ for every $x \in \mathbb{H}$.*

We denote by $\mathcal{P}_2^r(\mathbb{H})$ the class of regular probability measures that vanish on all d.c. hypersurfaces (and therefore on all σ -d.c. hypersurfaces).

We denote by $\mathcal{P}_2^{gr}(\mathbb{H})$ the class of G-regular probability measures, that vanish on all Gaussian null Borel subsets of \mathbb{H} .

Remark 2.4. *It is clear that $\mathcal{P}_2^{gr}(\mathbf{H}) \subset \mathcal{P}_2^{al}(\mathbf{H})$. It is useful to recall other important relations and some particular cases covered by the above definitions.*

1. *When $\mathbf{H} = \mathbb{R}$ is one-dimensional, the class of atomless and regular measures coincide, i.e. $\mathcal{P}_2^r(\mathbb{R}) = \mathcal{P}_2^{al}(\mathbb{R})$.*
2. *When \mathbf{H} has finite dimension d , $\mu \in \mathcal{P}_2^{gr}(\mathbf{H})$ if and only if it is absolutely continuous with respect to the d -dimensional Lebesgue measure.*
3. *Since d.c. hypersurfaces are Gaussian null, it is immediate to check that the class of G -regular measures $\mathcal{P}_2^{gr}(\mathbf{H})$ is included in $\mathcal{P}_2^r(\mathbf{H})$. Therefore we have*

$$\mathcal{P}_2^{gr}(\mathbf{H}) \subset \mathcal{P}_2^r(\mathbf{H}) \subset \mathcal{P}_2^{al}(\mathbf{H}), \quad \mathcal{P}_2^r(\mathbb{R}) = \mathcal{P}_2^{al}(\mathbb{R}). \quad (2.17)$$

The crucial property of regular measures $\mu \in \mathcal{P}_2^r(\mathbf{H})$ is that every convex and Lipschitz function $\varphi : \mathbf{H} \rightarrow \mathbb{R}$ is Gateaux-differentiable μ -almost everywhere. This follows by a nice result of Zajíček [Zaj79] (see also [BL00, Theorem 4.20]) showing that the set of points where $\varphi : \mathbf{H} \rightarrow \mathbb{R}$ is not Gateaux-differentiable is contained in an σ -d.c. hypersurface according to (2.16).

Using this property and the structure of optimal couplings, it is possible to prove the celebrated Brénier Theorem.

Theorem 2.5. *If $\mu_1 \in \mathcal{P}_2^r(\mathbf{H}), \mu_2 \in \mathcal{P}_2(\mathbf{H})$ then $\Gamma_o(\mu_1, \mu_2)$ contains a unique element γ which is deterministic, i.e. concentrated on the graph of a Borel map $f \in L^2(\mathbf{H}, \mu_1; \mathbf{H})$ with $f_{\#}\mu_1 = \mu_2$.*

We will also be interested to measurability properties of the above sets. We collect them in the following statement, in the case when \mathbf{H} has finite dimension.

Proposition 2.6 (Measurability of $\mathcal{P}_2^r(\mathbf{H})$ and $\mathcal{P}_2^{gr}(\mathbf{H})$). *Let us assume that \mathbf{H} has finite dimension d .*

1. *$\mathcal{P}_2^r(\mathbf{H})$ is a G_δ (thus Borel) subset of $\mathcal{P}_2(\mathbf{H})$.*
2. *$\mathcal{P}_2^{gr}(\mathbf{H})$ is a Borel subset of $\mathcal{P}_2(\mathbf{H})$.*

Proof. Claim 1 ($d = 1, \mathbf{H} = \mathbb{R}$). It is sufficient to note that $\mu \in \mathcal{P}_2^r(\mathbb{R})$ if and only if $\mu \times \mu(D) = 0$, where $D := \{(x, x) : x \in \mathbb{R}\}$ is the (closed) diagonal of \mathbb{R}^2 . Since the map $\mu \mapsto \mu \times \mu(D)$ is upper semicontinuous, we have $\mathcal{P}_2^r(\mathbb{R}) = \bigcap_{k \in \mathbb{N}_+} \{\mu \in \mathcal{P}_2(\mathbb{R}) : \mu \times \mu(D) > 1/k\}$.

Claim 1 ($\mathbf{H} = \mathbb{R}^d, d > 1$). Let \mathcal{B} be the collection of all the compact boxes $B = \prod_{k=1}^{d-1} [a_k, b_k] \subset \mathbb{R}^{d-1}$ with rational endpoints $a_k < b_k$, let \mathcal{R} a dense countable set of rotations R of \mathbb{R}^d . For every $L \in \mathbb{N}_+$ and $B \in \mathcal{B}$ we consider the set

$$C(B, L) := \{f : B \rightarrow \mathbb{R} : f \text{ convex, } L\text{-Lipschitz, } \sup_B |f| \leq L\}.$$

$C(B, L)$ is a compact set of the Polish space $C(B)$. For every $g, h \in C(B, L)$ and $R \in \mathcal{R}$ we consider the compact graph

$$S_{B,R}(g, h) := \{R(y, g(y) - h(y)) : y \in B\} \subset \mathbb{R}^d.$$

Since every d.c.-hypersurface can be covered by a countable collections of sets of the form $S_{B,R}(g, h)$ for g, h in some $C(B, L)$, we have

$$\begin{aligned} \mathcal{P}_2^r(\mathbb{R}^d) &= \bigcap \left\{ G(B, R, L) : B \in \mathcal{B}, R \in \mathcal{R}, L \in \mathbb{N}_+ \right\}, \\ G(B, R, L) &:= \left\{ \mu \in \mathcal{P}_2(\mathbb{R}^d) : \mu(S_{B,R}(g, h)) = 0 \text{ for every } g, h \in C(B, L) \right\}. \end{aligned}$$

It is then sufficient to show that each set $G(B, R, L)$ is a G_δ .

It is not difficult to check that the map $(\mu, g, h) \mapsto \mu(S_{B,R}(g, h))$ is jointly upper semicontinuous in $\mathcal{P}_2(\mathbb{R}^d) \times (C(B, L))^2$ (uniform convergence in $(C(B, L))^2$ implies Hausdorff convergence of the compact graphs $S_{B,R}(\cdot, \cdot)$) so that the function

$$\mu \mapsto U(\mu) := \sup_{g, h \in C(B, L)} \mu(S_{B,R}(g, h))$$

is upper semicontinuous as well, thanks to the compactness of $C(B, L)$. On the other hand

$$G(B, R, L) = \bigcap_{k \in \mathbb{N}_+} \left\{ \mu \in \mathcal{P}_2(\mathbb{R}^d) : U(\mu) < 1/k \right\}$$

so it is the intersection of a countable collection of open sets.

Claim 2. Let \mathbf{g} denote the standard Gaussian measure in \mathbb{R}^d and let C be the F_σ (thus Borel) set

$$C := \left\{ f \in L^1(\mathbb{R}^d, \mathbf{g}) : f \geq 0, \int f(x) d\mathbf{g}(x) = 1, \int |x|^2 f(x) d\mathbf{g}(x) < \infty \right\}$$

of the Polish space $L^1(\mathbb{R}^d; \mathbf{g}_d)$. The map $J : C \rightarrow \mathcal{P}_2(\mathbb{R}^d)$ defined by $J : f \mapsto f\mathbf{g}$, is continuous and injective so that its image $J(C) = \mathcal{P}_2^*(\mathbb{R}^d)$ is a Borel subset of $\mathcal{P}_2(\mathbb{R}^d)$ by Lusin Theorem. \square

2.3 Lagrangian representations.

Let us consider a standard Borel space (Q, \mathcal{F}_Q) endowed with a reference nonatomic Borel probability measure \mathbb{M} . We denote by \mathcal{H} the Hilbert space $L^2(Q, \mathcal{F}_Q, \mathbb{M}; \mathbb{H})$ endowed with the scalar product and norm

$$\langle X, Y \rangle_{\mathcal{H}} := \int \langle X(q), Y(q) \rangle d\mathbb{M}(q), \quad \|X\|_{\mathcal{H}}^2 := \langle X, X \rangle_{\mathcal{H}}. \quad (2.18)$$

A map $\mathbf{g} : Q \rightarrow Q$ is a *measure preserving isomorphism* (m.p.i.) if it is $\mathcal{F}_Q - \mathcal{F}_Q$ -measurable, there exists a set of full measure $Q_0 \subset Q$ such that $\mathbf{g}|_{Q_0}$ is injective and $\mathbf{g}_\# \mathbb{M} = \mathbb{M}$. We denote by $G(Q)$ the group of measure-preserving isomorphisms. A m.p.i. $\mathbf{g} \in G(Q)$ induces a linear isometry $\mathbf{g}^* : \mathcal{H} \rightarrow \mathcal{H}$ of \mathcal{H} by the map

$$\mathbf{g}^* X := X \circ \mathbf{g}. \quad (2.19)$$

There is a natural 1-Lipschitz surjective map $\iota : \mathcal{H} \rightarrow \mathcal{P}_2(\mathbb{H})$ given by

$$\iota(X) := X_\# \mathbb{M} \quad \text{for every } X \in \mathcal{H}, \quad (2.20)$$

and satisfying

$$\mathfrak{m}_2(\iota(X)) = \|X\|_{\mathcal{H}}, \quad \mathfrak{w}_2(\iota(X_1), \iota(X_2)) \leq \|X_1 - X_2\|_{\mathcal{H}}, \quad \langle X_1, X_2 \rangle_{\mathcal{H}} \leq [\iota(X_1), \iota(X_2)]. \quad (2.21)$$

Notice that

$$\mathbf{g} \in G(Q), \quad Y = X \circ \mathbf{g} \quad \Rightarrow \quad \iota(Y) = \iota(X), \quad (2.22)$$

i.e. $\iota \circ \mathbf{g}^* = \iota$ for every $\mathbf{g} \in G(Q)$. The next Lemma shows that we can considerably refine the previous inequalities. We refer to [CSS25, Sect. 3] for the proof.

Lemma 2.7. *For every $k \in \mathbb{N}$ the map $\iota^k : \mathcal{H}^k \rightarrow \mathcal{P}_2(\mathbb{H}^k)$ defined by*

$$\iota^k(X_1, X_2, \dots, X_k) := (X_1, X_2, \dots, X_k)_\# \mathbb{M} \quad (2.23)$$

is surjective. In particular, for every $\gamma \in \mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H})$ there exists a pair $(X_{\gamma,1}, X_{\gamma,2}) \in \mathcal{H} \times \mathcal{H}$ such that $\iota^2(X_{\gamma,1}, X_{\gamma,2}) = \gamma$ and for every $(X_1, X_2) \in \mathcal{H} \times \mathcal{H}$ we have

$$\iota^2(X_1, X_2) \in \mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H}) \Leftrightarrow \|X_1 - X_2\|_{\mathcal{H}} = \mathfrak{w}_2(\iota(X_1), \iota(X_2)) \Leftrightarrow \langle X_1, X_2 \rangle_{\mathcal{H}} = [\iota(X_1), \iota(X_2)]. \quad (2.24)$$

Moreover, for every $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{H})$ and $X_1, X_2 \in \mathcal{H}$ with $\iota(X_i) = \mu_i$ we have

$$[\mu_1, \mu_2] = \sup_{\mathbf{g} \in \mathbf{G}(\mathbb{Q})} \langle X_1, \mathbf{g}^* X_2 \rangle, \quad w_2^2(\mu_1, \mu_2) = \inf_{\mathbf{g} \in \mathbf{G}(\mathbb{Q})} \|X_1 - \mathbf{g}^* X_2\|_{\mathcal{H}}^2. \quad (2.25)$$

Finally, if $\mu_1 = \mu_2$ (i.e. $\iota(X_1) = \iota(X_2)$) then there exists a sequence $\mathbf{g}_n \in \mathbf{G}(\mathbb{Q})$ such that $\mathbf{g}_n^* X_1 \rightarrow X_2$ strongly in \mathcal{H} .

Every function $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ induces a function $\hat{\phi} := \phi \circ \iota$ on \mathcal{H} which satisfies the obvious law-invariance property

$$\iota(X) = \iota(Y) \quad \Rightarrow \quad \hat{\phi}(X) = \hat{\phi}(Y), \quad (2.26)$$

and, in particular, is invariant by the action of measure-preserving isomorphisms:

$$\hat{\phi}(\mathbf{g}^* X) = \hat{\phi}(X) \quad \text{for every } \mathbf{g} \in \mathbf{G}(\mathbb{Q}). \quad (2.27)$$

Lemma 2.8. *Let $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ and let $\hat{\phi} := \phi \circ \iota : \mathcal{H} \rightarrow \mathbb{R} \cup \{\infty\}$. We have*

1. ϕ is proper if and only if $\hat{\phi}$ is proper.
2. ϕ is l.s.c. if and only if $\hat{\phi}$ is l.s.c.
3. ϕ is continuous if and only if $\hat{\phi}$ is continuous.
4. ϕ is L -Lipschitz if and only if $\hat{\phi}$ is L -Lipschitz.

Moreover, suppose that $\Phi : \mathcal{H} \rightarrow \mathbb{R} \cup \{\infty\}$ is proper, l.s.c. and invariant by measure preserving isomorphisms, i.e. $\Phi \circ \mathbf{g}^* = \Phi$. Then Φ is law-invariant and $\Phi = \phi \circ \iota$ for a (unique) proper, l.s.c. function $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$.

Proof. The first claim and all the left-to-right implications of Claims 2,3,4 are trivial, thanks to (2.21). We thus consider only the right-to-left implications.

Concerning Claim 2, let us suppose that $\hat{\phi}$ is lower semicontinuous and let $(\mu_n)_{n \in \mathbb{N}}$ converging to μ in $\mathcal{P}_2(\mathbb{H})$ with $\phi(\mu_n) \leq c$. By (2.25) there exists a sequence X_n converging to X in \mathcal{H} such that $\mu_n = \iota(X_n)$ and $\mu = \iota(X)$. Since $\hat{\phi}$ is lower semicontinuous we deduce that $\phi(\mu) = \hat{\phi}(X) \leq c$ as well.

Claim 3 immediately follows by Claim 2 applied to ϕ and $-\phi$. Claim 4 follows from (2.24).

Let now Φ be as in the last statement of the Lemma and let $X_i \in \mathcal{H}$ with $\iota(X_1) = \iota(X_2)$. By Lemma 2.7 we can find a sequence $\mathbf{g}_n \in \mathbf{G}(\mathbb{Q})$ such that $\mathbf{g}_n^* X_1 \rightarrow X_2$ as $n \rightarrow \infty$ so that

$$\Phi(X_2) \leq \liminf_{n \rightarrow \infty} \Phi(\mathbf{g}_n^* X_1) = \liminf_{n \rightarrow \infty} \Phi(X_1) = \Phi(X_1).$$

Inverting the role of X_1 and X_2 we deduce that $\Phi(X_1) = \Phi(X_2)$. □

2.4 Lagrangian representation of the laws of random measures of $\mathcal{P}_2(\mathbb{H})$

The 1-Lipschitz and surjective law map $\iota : \mathcal{H} \rightarrow \mathcal{P}_2(\mathbb{H})$ provides a natural way to construct measures in $\mathcal{P}_2(\mathbb{H})$ (the space of laws of random measures in $\mathcal{P}_2(\mathbb{H})$) starting from measures in $\mathcal{P}_2(\mathcal{H})$ (the space of laws of random Lagrangian maps in \mathcal{H}). In fact, the corresponding push-forward transform $\iota_{\#}$ is a surjective map from $\mathcal{P}_2(\mathcal{H})$ to $\mathcal{P}_2(\mathbb{H})$ so that

$$\text{for every } \mathbf{m} \in \mathcal{P}_2(\mathcal{H}) \text{ the push-forward } \mathbf{M} = \iota_{\#} \mathbf{m} \text{ belongs to } \mathcal{P}_2(\mathbb{H}). \quad (2.28)$$

In this way, we will use (maps from) (\mathbb{Q}, \mathbb{M}) as a sort of ‘‘labelling’’ space for measures in \mathbb{H} and not as a source of randomness. The latter can be introduced by using a further measure \mathbf{m} , not in \mathbb{Q} but in the Hilbert space $L^2(\mathbb{Q}, \mathbb{M}; \mathbb{H})$; in turn \mathbf{m} can be represented as the law of a random vector ξ

defined in another (standard Borel) probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Thus in many situations it will be useful to deal with the pair of spaces $(\mathbb{Q}, \mathcal{F}_{\mathbb{Q}}, \mathbb{M})$ and $(\Omega, \mathcal{F}, \mathbb{P})$. Such a construction has been used, with different aims, in various contexts, see e.g. [KR24], [RS09; Stu11].

Let us briefly describe this simple construction: we can assume that

$$\begin{aligned} \text{the measure } \mathbf{m} \in \mathcal{P}_2(\mathcal{H}) \text{ is the law of a } \mathcal{H}\text{-valued random vector } \boldsymbol{\xi} \text{ defined in } (\Omega, \mathcal{F}, \mathbb{P}) : \\ \mathbf{m} = \boldsymbol{\xi}_{\#} \mathbb{P} \quad \text{and} \quad \mathbb{M} = (\iota \circ \boldsymbol{\xi})_{\#} \mathbb{P} \text{ is the law of the random measure } \mu = \boldsymbol{\xi}[\cdot]_{\#} \mathbb{M}. \end{aligned} \quad (2.29)$$

We first represent \mathbf{m} as the distribution of a $\mathcal{F} \otimes \mathcal{F}_{\mathbb{Q}}$ -measurable stochastic process.

Lemma 2.9. *If $\boldsymbol{\xi} : \Omega \rightarrow \mathcal{H}$ is a Borel vector field satisfying (2.29) then there exists a $\mathcal{F} \otimes \mathcal{F}_{\mathbb{Q}}$ -measurable stochastic process $\Xi : \Omega \times \mathbb{Q} \rightarrow \mathbb{H}$ such that*

$$\text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega \quad \Xi(\omega, \cdot) = \boldsymbol{\xi}[\omega] \quad \mathbb{M}\text{-a.e. in } \mathbb{Q}. \quad (2.30)$$

Proof. We select an orthonormal basis $(\mathbf{h}_k)_{k \in \mathbb{N}}$ of \mathbb{H} and set $\xi_k := \langle \boldsymbol{\xi}, \mathbf{h}_k \rangle$. ξ_k is a Borel map from Ω to $L^2(\mathbb{Q}, \mathbb{M})$ satisfying

$$\sum_k \int \|\xi_k[\omega]\|_{L^2(\mathbb{Q}, \mathbb{M})}^2 d\mathbb{P}(\omega) = \int \|\boldsymbol{\xi}[\omega]\|_{\mathcal{H}}^2 d\mathbb{P}(\omega) = \int \|X\|_{\mathcal{H}}^2 d\mathbf{m}(X) < \infty. \quad (2.31)$$

We can then set $\nu_k(\omega, B) := \int \xi_k[\omega] \chi_B d\mathbb{M}$ for every $\omega \in \Omega$ and $B \in \mathcal{F}_{\mathbb{Q}}$ obtaining a family of signed measures $\nu_k(\omega, \cdot)$ on $\mathcal{F}_{\mathbb{Q}}$ depending on ω in a \mathcal{F} -measurable way, with

$$|\nu_k|(\omega, B) \leq \|\xi_k[\omega]\|_{L^2(\mathbb{Q}, \mathbb{M})} \mathbb{M}(B)$$

By the Doob's measurable Radon-Nykodim theorem [DM82, Thm. 58] (see also [Bog07, Ex. 6.10.72]) we can find a $\mathcal{F} \times \mathcal{F}_{\mathbb{Q}}$ -measurable density $f_k : \Omega \times \mathbb{Q} \rightarrow \mathbb{R}$ such that $\nu_k(\omega, \cdot) = f_k(\omega, \cdot) \mathbb{M}$ so that

$$\xi_k[\omega] = f_k(\omega, \cdot) \quad \mathbb{M}\text{-a.e. and} \quad \int f_k^2(\omega, q) d\mathbb{M}(q) = \|\xi_k[\omega]\|_{L^2(\mathbb{Q}, \mathbb{M})}^2 \quad \text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega.$$

Setting

$$\Xi_n(\omega, q) := \sum_{k=1}^n f_k(\omega, q) \mathbf{h}_k$$

and using (2.31) it is not difficult to check that Ξ_n converges pointwise $\mathbb{P} \times \mathbb{M}$ -a.e. to an element $\Xi \in L^2(\Omega \times \mathbb{Q}, \mathbb{P} \times \mathbb{M}; \mathbb{H})$ which satisfies (2.30). \square

We can use Ξ to represent \mathbb{M} and its k -projections: for every $\omega \in \Omega$ the measure $\mu_{\omega} = \iota(\boldsymbol{\xi}(\omega)) \in \mathcal{P}(\mathbb{H})$ satisfies

$$\int_{\mathbb{H}} \zeta(x) d\mu_{\omega}(x) = \int_{\mathbb{Q}} \zeta(\Xi(\omega, q)) d\mathbb{M}(q) \quad (2.32)$$

and we can recover \mathbb{M} as the law with respect to \mathbb{P} of the random measure μ_{ω} :

$$\mathbb{M} = \int_{\Omega} \delta_{\mu_{\omega}} d\mathbb{P}(\omega) = \int_{\Omega} \delta_{\Xi(\omega, \cdot)_{\#} \mathbb{M}} d\mathbb{P}(\omega). \quad (2.33)$$

By using Ξ we can easily express k -projections of \mathbb{M} . For every $k \in \mathbb{N}$ and $\mathbb{M} \in \mathcal{P}_2(\mathbb{H})$ let us define

$$\text{pr}^k[\mathbb{M}] := \int \mu^{\otimes k} d\mathbb{M}(\mu) \in \mathcal{P}_2(\mathbb{H}^k), \quad (2.34)$$

which satisfies

$$\int_{\mathbb{H}^k} \zeta d \text{pr}^k[\mathbb{M}] = \int \left(\int_{\mathbb{H}^k} \zeta d\mu^{\otimes k} \right) d\mathbb{M}(\mu) \quad \text{for every bounded Borel } \zeta : \mathbb{H}^k \rightarrow \mathbb{R}. \quad (2.35)$$

We then have

$$\int_{\mathbb{H}^k} \zeta d \text{pr}^k[\mathbb{M}] = \int_{\Omega} \left(\int_{\mathbb{Q}} \zeta(\Xi(\omega, q_1), \dots, \Xi(\omega, q_k)) d\mathbb{M}^{\otimes k}(q_1, \dots, q_k) \right) d\mathbb{P}(\omega), \quad (2.36)$$

i.e.

$$\text{pr}^k[\mathbb{M}] = \Xi_{\#}^k(\mathbb{P} \otimes \mathbb{M}^{\otimes k}) \quad \text{where} \quad \Xi^k(\omega, q_1, \dots, q_k) := (\Xi(\omega, q_1), \dots, \Xi(\omega, q_k)). \quad (2.37)$$

Example 2.10 (Laws of Gaussian generated random measures (LGRRM)). Let us focus on the particular case of a nondegenerate centered Gaussian measure $\mathbf{g} \sim N(0, K)$ in \mathcal{H} with covariance operator K . By Karhunen-Loève expansion, we can find

1. an orthonormal basis $(\mathbf{E}_n)_{n \in \mathbb{N}_+}$ of \mathcal{H} , given by the eigenvectors of K ;
2. the corresponding sequence of nonnegative eigenvalues of K $(\lambda_n^2)_{n \in \mathbb{N}_+}$ with $K\mathbf{E}_n = \lambda_n^2 \mathbf{E}_n$ and $\Lambda := \sum_n \lambda_n^2 < \infty$
3. a sequence of independent Gaussian random variables $(\xi_n)_{n \in \mathbb{N}_+}$ defined in Ω with $\xi_n \sim N(0, \lambda_n^2)$

such that

$$\mathbf{g} = \xi_{\#} \mathbb{P}, \quad \xi := \sum_n \xi_n \mathbf{E}_n. \quad (2.38)$$

Since $\Xi_n(\omega, q) := \xi_n(\omega) \mathbf{E}_n(q)$ is an orthogonal system in $L^2(\Omega \times \mathbb{Q}, \mathbb{P} \otimes \mathbb{M})$ the series

$$\Xi(\omega, q) := \sum_n \Xi_n(\omega, q) = \sum_n \xi_n(\omega) \mathbf{E}_n(q), \quad \|\Xi_n\|_{L^2(\Omega \times \mathbb{Q}, \mathbb{P} \otimes \mathbb{M})} = \lambda_n \quad (2.39)$$

converges in $L^2(\Omega \times \mathbb{Q}, \mathbb{P} \otimes \mathbb{M})$ and provides a measurable version satisfying (2.30).

Example 2.11. We can also proceed in a slightly different way: assuming that \mathbb{Q} is a compact metrizable space, we can start from a \mathbb{H} -valued measurable process $\Xi_q = \Xi(\cdot, q)$ indexed by $q \in \mathbb{Q}$ with continuous paths, i.e. $q \mapsto \Xi(\omega, q) \in C^0(\mathbb{Q})$ for \mathbb{P} -a.e. ω . In this way the map $\xi : \omega \rightarrow \Xi(\omega, \cdot)$ can be considered as a measurable map from Ω to the Banach space $\mathcal{B} := C^0(\mathbb{Q}; \mathbb{H})$; assuming that $\xi \in L^2(\Omega, \mathbb{P}; \mathcal{B})$, the distribution of the process $\mathbf{m} = \xi_{\#} \mathbb{P}$ is a Radon measure in $\mathcal{P}_2(\mathcal{B})$. Any choice of diffuse measure $\mathbb{M} \in \mathcal{P}(\mathbb{Q})$ induces a push-forward map $\iota_{\mathbb{M}}(X) := X_{\#} \mathbb{M}$ and a measure $\mathbb{M} = (\iota_{\mathbb{M}})_{\#} \mathbf{m}$ as in (2.33). If in particular Ξ is a Gaussian process, then \mathbf{m} is a Gaussian measure in \mathcal{B} and thus also in \mathcal{H} .

3 Totally convex functionals

We first recall the definition of totally convex functionals on $\mathcal{P}_2(\mathbb{H})$ [CSS23a, Sec. 5]

Definition 3.1 (Totally convex functionals). *A functional $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ is totally convex if for every coupling $\mu \in \mathcal{P}_2(\mathbb{H} \times \mathbb{H})$ and $t \in [0, 1]$*

$$\phi(\mu_t) \leq (1-t)\phi(\mu_0) + t\phi(\mu_1) \quad t \in [0, 1], \quad \mu_t = (\pi_t^{1 \rightarrow 2})_{\#} \mu, \quad \pi_t^{1 \rightarrow 2}(x_1, x_2) := (1-t)x_1 + tx_2. \quad (3.1)$$

Equivalently, the lifted functional $\hat{\phi} := \phi \circ \iota$ is convex in \mathcal{H} .

We say that ϕ is totally λ -convex if $\phi - \frac{\lambda}{2} \mathbf{m}_2^2$ is totally convex (equivalently $\hat{\phi}$ is λ -convex in \mathcal{H}).

Notice that the lifting $\phi \rightarrow \hat{\phi}$ inherits also properness and lower semicontinuity. Conversely, if $\Phi : \mathcal{H} \rightarrow \mathbb{R} \cup \{\infty\}$ is convex, lower semicontinuous and invariant by m.p.i. then there exists a unique totally convex functional ϕ such that $\Phi = \hat{\phi} = \phi \circ \iota$.

If we restrict the convexity inequality (3.1) to displacement interpolation induced by *optimal couplings* μ , we obtain a larger class of functionals, which are called *geodesically (or displacement) convex*, according to [McC97]. However if $\dim(\mathbb{H}) \geq 2$, every *continuous* geodesically convex functionals ϕ is also totally convex [CSS23a, Thm. 5.5]. We quote here a few important examples:

Example 3.2 (Potential energy). If $f : \mathbb{H} \rightarrow \mathbb{R} \cup \{\infty\}$ is λ -convex (proper, l.s.c.), then

$$V_f : \mu \mapsto \int_{\mathbb{H}} f(x) d\mu(x)$$

is totally λ -convex (proper, l.s.c.) in $\mathcal{P}_2(\mathbb{H})$. In particular, \mathbf{m}_2^2 (corresponding to $f(x) = |x|^2$) is totally 1-convex.

Example 3.3 (Multiple interaction energy). If $g : \mathbb{H}^k \rightarrow \mathbb{R} \cup \{\infty\}$ is proper, convex, l.s.c., and invariant respect to arbitrary permutations of its entries, then

$$W_g : \mu \mapsto \int g(x_1, x_2, \dots, x_k) d\mu^{\otimes k}(x_1, \dots, x_k)$$

is totally convex (proper and l.s.c.) in $\mathcal{P}_2(\mathbb{H})$.

Example 3.4. For every $\nu \in \mathcal{P}_2(\mathbb{H})$ we define the maximal pairing functional $k_\nu : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R}$ defined by

$$k_\nu(\mu) := [\mu, \nu] \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{H}). \quad (3.2)$$

Proposition 3.5. *For every $\nu \in \mathcal{P}_2(\mathbb{H})$ the function k_ν is $m_2(\nu)$ -Lipschitz in $\mathcal{P}_2(\mathbb{H})$ and totally convex.*

Proof. We give two proofs. The first one is direct: let us fix $\nu, \mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{H})$ and let $\boldsymbol{\mu}^{12} \in \Gamma(\mu_1, \mu_2)$. Let us set

$$\pi_t^{1 \rightarrow 2} : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H}, \quad \pi_t^{1 \rightarrow 2}(x_1, x_2) := (1-t)x_1 + tx_2, \quad t \in [0, 1], \quad (3.3)$$

and $\mu_t^{1 \rightarrow 2} := (\pi_t^{1 \rightarrow 2})\# \boldsymbol{\mu}^{12}$. We have to prove that

$$k_\nu(\mu_t^{1 \rightarrow 2}) \leq (1-t)k_\nu(\mu_1) + tk_\nu(\mu_2) \quad \text{for every } t \in [0, 1]. \quad (3.4)$$

We select $\boldsymbol{\mu}_t \in \Gamma_o(\mu_t, \nu)$ and we apply [AGS08, Prop. 7.3.1] to find $\boldsymbol{\mu} \in \Gamma(\mu_1, \mu_2, \nu) \subset \mathcal{P}_2(\mathbb{H} \times \mathbb{H} \times \mathbb{H})$ such that $\pi_t^{12}\boldsymbol{\mu} = \boldsymbol{\mu}^{12}$ and $(\pi_t^{1 \rightarrow 2}, \pi^3)\# \boldsymbol{\mu} = \boldsymbol{\mu}_t$. It follows that

$$\begin{aligned} k_\nu(\mu_t^{1 \rightarrow 2}) &= \int \langle x, z \rangle d\boldsymbol{\mu}_t(x, z) = \int \langle \pi_t^{1 \rightarrow 2}(x_1, x_2), x_3 \rangle d\boldsymbol{\mu}(x_1, x_2, x_3) \\ &= \int \langle (1-t)x_1 + tx_2, x_3 \rangle d\boldsymbol{\mu}(x_1, x_2, x_3) \\ &= (1-t) \int \langle x_1, x_3 \rangle d\boldsymbol{\mu}(x_1, x_2, x_3) + t \int \langle x_2, x_3 \rangle d\boldsymbol{\mu}(x_1, x_2, x_3) \\ &\leq (1-t)k_\nu(\mu_1) + tk_\nu(\mu_2). \end{aligned}$$

The second argument uses the representation result (2.25); we fix $Y \in \mathcal{H}$ such that $\iota(Y) = \nu$ and observe that

$$\hat{k}_\nu(X) = k_\nu(\iota(X)) = \sup \left\{ \langle X, Y \circ g \rangle : g \in G(Q) \right\}, \quad (3.5)$$

so that \hat{k}_ν is the supremum of a family of L -Lipschitz (with $L = \|Y\|_{\mathcal{H}} = m_2(\nu)$) and linear functionals: therefore, it is convex and L -Lipschitz as well. \square

3.1 A Kantorovich version of the Legendre-Fenchel transform in $\mathcal{P}_2(\mathbb{H})$.

Recall that the Legendre-Fenchel transform of a function $\Phi : \mathcal{H} \rightarrow \mathbb{R} \cup \{\infty\}$ is defined as

$$\Phi^*(Y) := \sup_{X \in \mathcal{H}} \langle Y, X \rangle_{\mathcal{H}} - \Phi(X). \quad (3.6)$$

Inspired by this formula we define a corresponding transformation for functionals on $\mathcal{P}_2(\mathbb{H})$.

Definition 3.6 (Kantorovich-Legendre-Fenchel transformation in $\mathcal{P}_2(\mathbb{H})$). *Let $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow (-\infty, +\infty]$ be a proper function. We call Kantorovich-Legendre-Fenchel conjugate of ϕ the function $\phi^* : \mathcal{P}_2(\mathbb{H}) \rightarrow (-\infty, +\infty]$ defined by*

$$\phi^*(\nu) := \sup_{\mu \in \mathcal{P}_2(\mathbb{H})} [\nu, \mu] - \phi(\mu). \quad (3.7)$$

It is clear that ϕ^* is totally convex and lower semicontinuous, as the supremum of totally convex and continuous functions. It is also proper (i.e. not identically $= \infty$) if ϕ satisfies the lower bound

$$\phi(\mu) \geq -a + [\mu, \nu] \quad \text{for every } \mu \in \mathcal{P}_2(\mathbb{H}) \quad (3.8)$$

for some $a \in \mathbb{R}$ and $\nu \in \mathcal{P}_2(\mathbb{H})$.

Every geodesically convex function $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ is linearly bounded from below [MS20], in the sense that there exist constants $a, b \geq 0$ such that

$$\phi(\mu) \geq -a - b m_2(\mu) \quad (3.9)$$

When ϕ is totally convex, we immediately get a refined lower bound.

Lemma 3.7. *If $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ is totally convex and lower semicontinuous then there exist $a \in \mathbb{R}$ and $\nu \in \mathcal{P}_2(\mathbb{H})$ such that (3.8) holds. In particular $\phi^*(\nu) \leq a$ so that ϕ^* is proper.*

Proof. It is sufficient to observe that $\hat{\phi} := \phi \circ \iota$ is convex and lower semicontinuous in \mathcal{H} so that there exist $a \geq 0$ and $Y \in \mathcal{H}$ such that

$$\hat{\phi}(X) \geq -a - \langle X, Y \rangle \quad \text{for every } X \in \mathcal{H}.$$

Setting $\nu := \iota(Y) \in \mathcal{P}_2(\mathbb{H})$ and using the fact that $\hat{\phi}$ is invariant w.r.t. measure-preserving isomorphisms, we get for every $\mu = \iota(X) \in \mathcal{P}_2(\mathbb{H})$ and $\mathbf{g} \in \mathbf{G}(\mathbb{H})$

$$\phi(\mu) = \hat{\phi}(X) = \hat{\phi}(X \circ \mathbf{g}^{-1}) \geq -a + \langle X \circ \mathbf{g}^{-1}, Y \rangle = -a + \langle X, Y \circ \mathbf{g} \rangle.$$

Taking the supremum w.r.t. $\mathbf{g} \in \mathbf{G}(\mathbb{H})$ and recalling (2.21) and (2.25) we get

$$\phi(\mu) \geq -a + \sup_{\mathbf{g} \in \mathbf{G}(\mathbb{H})} \langle X, Y \circ \mathbf{g} \rangle = -a + [\nu, \mu].$$

□

We collect a few simple but relevant properties of the Kantorovich-Legendre-Fenchel transform in the following Theorem.

Theorem 3.8. *Let $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ be a proper function satisfying (3.8) for some $a \in \mathbb{R}$ and $\nu \in \mathcal{P}_2(\mathbb{H})$.*

1. *The function ϕ^* is proper, totally convex and lower semicontinuous.*
2. *ϕ^* satisfies the commutation property*

$$(\phi^*) \circ \iota = (\phi \circ \iota)^* \quad (3.10)$$

and it is the unique function satisfying (3.10).

3. *The function $\phi^{**} = (\phi^*)^*$ is the largest totally convex and lower semicontinuous function dominated by ϕ .*

Proof. Claim 1 is a direct consequence of Lemma 3.7 (for properness) and Proposition 3.5 (for convexity and lower semicontinuity).

In order to check (3.10) of Claim 2, we consider $Y \in \mathcal{H}$ and $\nu = \iota(Y)$; (2.25) yields

$$\begin{aligned} \phi^*(\nu) &= \sup_{\mu \in \mathcal{H}_2(\mathbb{H})} [\nu, \mu] - \phi(\mu) = \sup_{X \in \mathcal{H}_2(\mathbb{H})} [\nu, \iota(X)] - \phi(\iota(X)) = \sup_{\mathbf{g} \in \mathbf{G}(\mathbb{Q})} \sup_{X \in \mathcal{H}_2(\mathbb{H})} \langle Y, X \circ \mathbf{g} \rangle - \phi(\iota(X)) \\ &= \sup_{\mathbf{g} \in \mathbf{G}(\mathbb{Q})} \sup_{X \in \mathcal{H}_2(\mathbb{H})} \langle Y, X \circ \mathbf{g} \rangle - \phi(\iota(X \circ \mathbf{g})) = \sup_{X' \in \mathcal{H}_2(\mathbb{H})} \langle Y, X' \rangle - \phi(\iota(X')) = (\phi \circ \iota)^*(Y). \end{aligned}$$

The uniqueness follows by the law invariance of $(\phi \circ \iota)^*$. Claim 3 is then an obvious consequence of Claim 2 and the corresponding property for the Legendre-Fenchel transformation in \mathcal{H} . □

We can easily derive natural properties from the Hilbertian theory.

Corollary 3.9 (Kantorovich-Fenchel inequality). *For every $\gamma \in \Gamma(\mu, \nu)$ we have*

$$\int_{\mathbb{H}^2} \langle x, y \rangle d\gamma(x, y) \leq [\mu, \nu] \leq \phi(\mu) + \phi^*(\nu) \quad (3.11)$$

$$\phi^*(\nu) = \sup \left\{ \int_{\mathbb{H}^2} \langle x, y \rangle d\gamma(x, y) - \phi(\pi_{\#}^1 \gamma) : \gamma \in \mathcal{P}_2(\mathbb{H} \times \mathbb{H}), \pi_{\#}^2 \gamma = \nu \right\}. \quad (3.12)$$

Moreover

$$\int_{\mathbb{H}^2} \langle x, y \rangle d\gamma(x, y) = \phi(\mu) + \phi^*(\nu) \quad \Leftrightarrow \quad \begin{cases} \phi(\mu) + \phi^*(\nu) = [\mu, \nu], \\ \gamma \in \Gamma_o(\mu, \nu). \end{cases} \quad (3.13)$$

Corollary 3.10. *Let $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ be a proper function. The following properties are equivalent:*

1. ϕ is totally convex and lower semicontinuous;
2. $\phi \circ \iota$ is convex and lower semicontinuous in \mathcal{H} ;
3. $\phi = \phi^{**}$.
4. ϕ is $[\cdot, \cdot]$ -convex, i.e. there exists a set $G \subset \mathcal{P}_2(\mathbb{H}) \times \mathbb{R}$ such that

$$\phi(\mu) = \sup_{(\nu, a) \in G} [\mu, \nu] - a. \quad (3.14)$$

Proof. The implication $1 \Leftrightarrow 2$ is a consequence of the definition of total convexity and Lemma 2.8.

$1 \Leftrightarrow 3$ is a consequence of Claim 3 of Theorem 3.8. $3 \Rightarrow 4$ is also immediate by the definition of $(\phi^*)^*$; the converse implication is a consequence of the fact that the map $\mu \mapsto [\mu, \nu]$ is totally convex and continuous. \square

Thanks to the commutation identity (3.10), it is possible to derive interesting calculus rules for ϕ^* from the corresponding formulae for $\hat{\phi}^*$. We present two simple examples that will turn to be useful in the following. Let us first introduce the dilation maps and the Moreau-Yosida regularizations of ϕ .

Definition 3.11. *For every $a > 0$ we set $\mathfrak{d}_a : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathcal{P}_2(\mathbb{H})$*

$$\mathfrak{d}_a[\mu] := (a\mathbf{i})_{\#}\mu, \quad \mathfrak{d}_a[\mu](B) = \mu(a^{-1}B) \quad \text{for every Borel set } B \subset \mathbb{H}. \quad (3.15)$$

For every $\tau > 0$ and every proper l.s.c. and totally convex function $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ the τ -Moreau-Yosida regularization of ϕ is defined by

$$\phi_{\tau}(\mu) := \min_{\nu \in \mathcal{P}_2(\mathbb{H})} \frac{1}{2\tau} w_2^2(\mu, \nu) + \phi(\nu). \quad (3.16)$$

Dilations are clearly related to scalar multiplication of Lagrangian maps via the formula

$$\mu = \iota(X) \quad \Rightarrow \quad \mathfrak{d}_a[\mu] = \iota(aX). \quad (3.17)$$

The Yosida regularization (3.16) plays a crucial role in evolution problems via the JKO-Minimizing Movement scheme [JKO98; AGS08]. Existence of a minimizer of (3.16) for an arbitrary geodesically convex functional follows by the results of [NS21]; in the present case, we can invoke the following important commutation property

$$\widehat{\phi}_{\tau} = \hat{\phi}_{\tau} \quad (3.18)$$

where $\hat{\phi}_\tau$ is the Yosida regularization of $\hat{\phi}$ in \mathcal{H} defined by

$$\hat{\phi}_\tau(X) = \min_{Y \in \mathcal{H}} \frac{1}{2\tau} \|X - Y\|_{\mathcal{H}}^2 + \hat{\phi}(Y), \quad (3.19)$$

which also shows that ϕ_τ is totally convex as well. In order to check (3.18) we first observe that (2.21) yields

$$\mu = \iota(X) \quad \Rightarrow \quad \hat{\phi}_\tau(X) \geq \phi_\tau(\mu);$$

the converse inequality follows by the law invariance of $\hat{\phi}$ and (2.25).

Corollary 3.12. *Let $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ be a proper l.s.c. and totally convex function. For every $a, b > 0$ we have*

$$(a\phi)^* = a\phi^* \circ \mathfrak{d}_{a^{-1}}, \quad (3.20)$$

$$\left(a\phi + \frac{b}{2}\mathfrak{m}_2^2\right)^* = a(\phi^*)_{b/a} \circ \mathfrak{d}_{a^{-1}} \quad (3.21)$$

Proof. (3.20) follows from the corresponding identity for $\hat{\phi}$:

$$(a\hat{\phi})^*(X) = a\hat{\phi}^*(a^{-1}X).$$

(3.21) follows by the fact that

$$\left(a\phi + \frac{b}{2}\mathfrak{m}_2^2\right)^\wedge = a\hat{\phi} + \frac{b}{2}\|\cdot\|_{\mathcal{H}}^2$$

and by the corresponding formula for the Legendre transform of a quadratic perturbation of $\hat{\phi}$ in \mathcal{H} (which is a particular case of infimal convolution):

$$\left(a\hat{\phi} + \frac{b}{2}\tau\|\cdot\|_{\mathcal{H}}^2\right)^* = a(\hat{\phi}^*)_{b/a}(\cdot/a). \quad (3.22)$$

□

3.2 c-concave functions and their c-super differentials

As in the case of functions defined in a Hilbert space, we can connect the notion of totally convex functionals and Kantorovich-Legendre-Fenchel transform in $\mathcal{P}_2(\mathbb{H})$ with the corresponding notion of c-concavity and c-transform for the cost $\mathfrak{c} := \frac{1}{2}\mathfrak{w}_2^2$.

Let us first recall the main definition for a continuous and symmetric cost function $\mathfrak{c} : \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R}$: in our case we will mainly use the choice $\mathfrak{c} := \frac{1}{2}\mathfrak{w}_2^2$.

Definition 3.13 (Concave c-transform, c-concavity and c-superdifferential). *Let $f : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{-\infty\}$ be a proper function. Its concave c-transform $f^c : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{-\infty\}$ is defined by*

$$f^c(\nu) := \inf_{\mu \in \mathcal{P}_2(\mathbb{H})} \mathfrak{c}(\mu, \nu) - f(\mu). \quad (3.23)$$

f is c-concave if

$$\exists A \subset \mathcal{P}_2(\mathbb{H}) \times \mathbb{R} : \quad f(\mu) = \inf_{(\nu, a) \in A} \mathfrak{c}(\mu, \nu) - a. \quad (3.24)$$

If $\mu, \nu \in \mathcal{P}_2(\mathbb{H})$ with $f(\mu) \in \mathbb{R}$, we say that ν belongs to the c-superdifferential of f at μ , denoted by $\partial_c^+ f(\mu)$, if

$$f(\mu') - f(\mu) \leq \mathfrak{c}(\mu', \nu) - \mathfrak{c}(\mu, \nu) \quad \text{for every } \mu' \in \mathcal{P}_2(\mathbb{H}). \quad (3.25)$$

We collect a series of well known properties of the above notion.

Theorem 3.14. *Let $f : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{-\infty\}$ be a proper function.*

1. f is \mathbf{c} -concave iff $f = g^{\mathbf{c}}$ for some function $g : \mathcal{P}_2(\mathbf{H}) \rightarrow \mathbb{R} \cup \{-\infty\}$,
2. f is \mathbf{c} -concave if and only if $f = f^{\mathbf{c}\mathbf{c}} = (f^{\mathbf{c}})^{\mathbf{c}}$.
3. for every $\mu, \nu \in \mathcal{P}_2(\mathbf{H})$, $f(\mu) + f^{\mathbf{c}}(\nu) \leq \mathbf{c}(\mu, \nu)$
4. $\nu \in \partial_{\mathbf{c}}^+ f(\mu)$ if and only if $f(\mu) + f^{\mathbf{c}}(\nu) = \mathbf{c}(\mu, \nu)$.

We could introduce the analogous concepts of \mathbf{c} -convexity, \mathbf{c} -convex transform and \mathbf{c} -subdifferential: since in this context we will only use the pseudo-scalar cost $[\cdot, \cdot]$, $[\cdot, \cdot]$ -convexity is equivalent to total convexity by Corollary (3.10), so that we will keep the notation ϕ^* of Definition 3.6 for the $[\cdot, \cdot]$ -conjugate. We will just introduce the notion of $[\cdot, \cdot]$ -subdifferential: for every $\mu \in \mathcal{P}_2(\mathbf{H})$ with $\phi(\mu) \in \mathbb{R}$, we also set

$$\partial^- \phi(\mu) := \left\{ \nu \in \mathcal{P}_2(\mathbf{H}) : \phi(\mu') - \phi(\mu) \geq [\mu', \nu] - [\mu, \nu] \text{ for every } \mu' \in \mathcal{P}_2(\mathbf{H}) \right\}. \quad (3.26)$$

As for property 4 of Theorem 3.14 above, we easily get

$$\nu \in \partial^- \phi(\mu) \Leftrightarrow [\mu, \nu] = \phi(\mu) + \phi^*(\nu). \quad (3.27)$$

As in the case of the L^2 -Wasserstein metric in \mathbf{H} , we have the following simple but crucial properties.

Corollary 3.15. *Let us consider the cost function $\mathbf{c} := \frac{1}{2}\mathbf{w}_2^2$, let $\phi, \psi : \mathcal{P}_2(\mathbf{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ and \mathcal{U}, \mathcal{V} be defined as*

$$\mathcal{U} := \frac{1}{2}\mathbf{m}_2^2 - \phi, \quad \mathcal{V} := \frac{1}{2}\mathbf{m}_2^2 - \psi. \quad (3.28)$$

The following hold true:

1.
$$\psi = \phi^* \Leftrightarrow \mathcal{V} = \mathcal{U}^{\mathbf{c}} \quad (3.29)$$
2. ϕ is totally convex and lower semicontinuous if and only if \mathcal{U} is \mathbf{c} -concave (and thus upper semicontinuous).
3. For every $\mu, \nu \in \mathcal{P}_2(\mathbf{H})$ with $\phi(\mu) \in \mathbb{R}$ (and thus $\mathcal{U}(\mu) \in \mathbb{R}$ as well)

$$\nu \in \partial^- \phi(\mu) \Leftrightarrow \nu \in \partial_{\mathbf{c}}^+ \mathcal{U}(\mu). \quad (3.30)$$

Proof. We repeatedly use the identity in (2.12).

Claim 1. For every $\nu \in \mathcal{P}_2(\mathbf{H})$ (3.28) yields

$$\begin{aligned} \mathcal{U}^{\mathbf{c}}(\nu) &= \inf_{\mu \in \mathcal{P}_2(\mathbf{H})} \frac{1}{2}\mathbf{w}_2^2(\mu, \nu) - \mathcal{U}(\mu) = \inf_{\mu \in \mathcal{P}_2(\mathbf{H})} \frac{1}{2}\mathbf{w}_2^2(\mu, \nu) - (\mathbf{m}_2^2(\mu) - \phi(\mu)) \\ &= \inf_{\mu \in \mathcal{P}_2(\mathbf{H})} \frac{1}{2}\mathbf{m}_2^2(\nu) - [\mu, \nu] + \phi(\mu) = \frac{1}{2}\mathbf{m}_2^2(\nu) + \inf_{\mu \in \mathcal{P}_2(\mathbf{H})} (-[\mu, \nu] + \phi(\mu)) \\ &= \frac{1}{2}\mathbf{m}_2^2(\nu) - \sup_{\mu \in \mathcal{P}_2(\mathbf{H})} ([\mu, \nu] - \phi(\mu)) \\ &= \frac{1}{2}\mathbf{m}_2^2(\nu) - \phi^*(\nu). \end{aligned}$$

The second claim immediately follows by the first one, Claim 3 of Corollary 3.10 and Claim 2 of Theorem 3.14.

Claim 3 then follows by (3.27) and Claim 4 of Theorem 3.14. \square

In the next section we will discuss a more refined representation of $\partial^- \phi$, where we replace $\mathcal{P}_2(\mathbf{H}) \times \mathcal{P}_2(\mathbf{H})$ by $\mathcal{P}_2(\mathbf{H} \times \mathbf{H})$.

4 Multivalued probability fields

We recall the notion of multivalued probability vector field (MPVF) \mathbf{F} , introduced in a different context by [Pic19] and in full generality by [CSS23b]: they are a natural extension of the usual notion of vector field for studying the evolution of probability measures.

The simplest way to define a MPVF \mathbf{F} is just to consider it as a nonempty subset of $\mathcal{P}_2(\mathbb{H} \times \mathbb{H})$. We set $D(\mathbf{F}) := \{\pi_{\#}^1 \gamma : \gamma \in \mathbf{F}\}$ and for every $\mu \in D(\mathbf{F})$ we consider the sections $\mathbf{F}[\mu] := \{\gamma : \pi_{\#}^1 \gamma = \mu\}$. Sections define a multivalued map $\mathbf{F}[\cdot] : \mathcal{P}_2(\mathbb{H}) \rightarrow 2^{\mathcal{P}_2(\mathbb{H} \times \mathbb{H})}$ with the property

$$\gamma \in \mathbf{F}[\mu] \quad \Rightarrow \quad \pi_{\#}^1 \gamma = \mu.$$

When $\mathbf{F}[\cdot]$ is single valued we say that \mathbf{F} is a *probability vector field (PVF)*. A particular case, that play a crucial role, is given by *deterministic* PVFs: they are characterized by the property

$$\mathbf{F} \subset \mathcal{P}_2^{\text{det}}(\mathbb{H}^2) \quad \text{so that } \mathbf{F}[\mu] = (\mathbf{i} \times \mathbf{f})_{\#} \mu \quad \text{for a (unique) } \mathbb{H}\text{-valued map } \mathbf{f} \in L^2(\mathbb{H}, \mu; \mathbb{H}), \quad (4.1)$$

where we adopted the obvious notation $\mathcal{P}_2^{\text{det}}(\mathbb{H}^2) := \mathcal{P}_2(\mathbb{H}^2) \cap \mathcal{P}^{\text{det}}(\mathbb{H}^2)$. If \mathbf{F} is a deterministic PVF, for every $\mu \in D(\mathbf{F})$ we can thus define a nonlocal vectorfield

$$\mathbf{f}[\mu] = \mathbf{f}(\cdot, \mu) \in L^2(\mathbb{H}, \mu; \mathbb{H}) \quad \text{such that (4.1) holds;} \quad (4.2)$$

notice that

$$\pi_{\#}^2 \mathbf{F}[\mu] = \mathbf{f}[\mu]_{\#} \mu. \quad (4.3)$$

Every MPVF \mathbf{F} admits a Lagrangian representation (or lifting) $\hat{\mathbf{F}} \subset \mathcal{H} \times \mathcal{H}$ defined by

$$(X, Y) \in \hat{\mathbf{F}} \quad \Leftrightarrow \quad \iota^2(X, Y) = (X, Y)_{\#} \mathbb{M} \in \mathbf{F}. \quad (4.4)$$

It is immediate to check that $\hat{\mathbf{F}}$ is law invariant

$$(X, Y) \in \hat{\mathbf{F}}, \quad \iota^2(X', Y') = \iota^2(X, Y) \quad \Rightarrow \quad (X', Y') \in \hat{\mathbf{F}} \quad (4.5)$$

and thus invariant with respect to the action of measure-preserving isomorphisms:

$$(X, Y) \in \hat{\mathbf{F}} \quad \Rightarrow \quad (\mathbf{g}^* X, \mathbf{g}^* Y) \in \hat{\mathbf{F}} \quad \text{for every } \mathbf{g} \in \mathbf{G}(\mathbb{Q}). \quad (4.6)$$

Conversely, if a subset $\mathbf{B} \subset \mathcal{H} \times \mathcal{H}$ is invariant by m.p.i. and *closed*, then it is also law invariant and it is the Lagrangian representation of a unique MPVF \mathbf{F} [CSS25].

4.1 Totally monotone and cyclically monotone MPVF

Inspired by the the definition of totally monotone MPVF (introduced in [CSS23a]) we introduce here the corresponding notion of totally cyclically monotone MPVF. We will adopt the notation

$$\pi^n : (\mathbb{H}^2)^N \rightarrow \mathbb{H}^2, \quad \pi^n((x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)) = (x_n, y_n), \quad n = 1, \dots, N. \quad (4.7)$$

Definition 4.1 (Totally monotone and cyclically monotone MPVF). *A multivalued probability vector field (MPVF) $\mathbf{F} \subset \mathcal{P}_2(\mathbb{H}^2)$ is totally monotone if for every $\theta \in \mathcal{P}_2(\mathbb{H}^2 \times \mathbb{H}^2)$ with $\pi_{\#}^1 \theta \in \mathbf{F}$, $\pi_{\#}^2 \theta \in \mathbf{F}$ we have*

$$\int \langle y_2 - y_1, x_2 - x_1 \rangle d\theta(x_1, y_1; x_2, y_2) \geq 0. \quad (4.8)$$

\mathbf{F} is a maximal totally monotone MPVF if it is totally monotone and the inclusion $\mathbf{F} \subset \mathbf{G}$, \mathbf{G} totally monotone, yields $\mathbf{F} = \mathbf{G}$.

\mathbf{F} is totally cyclically monotone if for every $N \in \mathbb{N}$, $\theta \in \mathcal{P}_2((\mathbb{H}^2)^N)$, with $\pi_{\#}^n \theta \in \mathbf{F}$, $n = 1, \dots, N$, and $\sigma \in \mathbf{S}_N = \text{Sym}(\{1, \dots, N\})$ permutation

$$\sum_{n=1}^N \int \langle y_n, x_n - x_{\sigma(n)} \rangle d\theta(x_1, y_1; x_2, y_2; \dots; x_N, y_N) \geq 0. \quad (4.9)$$

In the case of a deterministic PVF induced by a nonlocal vector field \mathbf{f} as in (4.1),(4.2), the total monotonicity condition (4.9) reads as

$$\int \langle \mathbf{f}(x_2, \mu_2) - \mathbf{f}(x_1, \mu_1), x_2 - x_1 \rangle d\boldsymbol{\mu}(x_1, x_2) \geq 0 \quad \text{for every } \mu_i \in D(\mathbf{F}), \boldsymbol{\mu} \in \Gamma(\mu_1, \mu_2); \quad (4.10)$$

Similarly, (4.8) reads as

$$\sum_{n=1}^N \int \langle \mathbf{f}(x_n, \mu_n), x_n - x_{\sigma(n)} \rangle d\boldsymbol{\mu} \geq 0 \quad \text{for every } \mu_i \in D(\mathbf{F}), \boldsymbol{\mu} \in \Gamma(\mu_1, \mu_2, \dots, \mu_N). \quad (4.11)$$

Let $\hat{\mathbf{F}}$ be the Lagrangian representation of \mathbf{F} ; it is easy to check that

$$\mathbf{F} \text{ is totally monotone} \quad \Leftrightarrow \quad \hat{\mathbf{F}} \text{ is monotone in } \mathcal{H} \times \mathcal{H}, \quad (4.12)$$

and, according to the main result of [CSS25], we also have

$$\mathbf{F} \text{ is maximal totally monotone} \quad \Leftrightarrow \quad \hat{\mathbf{F}} \text{ is maximal monotone in } \mathcal{H} \times \mathcal{H}. \quad (4.13)$$

A maximal totally monotone MPVF has many important properties, see [CSS23a]. We quote here the most relevant ones for our discussion; the first one involves the (Borel) barycentric map $\mathbf{b} : \mathbb{H} \times \mathcal{P}_2(\mathbb{H} \times \mathbb{H}) \rightarrow \mathbb{H}$ defined by using the universal disintegration kernel (2.2):

$$\mathbf{b}(x, \gamma) := \int y d\mathcal{K}(x, \gamma)(y) = \int y d\kappa_x(y), \quad (4.14)$$

where $\kappa_x = \mathcal{K}(x, \gamma)$ is the disintegration of γ w.r.t. its first marginal.

Proposition 4.2. *Let \mathbf{F} be a maximal totally monotone MPVF.*

1. *If $\gamma \in \mathbf{F}$ and $\mu = \pi_{\sharp}^{\downarrow} \gamma \in D(\mathbf{F})$, then the barycentric projection $\mathbf{b}(\cdot, \gamma)$ of γ satisfies*

$$(\mathbf{i} \times \mathbf{b}(\cdot, \gamma))_{\sharp} \mu \in \mathbf{F}[\mu]. \quad (4.15)$$

2. *There exists a unique minimal section $\mathbf{F}^{\circ} \subset \mathbf{F} \cap \mathcal{P}_2^{\det}(\mathbb{H} \times \mathbb{H})$ and a Borel nonlocal vector field $\mathbf{f}^{\circ} = \mathbf{b}(\cdot, \mathbf{F}^{\circ}) : \mathbb{H} \times D(\mathbf{F}) \rightarrow \mathbb{H}$ such that for every $\mu \in D(\mathbf{F})$ we have $\mathbf{F}^{\circ}[\mu] = (\mathbf{i} \times \mathbf{f}^{\circ}(\cdot, \mu))_{\sharp} \mu$ and*

$$\int_{\mathbb{H}} |\mathbf{f}^{\circ}(x, \mu)|^2 d\mu(x) \leq \int_{\mathbb{H}} |\mathbf{b}(x, \gamma)|^2 d\mu(x) \leq \int_{\mathbb{H}} |y|^2 d\gamma(x, y) \quad \text{for every } \gamma \in \mathbf{F}[\mu]. \quad (4.16)$$

3. *If $\hat{\mathbf{F}}$ is the Lagrangian representation of \mathbf{F} , then $D(\hat{\mathbf{F}}) = (\iota_{\sharp})^{-1}(D(\mathbf{F}))$ and the minimal section $\hat{\mathbf{F}}^{\circ}$ of $\hat{\mathbf{F}}$ satisfies*

$$Y = \hat{\mathbf{F}}^{\circ}(X) \quad \Leftrightarrow \quad Y(q) = \mathbf{f}^{\circ}(X(q), \mu) = \mathbf{f}^{\circ}[\mu](X(q)) \quad \text{for } \mathbb{M}\text{-a.e. } q, \mu = \iota(X). \quad (4.17)$$

Let us focus now on totally cyclically monotone MPVFs: first of all, we have a simple lifting result.

Proposition 4.3 (Lifting of totally cyclically monotone MPVF). *A MPVF \mathbf{F} is totally cyclically monotone if and only if $\hat{\mathbf{F}}$ is cyclically monotone in $\mathcal{H} \times \mathcal{H}$.*

Proof. It is sufficient to observe that for every $\boldsymbol{\theta} \in \mathcal{P}_2((\mathbb{H}^2)^N)$ with $\pi_{\sharp}^n \boldsymbol{\theta} \in \mathbf{F}$, $n = 1, \dots, N$, we can find a Borel map $Z = ((X_1, Y_1), \dots, (X_N, Y_N)) : \mathbb{Q} \rightarrow (\mathbb{H}^2)^N$ such that $Z_{\sharp} \mathbb{M} = \boldsymbol{\theta}$, so that $(X_n, Y_n) \in \hat{\mathbf{F}}$ for every $n \in \{1, \dots, N\}$. Conversely, if $(X_n, Y_n) \in \hat{\mathbf{F}}$ then $\boldsymbol{\theta} = Z_{\sharp} \mathbb{M}$ belongs to $\mathcal{P}_2((\mathbb{H}^2)^N)$ and $\pi_{\sharp}^n \boldsymbol{\theta} \in \mathbf{F}$.

For every permutation $\sigma \in S_N$ we can then use the identity

$$\sum_{n=1}^N \int \langle y_n, x_n - x_{\sigma(n)} \rangle d\boldsymbol{\theta} = \sum_{n=1}^N \int \langle Y_n, X_n - X_{\sigma(n)} \rangle d\mathbb{M} = \sum_{n=1}^N \langle Y_n, X_n - X_{\sigma(n)} \rangle_{\mathcal{H}}$$

which shows the equivalence between condition (4.9) and the corresponding condition expressing the cyclical monotonicity of $\hat{\mathbf{F}}$ in $\mathcal{H} \times \mathcal{H}$. \square

As in the Hilbertian framework, we will show that totally cyclically monotone MPVFs are strictly linked with the notion of *total subdifferential*. Let us first recall the definition (we refer to [CSS23a] for a detailed comparison with other notion of subdifferentiability, in particular the ones introduced in [AGS08], from a metric perspective).

Definition 4.4 (Total subdifferential). *The total subdifferential of $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ is the set $\partial_t \phi \subset \mathcal{P}_2(\mathbb{H} \times \mathbb{H})$ characterized by the following property: a plan $\gamma \in \mathcal{P}_2(\mathbb{H}^2)$ belongs to $\partial_t \phi$ if and only if $\mu = \pi_{\#}^1 \gamma \in D(\phi)$ and for every $\nu \in D(\phi)$ and $\theta \in \Gamma(\gamma, \nu)$ we have*

$$\phi(\nu) - \phi(\mu) \geq \int_{\mathbb{H}^2 \times \mathbb{H}} \langle y_1, x_2 - x_1 \rangle d\theta(x_1, y_1; x_2). \quad (4.18)$$

Notice that if $\partial_t \phi$ is not empty then ϕ satisfies (3.8): if $\gamma \in \partial_t \phi[\mu]$, $\mu' := \pi_{\#}^2 \gamma$, $\gamma' \in \Gamma_o(\nu, \mu')$ we can apply the glueing Lemma to select $\theta \in \mathcal{P}_2(\mathbb{H}^3)$ with $\pi_{\#}^1 \theta = \gamma$, $\pi_{\#}^2 \theta = \gamma'$ so that (4.18) yields

$$\phi(\nu) \geq \phi(\mu) - \int_{\mathcal{H}^2} \langle y, x \rangle d\gamma(x, y) + \int_{\mathcal{H}^2} \langle y, z \rangle d\gamma'(y, z) = -a + [\nu, \mu']$$

where $a := \int_{\mathcal{H}^2} \langle y, x \rangle d\gamma(x, y) - \phi(\mu)$.

Applying [CSS23a, Proposition 5.2] we can show that there is a strong relation between the total subdifferential of ϕ and the (convex) subdifferential of its Lagrangian lifting $\hat{\phi} = \phi \circ \iota$ in \mathcal{H} .

Theorem 4.5 (Lifting of total subdifferentials). *Let $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ be a proper lower semicontinuous function with Lagrangian lift $\hat{\phi} := \phi \circ \iota$. The (possible empty) convex subdifferential $\partial \hat{\phi} \subset \mathcal{H} \times \mathcal{H}$ is law invariant and coincides with the Lagrangian representation of $\partial_t \phi$, i.e.*

$$\widehat{\partial_t \phi} = \partial \hat{\phi}. \quad (4.19)$$

In particular $\partial_t \phi$ is totally cyclically monotone and

$$\partial_t \phi \subset \partial_t(\phi^{**}), \quad \partial_t \phi[\mu] \neq \emptyset \Leftrightarrow \phi(\mu) = \phi^{**}(\mu), \quad \partial_t \phi[\mu] = \partial_t \phi^{**}[\mu] \neq \emptyset. \quad (4.20)$$

If moreover ϕ is also totally convex then $\partial_t \phi$ is maximal totally monotone and its minimal section $\partial_t^\circ \phi$ is associated with the minimal section $\partial^\circ \hat{\phi}$ through (4.17).

Proof. Since $\hat{\phi}$ is l.s.c., $\partial \hat{\phi}$ is a closed subset of $\mathcal{H} \times \mathcal{H}$. By Claim 1 of [CSS23a, Proposition 5.2] we deduce that $\partial \hat{\phi}$ is law invariant and we can then apply Claim 3 of [CSS23a, Proposition 5.2] (which uses only the law invariance of $\partial \hat{\phi}$) to deduce that $\partial \hat{\phi}$ is the Lagrangian representation of $\partial_t \phi$ according to (4.19). (4.20) then follow by the corresponding well known properties of $\partial \hat{\phi}$.

When ϕ is also totally convex, the same Proposition 5.2 shows that $\partial_t \phi$ is maximal totally monotone and (4.12) shows that it is totally cyclically monotone (since $\partial \hat{\phi}$ is cyclically monotone in $\mathcal{H} \times \mathcal{H}$). We conclude by applying Claim 3 of Proposition 4.2. \square

Thanks to the previous result, we can now extend the celebrated Rockafellar Theorem to totally cyclically monotone MPVF.

Theorem 4.6 (Totally cyclically maximal monotone MPVFs are total subdifferentials). *If \mathbf{F} is a totally cyclically monotone MPVF in $\mathcal{P}_2(\mathbb{H}^2)$ then there exists a proper, totally convex and lower semicontinuous function $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ such that $\mathbf{F} \subset \partial_t \phi$. In particular, \mathbf{F} has a maximal totally monotone extension which is cyclically monotone and every totally cyclically maximal monotone MPVF is the total subdifferential of a totally convex function.*

Proof. Let us denote by $\hat{\mathbf{F}}$ the Lagrangian representation of \mathbf{F} in \mathcal{H}^2 ; we know that $\hat{\mathbf{F}}$ is invariant by measure preserving isomorphisms, i.e.

$$(X, Y) \in \hat{\mathbf{F}} \quad \Rightarrow \quad (X \circ g, Y \circ g) \in \hat{\mathbf{F}} \quad \text{for every } g \in G(Q). \quad (4.21)$$

We fix $(X_0, Y_0) \in \hat{\mathbf{F}}$ and use Rockafellar construction [Roc66] to define

$$\Phi(X) := \sup \left\{ \langle X - X_N, Y_N \rangle_{\mathcal{H}} + \sum_{n=1}^N \langle X_n - X_{n-1}, Y_{n-1} \rangle_{\mathcal{H}} : N \in \mathbb{N}, (X_n, Y_n) \in \hat{\mathbf{F}} \right\}. \quad (4.22)$$

We know that Φ is convex, proper, lower semicontinuous, and $\hat{\mathbf{F}} \subset \partial\Phi$. Let us prove that Φ is invariant by measure-preserving isomorphisms: we use the fact that for every N -tuple $(X_n, Y_n) \in \hat{\mathbf{F}}$ we also have $(X_n \circ \mathbf{g}, Y_n \circ \mathbf{g}) \in \hat{\mathbf{F}}$ so that

$$\begin{aligned} \Phi(X \circ \mathbf{g}) &\geq \langle X \circ \mathbf{g} - X_N \circ \mathbf{g}, Y_N \circ \mathbf{g} \rangle_{\mathcal{H}} + \sum_{n=1}^N \langle X_n \circ \mathbf{g} - X_{n-1} \circ \mathbf{g}, Y_{n-1} \circ \mathbf{g} \rangle_{\mathcal{H}} \\ &= \langle X - X_N, Y_N \rangle_{\mathcal{H}} + \sum_{n=1}^N \langle X_n - X_{n-1}, Y_{n-1} \rangle_{\mathcal{H}} \end{aligned}$$

so that taking the supremum with respect to $(X_n, Y_n) \in \hat{\mathbf{F}}$ and $N \in \mathbb{N}$ we get

$$\Phi(X \circ \mathbf{g}) \geq \Phi(X).$$

Applying the same inequality to $X \circ \mathbf{g}$ with \mathbf{g} replaced by \mathbf{g}^{-1} we also obtain $\Phi(X \circ \mathbf{g} \circ \mathbf{g}^{-1}) = \Phi(X) \geq \Phi(X \circ \mathbf{g})$ so that $\Phi(X \circ \mathbf{g}) = \Phi(X)$ for every X and every $\mathbf{g} \in \mathbf{G}(\mathbb{Q})$. Since Φ is lower semicontinuous and invariant by the action of $\mathbf{G}(\mathbb{Q})$, Lemma 2.8 shows that $\Phi = \phi \circ \iota$ for a lower semicontinuous function $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$, which is clearly totally convex and satisfies $\mathbf{F} \subset \partial_t \phi$ by Proposition 4.5. \square

We briefly discuss the case when ϕ is differentiable, in a suitable Wasserstein sense. First of all by using a (Borel) barycentric map \mathbf{b} as in (4.14) we can define the nonlocal deterministic field $\nabla_W \phi : \mathbb{H} \times D(\partial_t \phi) \rightarrow \mathbb{H}$ as in Proposition 4.2:

$$\nabla_W \phi(x, \mu) = \nabla_W \phi[\mu](x) := \mathbf{b}(x, \partial_t^\circ \phi[\mu]). \quad (4.23)$$

Recall that if a convex function $\psi : \mathcal{H} \rightarrow \mathbb{R} \cup \{\infty\}$ is Gateaux differentiable at $X \in D(\psi)$ then $\partial\psi(X)$ is a singleton (thus coinciding with the minimal section $\partial^\circ\psi$). The converse is also true if in addition ψ is continuous at X [ET76, Chap. I, Prop. 5.3]. We can say that

$$\phi \text{ is } W\text{-differentiable at } \mu \text{ if } \partial_t \phi[\mu] \text{ contains a unique element,} \quad (4.24)$$

which in turn coincides with the minimal section, expressed through $\nabla_W \phi$.

Proposition 4.7. *Let $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ be a proper, l.s.c., and totally convex function.*

1. ϕ is W -differentiable at μ if and only if $\partial\hat{\phi}$ is a singleton at every $X \in \iota^{-1}(\mu)$.
2. If ϕ is continuous at μ then it is W -differentiable if and only if $\hat{\phi}$ is Gateaux-differentiable at every point of $\iota^{-1}(\mu)$.
3. For every $\tau > 0$ ϕ_τ is W -differentiable everywhere, the map $(x, \mu) \mapsto \nabla_W \phi_\tau(x, \mu)$ is everywhere defined and continuous in

$$\mathcal{S}(\mathbb{H}) := \left\{ (x, \mu) \in \mathbb{H} \times \mathcal{P}_2(\mathbb{H}) : x \in \text{supp}(\mu) \right\} \quad (4.25)$$

for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ the map $x \mapsto \nabla_W \phi_\tau(x, \mu)$ is τ^{-1} -Lipschitz in $\text{supp}(\mu)$, and the map $\mu \mapsto \nabla_W \phi_\tau(\cdot, \mu)_{\#} \mu$ is τ^{-1} -Lipschitz in $\mathcal{P}_2(\mathbb{H})$.

Proof. To prove the first claim it is sufficient to recall that

$$(X, Y) \in \partial \hat{\phi} \quad \Leftrightarrow \quad (X, Y)_{\sharp} \mathbb{M} \in \partial_t \phi$$

$\partial \phi[\mu]$ contains just one element γ if and only if it is reduced to the minimal section, i.e. $\gamma = (i \times \nabla_W \phi(\cdot, \mu))_{\sharp} \mu$ so that $\iota^2(X, Y) = \gamma$ implies $Y = \nabla_W(X, \mu)$ and therefore $\partial \phi(X)$ is reduced to a singleton for every $X \in \iota^{-1}(\mu)$.

The second claim then follows since continuity of ϕ at μ implies continuity of $\hat{\phi}$ at every $X \in \iota^{-1}(\mu)$.

Finally, the third claim follows by (3.18) and the well known properties of the Moreau-Yosida regularization in the Hilbert space \mathcal{H} : $\hat{\phi}_\tau$ is of class $C^{1,1}$ and its differential $D\hat{\phi}_\tau$ (which provides the unique element of $\partial \hat{\phi}_\tau$) is τ^{-1} -Lipschitz. We can then apply the results of [CSS25, Section 4] and [CSS23a, Section 5] to conclude. \square

4.2 Marginal projections of MPVF

We notice that a MPVF $\mathbf{F} \subset \mathcal{P}_2(\mathbb{H} \times \mathbb{H})$ induces a subset $\mathfrak{F} \subset \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H})$ by applying marginal projections

$$\mathfrak{F} = (\pi_{\sharp}^1, \pi_{\sharp}^2) \mathbf{F}, \quad \text{or, equivalently,} \quad (\mu, \nu) \in \mathfrak{F} \quad \Leftrightarrow \quad \mathbf{F} \cap \Gamma(\mu, \nu) \neq \emptyset. \quad (4.26)$$

It is therefore natural to investigate the relations between the total subdifferential $\partial_t \phi$ (a MPVF) and the $[\cdot, \cdot]$ -subdifferential of ϕ (a subset of $\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H})$).

We start with a nice characterization of the total subdifferential of ϕ in terms of its Kantorovich-Legendre-Fenchel transform.

Theorem 4.8 (Total subdifferential and Kantorovich-Legendre-Fenchel transform). *A plan $\gamma \in \mathcal{P}_2(\mathbb{H}^2)$ belongs to $\partial_t \phi[\mu]$ if and only if $\mu = \pi_{\sharp}^1 \gamma$, $\nu = \pi_{\sharp}^2 \gamma \in D(\phi^*)$ and*

$$\int_{\mathbb{H}^2} \langle x, y \rangle d\gamma(x, y) = \phi(\mu) + \phi^*(\nu). \quad (4.27)$$

In particular $\gamma \in \Gamma_o(\mu, \nu)$ and $\int_{\mathbb{H}^2} \langle x, y \rangle d\gamma(x, y) = [\mu, \nu]$.

Proof. We have seen that $\gamma \in \partial_t \phi[\mu]$ if and only if for every $(X, Y) \in \mathcal{H}^2$ with $\iota^2(X, Y) = \gamma$ we have $(X, Y) \in \partial \hat{\phi}$. In turn, this is equivalent to

$$\langle X, Y \rangle_{\mathcal{H}} = \hat{\phi}(X) + (\hat{\phi})^*(Y) = \phi(\mu) + \phi^*(\nu).$$

Since $\langle X, Y \rangle_{\mathcal{H}} = \int \langle x, y \rangle d\gamma$ we obtain (4.27). \square

Combining Theorem 4.8 with (3.27) we can immediately link the total subdifferential $\partial_t \phi$ with the $[\cdot, \cdot]$ subdifferential $\partial^- \phi$.

Corollary 4.9 (Total and $[\cdot, \cdot]$ subdifferential). *Let $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ be a function satisfying the lower bound (3.8), $\mu, \nu \in \mathcal{P}_2(\mathbb{H})$ with $\phi(\mu) \in \mathbb{R}$. Then*

$$\gamma \in \partial_t \phi[\mu] \quad \Leftrightarrow \quad \nu \in \partial^- \phi(\mu), \quad \gamma \in \Gamma_o(\mu, \nu). \quad (4.28)$$

In particular $\partial_t \phi$ can be obtained as the image of the graph of $\partial^- \phi$ in $\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H})$ through the multivalued map $\Gamma_o : \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H}) \rightrightarrows \mathcal{P}_2(\mathbb{H} \times \mathbb{H})$.

Conversely, the graph of $\partial^- \phi$ can be obtained as the image of $\partial_t \phi \subset \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H})$ through the map $\pi_{\sharp}^1 \times \pi_{\sharp}^2$.

It is worth highlighting the following consequence of the previous result, which does not immediately appear from the definition of total subdifferential.

Corollary 4.10 (Total subdifferentials are optimal couplings). *For every function $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ its total subdifferential $\partial_t \phi$ is contained in the set of optimal couplings $\mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H})$.*

The above corollary suggests a lifting procedure of a subset $\mathfrak{F} \subset \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H})$ which defines a MPVF \mathbf{F} by the formula

$$\mathbf{F} := \left\{ \gamma \in \mathcal{P}_2(\mathbb{H}^2) : \gamma \in \Gamma_o(\mu, \nu) \text{ for some } (\mu, \nu) \in \mathfrak{F} \right\} = (\pi_{\sharp}^1, \pi_{\sharp}^2)^{-1}(\mathfrak{F}) \cap \mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H}). \quad (4.29)$$

Theorem 4.11 (Total cyclical monotonicity and w_2^2 -cyclical monotonicity). *If the MPVF \mathbf{F} is totally cyclically monotone then the set $\mathfrak{F} = (\pi_{\sharp}^1, \pi_{\sharp}^2)\mathbf{F}$ defined as in (4.26) is w_2^2 -cyclically monotone in $\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H})$ in the sense that for every $N \in \mathbb{N}$, every choice of pairs $(\mu_1, \nu_1), \dots, (\mu_N, \nu_N) \in \mathfrak{F}$ and every permutation $\sigma \in S_N$*

$$\sum_{n=1}^N w^2(\mu_n, \nu_n) \leq \sum_{n=1}^N w^2(\mu_n, \nu_{\sigma(n)}). \quad (4.30)$$

Conversely, if \mathfrak{F} is w_2^2 -cyclically monotone in $\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H})$ according to (4.30) then \mathbf{F} defined by (4.29) is totally cyclically monotone.

Proof. Let us first observe that for arbitrary choice of $(\mu_n, \nu_n) \in \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H})$ the identity (2.12) yields

$$\frac{1}{2} \sum_{n=1}^N w^2(\mu_n, \nu_{\sigma(n)}) - \frac{1}{2} \sum_{n=1}^N w^2(\mu_n, \nu_n) = \sum_{n=1}^N \left([\mu_n, \nu_n] - [\mu_n, \nu_{\sigma(n)}] \right).$$

Suppose that \mathbf{F} is totally cyclically monotone; by Theorem 4.6 we can find a proper totally convex l.s.c. function $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ such that $\mathbf{F} \subset \partial_t \phi$ so that in particular all the elements of \mathbf{F} are optimal couplings of $\mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H})$. For every pair $(\mu_n, \nu_n) \in \mathfrak{F}$, $n = 1, \dots, N$ we thus get

$$\begin{aligned} \sum_{n=1}^N [\mu_n, \nu_n] - [\mu_n, \nu_{\sigma(n)}] &\geq \sum_{n=1}^N \phi(\mu_n) + \phi^*(\nu_n) - \left(\phi(\mu_n) + \phi^*(\nu_{\sigma(n)}) \right) \\ &= \sum_{n=1}^N \phi^*(\nu_n) - \sum_{n=1}^N \phi^*(\nu_{\sigma(n)}) = 0. \end{aligned}$$

Conversely, if \mathfrak{F} is w_2^2 -cyclically monotone and \mathbf{F} is defined as in (4.29) we have for every $\theta \in \mathcal{P}_2((\mathbb{H}^2)^N)$ with $\pi_{\sharp}^n \theta = \gamma_n \in \mathbf{F}$

$$\begin{aligned} \sum_{n=1}^N \int \langle y_n, x_n - x_{\sigma(n)} \rangle d\theta &= \sum_{n=1}^N \int \langle y_n, x_n \rangle d\theta - \sum_{n=1}^N \int \langle y_n, x_{\sigma(n)} \rangle d\theta \\ &= \sum_{n=1}^N \int \langle y, x \rangle d\gamma_n(x, y) - \sum_{n=1}^N \int \langle y_n, x_{\sigma(n)} \rangle d\theta \\ &= \sum_{n=1}^N [\mu_n, \nu_n] - \sum_{n=1}^N \int \langle y_n, x_{\sigma(n)} \rangle d\theta \\ &\geq \sum_{n=1}^N [\mu_n, \nu_n] - \sum_{n=1}^N [\mu_n, \nu_{\sigma(n)}] \geq 0 \end{aligned}$$

□

5 L^2 -Random Optimal Transport

In this section we will apply the main results of Section 3 for totally convex functions and of Section 4 for totally cyclically monotone MPVF to optimal transport in $\mathcal{P}_2(\mathbb{H})$.

5.1 Random couplings and couplings of random measures

Let us first observe that to every random coupling law $\mathbf{P} \in \mathfrak{P}_2(\mathbb{H} \times \mathbb{H}) = \mathcal{P}_2(\mathcal{P}_2(\mathbb{H} \times \mathbb{H}))$ we can associate a coupling between (laws of) random measures Π by the formula

$$\Pi = (\pi_{\#}^1, \pi_{\#}^2)_{\#} \mathbf{P}, \quad \Pi = \int \delta_{\pi_{\#}^1 \gamma} \otimes \delta_{\pi_{\#}^2 \gamma} d\mathbf{P}(\gamma). \quad (5.1)$$

If $\mathbf{M}_i = \pi_{\#}^i \Pi$, it is not difficult to check that

$$\mathbb{W}_2^2(\mathbf{M}_1, \mathbf{M}_2) \leq \int \mathbb{w}_2^2(\mu_1, \mu_2) d\Pi(\mu_1, \mu_2) \leq \int \left(\int |x_1 - x_2|^2 d\gamma(x_1, x_2) \right) d\mathbf{P}(\gamma). \quad (5.2)$$

Conversely, using the fact that the continuous map

$$\pi_{\#}^1 \times \pi_{\#}^2 : \mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H}) \rightarrow \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H}) \quad \text{is surjective,} \quad (5.3)$$

given a coupling $\Pi \in \mathcal{P}_2(\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H}))$ we can find a $\mathbf{P} \in \mathcal{P}_2(\mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H}))$ (thus concentrated on optimal couplings) such that

$$\Pi = (\pi_{\#}^1, \pi_{\#}^2)_{\#} \mathbf{P}, \quad \int \mathbb{w}_2^2(\mu_1, \mu_2) d\Pi(\mu_1, \mu_2) = \int \left(\int |x_1 - x_2|^2 d\gamma(x_1, x_2) \right) d\mathbf{P}(\gamma). \quad (5.4)$$

We immediately deduce an equivalent characterization of \mathbb{W}_2 in terms of random coupling laws. We write $\pi_{\#}^i := (\pi_{\#}^i)_{\#}$ and we call

$$\text{R}\Gamma(\mathbf{M}_1, \mathbf{M}_2) := \left\{ \mathbf{P} \in \mathfrak{P}_2(\mathbb{H} \times \mathbb{H}), \pi_{\#}^i \mathbf{P} = \mathbf{M}_i \right\}. \quad (5.5)$$

Proposition 5.1 (Random couplings formulation of OT between random measures). *For every $\mathbf{M}_1, \mathbf{M}_2 \in \mathfrak{P}_2(\mathbb{H})$ we have*

$$\mathbb{W}_2^2(\mathbf{M}_1, \mathbf{M}_2) = \min \left\{ \int \left(\int |x_1 - x_2|^2 d\gamma(x_1, x_2) \right) d\mathbf{P}(\gamma) : \mathbf{P} \in \text{R}\Gamma(\mathbf{M}_1, \mathbf{M}_2) \right\}. \quad (5.6)$$

Moreover, \mathbf{P} is optimal for (5.6) if and only if \mathbf{P} is concentrated on the optimal couplings of $\mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H})$ and $\Pi = (\pi_{\#}^1, \pi_{\#}^2)_{\#} \mathbf{P} \in \Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$.

We denote by $\text{R}\Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$ the set of optimal random couplings for (5.6).

Proof. By (5.2) it is clear that the infimum of the quantities in the right-hand side of (5.6) is larger than $\mathbb{W}_2^2(\mathbf{M}_1, \mathbf{M}_2)$ and that if \mathbf{P} satisfies the optimal condition

$$\mathbb{W}_2^2(\mathbf{M}_1, \mathbf{M}_2) = \int \left(\int |x_1 - x_2|^2 d\gamma(x_1, x_2) \right) d\mathbf{P}(\gamma),$$

then $\Pi = (\pi_{\#}^1, \pi_{\#}^2)_{\#} \mathbf{P}$ is optimal thanks to (5.2).

On the other hand, we can prove the equality by choosing an optimal coupling $\Pi \in \Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$ and a corresponding lift \mathbf{P} as in (5.4). \square

5.2 L^2 -Optimal Transport for laws of random measures

Let us first introduce the natural pairing in $\mathfrak{P}_2(\mathbb{H})$ associated with the maximal correlation pairing $[\cdot, \cdot]$.

Definition 5.2. *For every $\mathbf{M}_1, \mathbf{M}_2 \in \mathfrak{P}_2(\mathbb{H})$ we set*

$$[\mathbf{M}_1, \mathbf{M}_2] := \max \left\{ \int_{\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H})} [\mu, \nu] d\Pi(\mu, \nu) : \Pi \in \Gamma(\mathbf{M}_1, \mathbf{M}_2) \right\}. \quad (5.7)$$

Recalling the identity (2.12) we immediately have the corresponding property at the level of W_2 :

Lemma 5.3. *For every $M_1, M_2 \in \mathcal{P}_2(\mathbb{H})$ we have*

$$W_2^2(M_1, M_2) = M_2^2(M_1) + M_2^2(M_2) - 2\llbracket M_1, M_2 \rrbracket. \quad (5.8)$$

In particular, the class of optimal couplings for (5.7) coincides with the class $\Gamma_o(M_1, M_2)$ of optimal couplings for (2.14). Moreover, we have the equivalent formulation

$$\llbracket M_1, M_2 \rrbracket = \max \left\{ \int \left(\int \langle x_1, x_2 \rangle d\gamma(x_1, x_2) \right) dP(\gamma) : P \in R\Gamma(M_1, M_2) \right\} \quad (5.9)$$

whose solution is provided by the same class $R\Gamma_o(M_1, M_2)$ of optimal random couplings of (5.6).

The general duality result for Optimal Transport and Corollary 3.15 then yield:

Theorem 5.4 (Optimal Kantorovich potentials for ROT). *For every proper lower semicontinuous function $\zeta : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$*

$$\int \zeta(\mu) dM_1(\mu) + \int \zeta^*(\nu) dM_2(\nu) \geq \llbracket M_1, M_2 \rrbracket \quad (5.10)$$

and there exists a totally convex, lower semicontinuous and proper function $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ such that

$$\int \phi(\mu) dM_1(\mu) + \int \phi^*(\nu) dM_2(\nu) = \llbracket M_1, M_2 \rrbracket. \quad (5.11)$$

The corresponding potential $\mathcal{U} = \frac{1}{2}m_2^2 - \phi$ defined as in (3.28) satisfy

$$\mathcal{V} = \mathcal{U}^c = \frac{1}{2}m_2^2 - \phi^*, \quad \int \mathcal{U}(\mu) dM_1(\mu) + \int \mathcal{U}^c(\nu) dM_2(\nu) = \frac{1}{2}W_2^2(M_1, M_2), \quad (5.12)$$

with respect to the cost $c := \frac{1}{2}w_2^2$

We collect now the main results concerning optimality and duality.

Theorem 5.5 (Optimality conditions). *Let $M_1, M_2 \in \mathcal{P}_2(\mathbb{H})$, $\Pi \in \Gamma(M_1, M_2) \subset \mathcal{P}_2(\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H}))$, $P \in \mathcal{P}_2(\mathbb{H} \times \mathbb{H})$ supported in $\mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H})$ and associated with Π via $(\pi_{\sharp}^1, \pi_{\sharp}^2)_{\sharp} P = \Pi$ as in (5.4), so that in particular*

$$(\pi_{\sharp}^1, \pi_{\sharp}^2)(\text{supp } P) \subset \text{supp } \Pi \subset \overline{(\pi_{\sharp}^1, \pi_{\sharp}^2)(\text{supp } P)}, \quad \text{supp } P \subset (\pi_{\sharp}^1, \pi_{\sharp}^2)^{-1}(\text{supp } \Pi) \cap \mathcal{P}_{2,o}(\mathbb{H}^2). \quad (5.13)$$

The following properties are equivalent:

1. Π is an optimal plan in $\Gamma_o(M_1, M_2)$ for W_2 or, equivalently, for $\llbracket \cdot, \cdot \rrbracket$.
2. $\text{supp}(\Pi)$ is w_2^2 -cyclically monotone (recall Theorem 4.11).
3. P is an optimal random coupling law in $R\Gamma_o(M_1, M_2)$ for W_2 (according to (5.6)) or, equivalently, for $\llbracket \cdot, \cdot \rrbracket$ (according to (5.9)).
4. $\text{supp}(P)$ is totally cyclically monotone.
5. There exists a totally convex, lower semicontinuous and proper function $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ such that

$$\phi(\mu) + \phi^*(\nu) = [\mu, \nu] \quad \text{for } \Pi\text{-a.e. } (\mu, \nu) \in \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H}), \quad (5.14)$$

i.e. $\text{supp}(\Pi) \subset \partial^- \phi$. Moreover such a property holds for every pair of optimal Kantorovich potentials satisfying (5.11).

6. There exists a totally convex, lower semicontinuous and proper function $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R} \cup \{\infty\}$ such that

$$\phi(\pi_{\sharp}^1 \gamma) + \phi^*(\pi_{\sharp}^2 \gamma) = \int \langle x_1, x_2 \rangle d\gamma \quad \text{for } \mathbb{P}\text{-a.e. } \gamma \in \mathcal{P}_2(\mathbb{H} \times \mathbb{H}), \quad (5.15)$$

i.e. $\text{supp}(\mathbb{P}) \subset \partial_t \phi$. Moreover, such a property holds for every pair of optimal Kantorovich potentials satisfying (5.11).

Proof. The equivalence $1 \Leftrightarrow 2$ follows by the general theory of optimal transport.

The equivalence $1 \Leftrightarrow 3$ follows by Proposition 5.1.

Clearly $2 \Rightarrow 3$ by the last inclusion of (5.13) and the second part of Theorem 4.11. On the other hand, if $\text{supp}(\mathbb{P})$ is totally cyclically monotone the same Theorem 4.11 and the inclusion $\text{supp } \Pi \subset \overline{(\pi_{\sharp}^1, \pi_{\sharp}^2) \text{supp } \mathbb{P}}$ shows that $\text{supp } \Pi$ is w_2^2 -cyclically monotone, and therefore Π is optimal.

In order to prove the implication $1 \Rightarrow 5$ it is sufficient to select an optimal totally convex proper and l.s.c. function ϕ satisfying (5.11). The optimality of Π yields

$$\int \left(\phi(\mu_1) + \phi^*(\mu_2) - [\mu_1, \mu_2] \right) d\Pi(\mu_1, \mu_2) = 0 \quad (5.16)$$

which implies (5.14) thanks to the Kantorovich-Fenchel inequality (3.11). On the other hand, if Π satisfies (5.14) we get (5.16) and the optimality of Π thanks to (5.10).

A similar argument shows the equivalence with Claim 6. \square

As in the usual deterministic case, when \mathbb{M}_2 is concentrated on a set of measures with uniformly bounded quadratic moment, we can find a Lipschitz totally convex optimal Kantorovich potential.

Corollary 5.6. *Let us suppose that there exists $R > 0$ such that*

$$m_2(\mu) \leq R \quad \text{for } \mathbb{M}_2\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{H}). \quad (5.17)$$

Then we can find a totally convex R -Lipschitz function $\phi_R : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R}$ satisfying (5.11).

Proof. Let ϕ as in Claim 5 of Theorem 5.5 and let us set

$$\psi_R(\nu) := \phi^*(\nu) \quad \text{if } m_2(\nu) \leq R, \quad \psi_R(\nu) := +\infty \quad \text{otherwise.}$$

Clearly ψ_R is proper, totally convex and lower semicontinuous; the pair (ϕ, ψ_R) satisfies

$$\phi(\mu) + \psi_R(\nu) \geq [\mu, \nu], \quad \phi(\mu) + \psi_R(\nu) = [\mu, \nu] \quad \text{for } \Pi\text{-a.e. } (\mu, \nu) \in \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H}).$$

It is easy to check that $\phi_R := (\psi_R)^*$ still provides an optimal totally convex Kantorovich potential with $(\phi_R)^* = \psi_R$; ϕ_R is also R -Lipschitz since

$$\phi_R(\mu) = \sup \left\{ [\mu, \nu] - \phi^*(\nu) : m_2(\nu) \leq R \right\}. \quad (5.18)$$

\square

5.3 The structure of minimal geodesics in $\mathfrak{P}_2(\mathbb{H})$

We can apply the previous results to identify minimal geodesics in $\mathfrak{P}_2(\mathbb{H})$ and to characterize their transport structure. Let us first recall that a minimal geodesic $(\mathbb{M}_t)_{t \in [0,1]}$ connecting \mathbb{M}_0 to \mathbb{M}_1 is a curve in $\mathfrak{P}_2(\mathbb{H})$ satisfying

$$W_2(\mathbb{M}_s, \mathbb{M}_t) = |t - s| W_2(\mathbb{M}_0, \mathbb{M}_1) \quad \text{for every } s, t \in [0, 1]. \quad (5.19)$$

We also say that $\tilde{\mathbb{M}}$ is a t -intermediate point between \mathbb{M}_0 and \mathbb{M}_1 , $t \in (0, 1)$, if

$$W_2(\mathbb{M}_0, \tilde{\mathbb{M}}) = t W_2(\mathbb{M}_0, \mathbb{M}_1), \quad W_2(\tilde{\mathbb{M}}, \mathbb{M}_1) = (1 - t) W_2(\mathbb{M}_0, \mathbb{M}_1). \quad (5.20)$$

We will use the interpolating maps $\pi_t^{1 \rightarrow 2}$ of (1.11) and we will consider Borel maps defined in sets of the form (recall (4.25))

$$\mathcal{S}(\mathbf{H}, D) := \left\{ (x, \mu) \in \mathbf{H} \times D : x \in \text{supp } \mu \right\}, \quad D \text{ Borel subset of } \mathcal{P}_2(\mathbf{H}), \quad \mathcal{S}(\mathbf{H}) = \mathcal{S}(\mathbf{H}, \mathcal{P}_2(\mathbf{H})), \quad (5.21)$$

which are Borel subsets of $\mathbf{H} \times \mathcal{P}_2(\mathbf{H})$ [CSS23a, (4.23)].

Theorem 5.7 (Structure of minimal geodesics in $\mathfrak{P}_2(\mathbf{H})$). *Let $\mathbf{M}_0, \mathbf{M}_1 \in \mathfrak{P}_2(\mathbf{H})$, let ϕ, ϕ^* be a pair of optimal Kantorovich potentials for \mathbf{M}_i , and let $\mathbf{F} = \mathfrak{D}_t \phi \subset \mathcal{P}_2(\mathbf{H} \times \mathbf{H})$.*

1. *For every optimal random coupling law $\mathbf{P} \in \text{R}\Gamma_o(\mathbf{M}_0, \mathbf{M}_1)$ the curve*

$$\mathbf{M}_t := (\pi_t^{1 \rightarrow 2})_{\#} \mathbf{P}, \quad t \in [0, 1] \quad \text{is a minimal geodesic.} \quad (5.22)$$

2. *For every $t \in (0, 1)$ the set $\mathbf{F}_t := (\pi_t^{1 \rightarrow 2})_{\#}(\mathbf{F})$ is closed in $\mathcal{P}_2(\mathbf{H})$ and there exists two uniquely characterized continuous maps $\mathbf{f}_{t,i} : \mathcal{S}(\mathbf{H}, \mathbf{F}_t) \rightarrow \mathbf{H}$, $i = 0, 1$, inverting $(\pi_t^{1 \rightarrow 2})_{\#}$ in the sense that*

$$\gamma \in \mathbf{F}, \quad \mu = (\pi_t^{1 \rightarrow 2})_{\#} \gamma \quad \Rightarrow \quad \gamma = (\mathbf{f}_{t,0}(\cdot, \mu), \mathbf{f}_{t,1}(\cdot, \mu))_{\#} \mu. \quad (5.23)$$

Moreover, $\mathbf{f}_{t,i}(\cdot, \mu)$ is Lipschitz in $\text{supp}(\mu)$ and cyclically monotone in \mathbf{H} , the maps $\mathcal{F}_{t,i} : \mu \mapsto \mathbf{f}_{t,i}(\cdot, \mu)_{\#} \mu$ are Lipschitz from \mathbf{F}_t to $\mathcal{P}_2(\mathbf{H})$ and cyclically monotone in $\mathcal{P}_2(\mathbf{H})$.

3. *If $t \in (0, 1)$ and $\tilde{\mathbf{M}}$ is a t -intermediate point between \mathbf{M}_0 and \mathbf{M}_1 then $\text{supp}(\tilde{\mathbf{M}}) \subset \mathbf{F}_t$ and the formula*

$$\tilde{\mathbf{P}} = (\mathcal{G}_t)_{\#} \tilde{\mathbf{M}} \quad \text{where} \quad \mathcal{G}_t(\mu) := (\mathbf{f}_{t,0}(\cdot, \mu), \mathbf{f}_{t,1}(\cdot, \mu))_{\#} \mu \quad (5.24)$$

provides the unique $\tilde{\mathbf{P}} \in \text{R}\Gamma_o(\mathbf{M}_0, \mathbf{M}_1)$ such that $\tilde{\mathbf{M}} = (\pi_t^{1 \rightarrow 2})_{\#} \tilde{\mathbf{P}}$ and correspondingly the unique geodesic $(\mathbf{M}_s)_{s \in [0,1]}$ connecting \mathbf{M}_0 to \mathbf{M}_1 such that $\mathbf{M}_t = \tilde{\mathbf{M}}$. Moreover $\text{R}\Gamma_o(\tilde{\mathbf{M}}, \mathbf{M}_i)$, $i = 0, 1$, contains the unique element $\tilde{\mathbf{P}}_{t,i}$ given by

$$\begin{aligned} \tilde{\mathbf{P}}_{t,0} &= (\pi_t^1, \pi_t^{1 \rightarrow 2})_{\#} \tilde{\mathbf{P}} = (\mathcal{G}_{t,0})_{\#} \tilde{\mathbf{M}}, \quad \mathcal{G}_{t,0}(\mu) = (\mathbf{f}_{t,0}(\cdot, \mu), \mathbf{i})_{\#} \mu \\ \tilde{\mathbf{P}}_{t,1} &= (\pi_t^{1 \rightarrow 2}, \pi_t^2)_{\#} \tilde{\mathbf{P}} = (\mathcal{G}_{t,1})_{\#} \tilde{\mathbf{M}}, \quad \mathcal{G}_{t,1}(\mu) = (\mathbf{i}, \mathbf{f}_{t,1}(\cdot, \mu))_{\#} \mu \end{aligned} \quad (5.25)$$

which is concentrated on deterministic optimal couplings, and

$$\tilde{\mathbf{I}}_{t,i} = (\pi_{\#}^1, \pi_{\#}^2)_{\#} \tilde{\mathbf{P}}_{t,i} = (\text{Id} \times \mathcal{F}_{t,i})_{\#} \tilde{\mathbf{M}}, \quad i = 0, 1, \quad (5.26)$$

is the unique optimal coupling in $\Gamma_o(\tilde{\mathbf{M}}, \mathbf{M}_i)$.

4. *For every $t \in (0, 1)$ the conjugate functions (recall the definition of Moreau-Yosida regularization (3.16))*

$$\phi^t := \frac{1-t}{2} \mathbf{m}_2^2 + t\phi, \quad (\phi^t)^* = t(\hat{\phi}^*)_{1/t-1} \circ \mathfrak{d}_{t-1} \quad (5.27)$$

provide a pair of optimal Kantorovich potentials for \mathbf{M}_0 and any t -intermediate point $\tilde{\mathbf{M}}$ between \mathbf{M}_0 and \mathbf{M}_1 . Similarly

$$(1-t)\phi_{t/(1-t)} \circ \mathfrak{d}_{(1-t)^{-1}}, \quad (1-t)\phi^* + \frac{t}{2} \mathbf{m}_2^2 \quad (5.28)$$

is a pair of optimal Kantorovich potentials for $\tilde{\mathbf{M}}$ and \mathbf{M}_1 .

Remark 5.8. *The above Theorem recovers in a much more precise form various results that hold for $\mathcal{P}_2(\mathbf{X})$ in suitable classes of metric spaces \mathbf{X} , see [Vil09, Chap. 7]. See in particular the nonbranching property (stated in locally compact spaces) [Vil09, Corollary 7.32] for Claim 2, the ‘‘interpolation of prices’’ [Vil09, Theorem 7.36] concerning Claim 3, and the related bibliographical notes.*

Proof. Claim 1. For $0 \leq s < t \leq 1$ we define the maps $\pi_{s,t}^{1 \rightarrow 2} : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{H} \times \mathbb{H}$

$$\pi_{s,t}^{1 \rightarrow 2}(x_0, x_1) := (\pi_s^{1 \rightarrow 2}(x_0, x_1), \pi_t^{1 \rightarrow 2}(x_0, x_1));$$

we observe that

$$|\pi_s^{1 \rightarrow 2}(x_0, x_1) - \pi_t^{1 \rightarrow 2}(x_0, x_1)| = |t - s| \cdot |x_1 - x_0|$$

so that the random coupling $\mathbb{P}_{s,t} = (\pi_{s,t}^{1 \rightarrow 2})_{\#} \mathbb{P}$ belongs to $\text{R}\Gamma(\mathbb{M}_s, \mathbb{M}_t)$ and thus yields

$$\mathbb{W}_2(\mathbb{M}_s, \mathbb{M}_t) \leq |t - s| \mathbb{W}_2(\mathbb{M}_0, \mathbb{M}_1) \quad \text{if } 0 \leq s < t \leq 1.$$

The triangle inequality then yields (5.19).

Claim 2. Let us consider the Lagrangian lifting $\hat{\mathbf{F}} = \partial \hat{\phi} \subset \mathcal{H} \times \mathcal{H}$ of the total subdifferential $\mathbf{F} = \partial_t \phi$ of ϕ and let us set

$$I_t(X_0, X_1) := (1 - t)X_0 + tX_1, \quad \hat{\mathbf{F}}_t = I_t(\hat{\mathbf{F}}) = \left\{ \tilde{X} = (1 - t)X_0 + tX_1 : (X_0, X_1) \in \hat{\mathbf{F}} \right\}$$

which is clearly a set invariant by m.p.i. By monotonicity, if $\tilde{X} = I_t(X_0, X_1)$, $\tilde{X}' = I_t(X'_0, X'_1)$ for $(X_0, X_1), (X'_0, X'_1) \in \hat{\mathbf{F}}$ we have

$$|\tilde{X} - \tilde{X}'|^2 \geq (1 - t)|X_0 - X'_0|^2 + t|X_1 - X'_1|^2, \quad (5.29)$$

which shows that $\hat{\mathbf{F}}_t$ is closed and there exist Lipschitz maps $F_{t,i} : \hat{\mathbf{F}}_t \rightarrow \mathbb{H}$ such that

$$\tilde{X} = I_t(X_0, X_1), (X_0, X_1) \in \hat{\mathbf{F}} \quad \Rightarrow \quad X_i = F_{t,i}(\tilde{X}). \quad (5.30)$$

Since $F_{t,i}$ are also invariant by m.p.i., the general extension and representation Theorem 4.8 of [CSS25] shows that there is a unique pair of continuous maps $\mathbf{f}_{t,i} : \mathcal{S}(\mathbb{H}, \mathbf{F}_t) \rightarrow \mathbb{H}$ representing $F_{t,i}$ as

$$F_{t,i}[\tilde{X}](q) = \mathbf{f}_{t,i}(\tilde{X}(q), \iota(\tilde{X})) \quad (5.31)$$

with the properties stated in Claim 2. (5.31) clearly yields (5.23) since $\text{supp}(\mu) \subset \mathbf{F}_t \subset \iota(\hat{\mathbf{F}}_t)$. The cyclical monotonicity of $\mathbf{f}_{t,i}(\cdot, \mu)$ and of $\mathcal{F}_{t,i}$ follows from the corresponding cyclical monotonicity of the maps $F_{t,i}$ in \mathcal{H} , which in turn follows by the fact that they are the inverse of the cyclically monotone sets

$$\begin{aligned} \hat{\mathbf{F}}_{t,0} &:= (I_0, I_t)(\hat{\mathbf{F}}) = \left\{ (X_0, (1 - t)X_0 + tX_1) : (X_0, X_1) \in \hat{\mathbf{F}} \right\}, \\ \hat{\mathbf{F}}_{t,1} &:= (I_1, I_t)(\hat{\mathbf{F}}) = \left\{ (X_1, (1 - t)X_0 + tX_1) : (X_0, X_1) \in \hat{\mathbf{F}} \right\}. \end{aligned} \quad (5.32)$$

Claim 3. Let $\Pi_0 \in \Gamma_o(\mathbb{M}_0, \tilde{\mathbb{M}})$ and $\Pi_1 \in \Gamma_o(\tilde{\mathbb{M}}, \mathbb{M}_1)$. By the glueing Lemma we find a tri-plan $\Pi \in \Gamma(\mathbb{M}_0, \tilde{\mathbb{M}}, \mathbb{M}_1)$ such that $\pi_{\#}^{1,2} \Pi = \Pi_0$, $\pi_{\#}^{2,3} \Pi = \Pi_1$. Consider now the closed set

$$\Omega := \left\{ \gamma \in \mathcal{P}_2(\mathbb{H} \times \mathbb{H} \times \mathbb{H}) : \pi_{\#}^{1,2} \gamma \text{ and } \pi_{\#}^{2,3} \gamma \text{ belong to } \mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H}) \right\}. \quad (5.33)$$

Since the map $(\pi_{\#}^1 \times \pi_{\#}^2 \times \pi_{\#}^3) : \Omega \rightarrow (\mathcal{P}_2(\mathbb{H}))^3$ is surjective, we can find $\mathbb{P} \in \mathcal{P}_2(\Omega) \subset \mathcal{P}_2(\mathbb{H} \times \mathbb{H} \times \mathbb{H})$ such that $(\pi_{\#}^1 \times \pi_{\#}^2 \times \pi_{\#}^3)_{\#} \mathbb{P} = \Pi$.

We have

$$\pi_{\#}^{1,2}(\mathbb{P}) \in \text{R}\Gamma_o(\mathbb{M}_0, \tilde{\mathbb{M}}), \quad \pi_{\#}^{2,3}(\mathbb{P}) \in \text{R}\Gamma_o(\tilde{\mathbb{M}}, \mathbb{M}_1)$$

and by the elementary inequality $(a + b)^2 \leq \frac{1}{t}a^2 + \frac{1}{1-t}b^2$

$$\begin{aligned} \mathbb{W}_2^2(\mathbb{M}_0, \mathbb{M}_1) &\leq \int \int |x_1 - x_3|^2 d\gamma \mathbb{P}(\gamma) \\ &\leq \frac{1}{t} \int \int |x_1 - x_2|^2 d\gamma \mathbb{P}(\gamma) + \frac{1}{1-t} \int \int |x_2 - x_3|^2 d\gamma \mathbb{P}(\gamma) \\ &= \frac{1}{t} \mathbb{W}_2^2(\mathbb{M}_0, \tilde{\mathbb{M}}) + \frac{1}{1-t} \mathbb{W}_2^2(\tilde{\mathbb{M}}, \mathbb{M}_1) = \mathbb{W}_2^2(\mathbb{M}_0, \mathbb{M}_1) \end{aligned}$$

we deduce that $\tilde{\mathbb{P}} := \pi_{\#\#}^{1,3} \mathbb{P} \in \text{R}\Gamma_o(\mathbb{M}_0, \mathbb{M}_1)$ and

$$|x_1 - x_3|^2 = \frac{1}{t}|x_1 - x_2|^2 + \frac{1}{1-t}|x_2 - x_3|^2 \text{ on } \text{supp}(\gamma) \text{ for P-a.e. } \gamma,$$

so that P-a.e. γ is supported in the set

$$\mathbb{H}_t^3 := \left\{ (x_1, x_2, x_3) \in \mathbb{H}^3 : x_2 = (1-t)x_1 + tx_3 \right\} \quad (5.34)$$

and therefore $\tilde{\mathbb{M}} = (\pi_t^{1 \rightarrow 2})_{\#\#} \tilde{\mathbb{P}}$.

In order to show that $\tilde{\mathbb{P}}$ is unique, we observe that any random coupling in $\text{R}\Gamma_o(\mathbb{M}_0, \mathbb{M}_1)$ has support in \mathbf{F} so that $\tilde{\mathbb{M}}$ has support in \mathbf{F}_t and we can then apply (5.23) which yields (5.24). The above discussion also shows that $\text{R}\Gamma_o(\tilde{\mathbb{M}}, \mathbb{M}_i)$ and $\Gamma_o(\tilde{\mathbb{M}}, \mathbb{M}_i)$ are uniquely characterized by (5.25) and (5.26)

Claim 4. In order to check (5.27) (the argument for (5.28) is similar) we use the fact that $\tilde{\mathbb{M}}$ is supported in $\mathbf{F}_t = \iota(\hat{\mathbf{F}}_t)$. On the other hand, the set $\hat{\mathbf{F}}_{t,0}$ defined by (5.32) is the graph of the subdifferential of the Lagrangian lifting of ϕ^t

$$\hat{\phi}^t(X) := \frac{1-t}{2} \|X\|_{\mathcal{H}}^2 + t\hat{\phi}(X)$$

whose Legendre-Fenchel transform can be expressed in terms of the Moreau-Yosida regularization of $\hat{\phi}^*$ (recall (3.19) and (3.22)) and corresponds to the Lagrangian lifting of the functional $(\phi^t)^*$ given in (5.27). Since

$$\hat{\phi}^t(X) + (\hat{\phi}^t)^*(Y) = \langle X, Y \rangle_{\mathcal{H}} = [\iota(X), \iota(Y)] \text{ for every } (X, Y) \in \hat{\mathbf{F}}_{t,0}$$

we get the proof of the claim. \square

5.4 Lifting (laws of) random measures of $\mathfrak{P}_2(\mathbb{H})$ to measures on Lagrangian maps in $\mathcal{P}_2(\mathcal{H})$.

In the previous sections, we exploited the lifting technique of Proposition 5.1 in order to describe optimal couplings in $\mathcal{P}_2(\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H}))$ in terms of laws of random optimal couplings in $\mathfrak{P}_2(\mathbb{H} \times \mathbb{H})$.

There is another lifting technique which is induced by the 1-Lipschitz and surjective law map $\iota : \mathcal{H} \rightarrow \mathcal{P}_2(\mathbb{H})$. The corresponding push-forward transformation $\iota_{\#}$ still provides a surjective map from $\mathcal{P}_2(\mathcal{H})$ (the space of measures on Lagrangian maps of \mathcal{H}) to $\mathfrak{P}_2(\mathbb{H})$ (the space of (laws of) random measures), so that it is natural to study the relations between optimal transport problems in $\mathfrak{P}_2(\mathbb{H})$ and in $\mathcal{P}_2(\mathcal{H})$.

First of all, since ι is 1-Lipschitz, we observe that for every $\mathbf{m}_i \in \mathcal{P}_2(\mathcal{H})$, $i = 1, 2$, we have

$$\mathbf{M}_i = \iota_{\#} \mathbf{m}_i \quad \Rightarrow \quad \mathbb{W}_2(\mathbf{M}_1, \mathbf{M}_2) \leq \mathbb{W}_{2, \mathcal{H}}(\mathbf{m}_1, \mathbf{m}_2). \quad (5.35)$$

Similarly, it is not difficult to check that given a coupling $\mathbf{p} \in \mathcal{P}_2(\mathcal{H} \times \mathcal{H})$ and setting $\iota_i := \iota \circ \pi^i$, we have

$$\mathbb{H} = (\iota_1, \iota_2)_{\#} \mathbf{p} \quad \Rightarrow \quad \int \mathbb{w}_2^2(\mu_1, \mu_2) d\mathbb{H}(\mu_1, \mu_2) \leq \int \|X_1 - X_2\|_{\mathcal{H}}^2 d\mathbf{p}(X_1, X_2). \quad (5.36)$$

Eventually, still starting from $\mathbf{p} \in \mathcal{P}_2(\mathcal{H} \times \mathcal{H})$ and using $\iota^2(X, Y) := (X, Y)_{\#} \mathbb{M}$, we have

$$\mathbb{P} = \iota_{\#}^2 \mathbf{p} \quad \Rightarrow \quad \int \int |x_1 - x_2|^2 d\gamma d\mathbb{P}(\gamma) = \int \|X_1 - X_2\|_{\mathcal{H}}^2 d\mathbf{p}(X_1, X_2). \quad (5.37)$$

Since ι^2 is surjective from $\mathcal{H} \times \mathcal{H}$ to $\mathcal{P}_2(\mathbb{H} \times \mathbb{H})$, $\iota_{\#}^2$ is surjective as well, so that given $\mathbb{P} \in \mathcal{P}_2(\mathcal{P}_2(\mathbb{H} \times \mathbb{H}))$ it is always possible to find a lifting \mathbf{p}^ℓ such that $\mathbb{P} = \iota_{\#}^2 \mathbf{p}^\ell$.

If we want to lift Π so that (5.36) holds as an equality, we can first select $\mathbf{P} \in \mathcal{P}_2(\mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H}))$ so that (5.4) holds: the lifting \mathbf{p}^ℓ satisfies the identity

$$\Pi = (\iota_1, \iota_2)_\# \mathbf{p}^\ell, \quad \int \mathfrak{w}_2^2(\mu_1, \mu_2) d\Pi(\mu_1, \mu_2) = \int \|X_1 - X_2\|_{\mathcal{H}}^2 d\mathbf{p}^\ell(X_1, X_2). \quad (5.38)$$

If moreover $\Pi \in \Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$ then $\mathbf{P} \in \text{R}\Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$ and setting $\mathbf{m}_1^\ell = \pi_\#^1 \mathbf{p}^\ell, \mathbf{m}_2^\ell = \pi_\#^2 \mathbf{p}^\ell$ we get

$$\mathbf{M}_i = \iota_\# \mathbf{m}_i^\ell, \quad \mathbf{W}_2(\mathbf{M}_1, \mathbf{M}_2) = \mathbf{W}_{2,\mathcal{H}}(\mathbf{m}_1^\ell, \mathbf{m}_2^\ell). \quad (5.39)$$

We recap the above argument in the next proposition.

Proposition 5.9 (From random OT to OT in \mathcal{H}). *For every pair $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{P}_2(\mathbb{H})$ there exists a pair $\mathbf{m}_1^\ell, \mathbf{m}_2^\ell \in \mathcal{P}_2(\mathcal{H})$ such that (5.39) holds, so that*

$$\mathbf{W}_2(\mathbf{M}_1, \mathbf{M}_2) = \min \left\{ \mathbf{W}_{2,\mathcal{H}}(\mathbf{m}_1, \mathbf{m}_2) : \mathbf{m}_i \in \mathcal{P}_2(\mathcal{H}), \iota_\# \mathbf{m}_i = \mathbf{M}_i \right\}. \quad (5.40)$$

If $\mathbf{m}_1^\ell, \mathbf{m}_2^\ell \in \mathcal{P}_2(\mathcal{H})$ are minimizers of (5.40) and $\mathbf{p}^\ell \in \Gamma_o(\mathbf{m}_1^\ell, \mathbf{m}_2^\ell)$ in $\mathcal{P}_2(\mathcal{H} \times \mathcal{H})$ then $\mathbf{P} = \iota_\#^2 \mathbf{p}^\ell \in \text{R}\Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$ and $\Pi = (\pi_\#^1, \pi_\#^2) \mathbf{P} = (\iota_1, \iota_2)_\# \mathbf{p}^\ell \in \Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$.

We want to highlight that the optimal lifted measures $\mathbf{m}_1^\ell, \mathbf{m}_2^\ell$ given by the previous proposition typically depend on both the measures $\mathbf{M}_1, \mathbf{M}_2$ and in general we cannot fix an arbitrary \mathbf{m}_1^ℓ such that $\iota_\# \mathbf{m}_1^\ell = \mathbf{M}_1$. We want to find a sufficient condition on $\mathbf{M}_1, \mathbf{M}_2$ for which the following property holds:

$$\text{for every } \mathbf{m}_1 \in \mathcal{P}_2(\mathcal{H}) \text{ such that } \iota_\# \mathbf{m}_1 = \mathbf{M}_1 \text{ there exists } \mathbf{m}_2 \in \mathcal{P}_2(\mathcal{H}) \text{ with} \quad (5.41)$$

$$\iota_\# \mathbf{m}_2 = \mathbf{M}_2, \quad \mathbf{W}_2(\mathbf{M}_1, \mathbf{M}_2) = \mathbf{W}_{2,\mathcal{H}}(\mathbf{m}_1, \mathbf{m}_2).$$

A crucial role in this respect is played by the set of ‘‘deterministic’’ couplings $\mathcal{P}_{2,o}^{\text{det}}(\mathbb{H} \times \mathbb{H})$ which are concentrated on maps:

$$\mathcal{P}_2^{\text{det}}(\mathbb{H} \times \mathbb{H}) := \left\{ (i \times \mathbf{f})_\# \mu : \mu \in \mathcal{P}_2(\mathbb{H}), \mathbf{f} \in L^2(\mathbb{H}, \mu; \mathbb{H}) \right\}, \quad (5.42)$$

$$\mathcal{P}_{2,o}^{\text{det}}(\mathbb{H} \times \mathbb{H}) := \mathcal{P}_2^{\text{det}}(\mathbb{H} \times \mathbb{H}) \cap \mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H}).$$

First of all, we will show a simple condition for which there exists an optimal coupling $\mathbf{P} \in \text{R}\Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$ which is concentrated on $\mathcal{P}_{2,o}^{\text{det}}(\mathbb{H} \times \mathbb{H})$.

Lemma 5.10. *An optimal coupling $\Pi \in \mathcal{P}_{2,o}(\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H}))$ satisfies the property*

$$\text{for } \Pi\text{-a.e. } (\mu_1, \mu_2) \quad \Gamma_o(\mu_1, \mu_2) \cap \mathcal{P}_2^{\text{det}}(\mathbb{H} \times \mathbb{H}) \neq \emptyset \quad (5.43)$$

if and only if

$$\text{there exists } \mathbf{P} \in \mathcal{P}_2(\mathcal{P}_{2,o}^{\text{det}}(\mathbb{H} \times \mathbb{H})) \text{ such that } \Pi = (\pi_\#^1, \pi_\#^2)_\# \mathbf{P}. \quad (5.44)$$

Proof. Let us set $\mathcal{O} := \left\{ (\mu_1, \mu_2) \in \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H}) : \Gamma_o(\mu_1, \mu_2) \cap \mathcal{P}_2^{\text{det}}(\mathbb{H} \times \mathbb{H}) \neq \emptyset \right\}$. We can equivalently characterize \mathcal{O} as the image of the Borel set $\mathcal{P}_{2,o}^{\text{det}}(\mathbb{H} \times \mathbb{H})$ through the continuous map $\pi_\#^1 \times \pi_\#^2$, so that \mathcal{O} is a Souslin (and therefore universally measurable) set.

(5.43) just says that Π is concentrated on \mathcal{O} , and therefore its equivalence with (5.44) follows by Theorem 2.1. \square

Theorem 5.11. *Let $\Pi \in \Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$ satisfy (5.43), let \mathbf{P} as in (5.44), and let $\mathbf{m}_1 \in \mathcal{P}_2(\mathcal{H})$ such that $\iota_\# \mathbf{m}_1 = \mathbf{M}_1$. Then there exists $\mathbf{p}^\ell \in \mathcal{P}_{2,o}(\mathcal{H} \times \mathcal{H})$ such that $\pi_\#^1 \mathbf{p}^\ell = \mathbf{m}_1$ and $\iota_\#^2 \mathbf{p}^\ell = \mathbf{P}$, so that, in particular, $(\iota_1, \iota_2)_\# \mathbf{p}^\ell = \Pi$ and setting $\mathbf{m}_2 = \pi_\#^2 \mathbf{p}^\ell$ we have*

$$\iota_\# \mathbf{m}_2 = \mathbf{M}_2, \quad \mathbf{W}_2(\mathbf{M}_1, \mathbf{M}_2) = \mathbf{W}_{2,\mathcal{H}}(\mathbf{m}_1, \mathbf{m}_2). \quad (5.45)$$

Proof. Let us first consider the closed subset \mathcal{A} of $\mathcal{H} \times \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H} \times \mathbb{H})$ defined by

$$\mathcal{A} := \left\{ (X, \mu, \gamma) \in \mathcal{H} \times \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H} \times \mathbb{H}) : \iota(X) = \mu = \pi_{\#}^1(\gamma) \right\}, \quad (5.46)$$

and the map $A : \mathcal{H} \times \mathcal{H} \rightarrow \mathcal{A}$

$$A(X, Y) := (X, \iota(X), \iota^2(X, Y)) = (X, X_{\#}\mathbb{M}, (X, Y)_{\#}\mathbb{M}). \quad (5.47)$$

A is continuous but in general it is not surjective. However, it is not difficult to check that the image of A contains the Borel set

$$\begin{aligned} \mathcal{A}^{\det} &:= \left\{ (X, \mu, \gamma) \in \mathcal{H} \times \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2^{\det}(\mathbb{H} \times \mathbb{H}) : \iota(X) = \mu = \pi_{\#}^1(\gamma) \right\} \\ &= \mathcal{A} \cap \left(\mathcal{H} \times \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2^{\det}(\mathbb{H} \times \mathbb{H}) \right). \end{aligned} \quad (5.48)$$

In fact, if $(X, \mu, \gamma) \in \mathcal{A}^{\det}$ then $\mu = \pi_{\#}^1\gamma$ and we can find a map $f_{\gamma} \in L^2(\mathbb{H}, \mu; \mathbb{H})$ such that $\gamma = (i \times f_{\gamma})_{\#}\mu$. Defining $Y := f_{\gamma} \circ X$ we immediately see that $\iota^2(X, Y) = \gamma$, so that $A(X, Y) = (X, \mu, \gamma)$.

Let us now set $\bar{\mathbf{m}}_1 := (\text{Id} \times \iota)_{\#}\mathbf{m}_1 \in \mathcal{P}_2(\mathcal{H} \times \mathcal{P}_2(\mathbb{H}))$ and $\bar{\mathbf{P}} := (\pi_{\#}^1 \times \text{Id})_{\#}\mathbf{P} \in \mathcal{P}_2(\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2^{\det}(\mathbb{H} \times \mathbb{H}))$. By assumption

$$\pi_{\#}^2 \bar{\mathbf{m}}_1 = \iota_{\#}\mathbf{m}_1 = \mathbf{M}_1, \quad \pi_{\#}^1 \bar{\mathbf{P}} = (\pi_{\#}^1)_{\#}\mathbf{P} = \mathbf{M}_1$$

so that, by the gluing Lemma, we can find a plan $\mathbf{Q} \in \mathcal{P}_2(\mathcal{H} \times \mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2^{\det}(\mathbb{H} \times \mathbb{H}))$ such that $\pi_{\#}^1{}^2 \mathbf{Q} = \bar{\mathbf{m}}_1$ and $\pi_{\#}^1{}^3 \mathbf{Q} = \bar{\mathbf{P}}$. By construction, $\bar{\mathbf{m}}_1$ is concentrated on the set

$$\left\{ (X, \mu) \in \mathcal{H} \times \mathcal{P}_2(\mathbb{H}) : \iota(X) = \mu \right\}$$

and $\bar{\mathbf{P}}$ is concentrated on the set

$$\left\{ (\mu, \gamma) \in \mathcal{H} \times \mathcal{P}_2^{\det}(\mathbb{H} \times \mathbb{H}) : \mu = \pi_{\#}^1\gamma \right\}$$

we deduce that \mathbf{Q} is concentrated on \mathcal{A}^{\det} .

By Theorem 2.1 we can find a probability measure $\mathbf{p}^{\ell} \in \mathcal{P}_2(\mathcal{H} \times \mathcal{H})$ such that $A_{\#}\mathbf{p}^{\ell} = \mathbf{Q}$. By the very definition of A we get

$$\pi_{\#}^1 \mathbf{p}^{\ell} = (\pi^1 \circ A)_{\#}\mathbf{p}^{\ell} = \pi_{\#}^1 \mathbf{Q} = \mathbf{m}_1, \quad \iota_{\#}^2 \mathbf{p}^{\ell} = (\pi^3 \circ A)_{\#}\mathbf{p}^{\ell} = \pi_{\#}^3 \mathbf{Q} = \mathbf{P}$$

and the thesis follows. \square

Corollary 5.12. *Let $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{P}_2(\mathbb{H})$ and let us suppose that \mathbf{M}_1 is concentrated on $\mathcal{P}_2^{\ell}(\mathbb{H})$. Then (5.41) holds.*

6 Random Gaussian-null sets and strict Monge formulation of OT via nonlocal totally cyclically monotone fields

In this last section we want to address the uniqueness and the Monge formulation of the L^2 -OT problem in $\mathcal{P}_2(\mathbb{H})$. These questions can be settled at the usual level of couplings of (laws of) random measures. We then seek for conditions on $\mathbf{M}_i \in \mathcal{P}_2(\mathbb{H})$ ensuring that the class of optimal coupling $\Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$ contains a unique element Π which is concentrated on the graph of a Borel map $\mathcal{F} : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathcal{P}_2(\mathbb{H})$, so that

$$\Pi = (\text{Id} \times \mathcal{F})_{\#}\mathbf{M}_1, \quad \mathbf{M}_2 = \mathcal{F}_{\#}\mathbf{M}_1, \quad W_2^2(\mathbf{M}_1, \mathbf{M}_2) = \int_{\mathcal{P}_2(\mathbb{H})} w_2^2(\mu, \mathcal{F}(\mu)) d\mathbf{M}_1(\mu); \quad (6.1)$$

this property is equivalent to asking for $\Pi \in \mathcal{P}_2^{\det}(\mathcal{P}_2(\mathbb{H}) \times \mathcal{P}_2(\mathbb{H}))$.

A second formulation involves random couplings and provides a more refined description of \mathcal{F} : we look for conditions ensuring that $\text{R}\Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$ contains a unique element \mathbf{P} that is concentrated on the graph of a deterministic totally cyclically monotone field $\mathbf{f} : \mathbb{H} \times \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{H}$.

6.1 The strict Monge formulation

In order to describe such a construction, we first observe that every $\mathbf{P} \in \mathfrak{P}_2(\mathbb{H} \times \mathbb{H})$ can be disintegrated with respect to $\mathbf{M}_1 = \pi_{\sharp}^1 \mathbf{P}$ to obtain a Borel family $\mathbf{P}_\mu \in \mathfrak{P}_2(\mathbb{H} \times \mathbb{H})$ indexed by $\mu \in \mathcal{P}_2(\mathbb{H})$ and concentrated on the set of couplings $\Gamma(\mu) := \{\gamma \in \mathcal{P}_2(\mathbb{H} \times \mathbb{H}) : \pi_{\sharp}^1 \gamma = \mu\}$ for \mathbf{M}_1 -a.e. $\mu \in \mathcal{P}_2(\mathbb{H})$. When $\text{supp}(\mathbf{P}) \subset \mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H})$ (as in the case of optimal couplings) then $\text{supp}(\mathbf{P}_\mu) \subset \Gamma_o(\mu) := \Gamma(\mu) \cap \mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H})$.

If $\Pi = (\pi_{\sharp}^1, \pi_{\sharp}^2)_{\sharp} \mathbf{P}$ then \mathbf{P}_μ characterizes the family of measures in $\mathfrak{P}_2(\mathbb{H})$ arising from the disintegration $(\Pi_\mu)_{\mu \in \mathcal{P}_2(\mathbb{H})}$ of Π with respect to its first marginal through the formula

$$\Pi_\mu = \pi_{\sharp}^2 \mathbf{P}_\mu \quad \text{for } \mathbf{M}_1\text{-a.e. } \mu. \quad (6.2)$$

If π_{\sharp}^1 is essentially injective with respect to \mathbf{P} then there is a Borel map $\mathcal{G} : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathfrak{P}_2(\mathbb{H} \times \mathbb{H})$ such that for \mathbf{M}_1 -a.e. μ \mathbf{P}_μ is concentrated on a unique coupling $\gamma = \mathcal{G}(\mu)$ with first marginal μ , so that Π is deterministic as in (6.1) and we have

$$\mathbf{P}_\mu = \delta_{\mathcal{G}(\mu)}, \quad \mathcal{G}(\mu) \in \Gamma(\mu, \mathcal{F}(\mu)), \quad \mathcal{F}(\mu) = \pi_{\sharp}^2 \mathcal{G}(\mu), \quad \mathbf{M}_2 = (\pi_{\sharp}^2 \circ \mathcal{G})_{\sharp} \mathbf{M}_1 = \pi_{\sharp}^2 \mathcal{G}_{\sharp} \mathbf{M}_1. \quad (6.3)$$

We can then apply to \mathbf{P}_μ the disintegration map \mathcal{K} , obtaining the decomposition $\mathcal{G}(\mu) = \mu \otimes \kappa_{x,\mu}$, $\kappa_{x,\mu} = \mathcal{K}(x, \mathcal{G}(\mu))$. If \mathbf{P} is concentrated on $\mathcal{P}_2^{\text{det}}(\mathbb{H} \times \mathbb{H})$ then \mathbf{P}_μ is concentrated on $\mathcal{P}_2^{\text{det}}(\mathbb{H} \times \mathbb{H})$ as well, so that $\kappa_{x,\mu} = \delta_{\mathbf{f}(x,\mu)}$ for some Borel map $\mathbf{f} : \mathbb{H} \times \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{H}$. We then obtain

$$\mathcal{G}(\mu) = (\mathbf{i} \times \mathbf{f}(\cdot, \mu))_{\sharp} \mu, \quad \mathcal{F}(\mu) = \mathbf{f}(\cdot, \mu)_{\sharp} \mu. \quad (6.4)$$

Since $\mathcal{F}(\mu) \in \mathcal{P}_2(\mathbb{H})$ and $\mathbf{M}_2 \in \mathfrak{P}_2(\mathbb{H})$ we have

$$\begin{aligned} \int_{\mathbb{H}} |\mathbf{f}(x, \mu)|^2 d\mu(x) &= m_2^2(\mathcal{F}(\mu)) < \infty \quad \text{for } \mathbf{M}_1\text{-a.e. } \mu \in \mathcal{P}_2(\mathbb{H}) \\ \int_{\mathcal{P}_2(\mathbb{H})} \left(\int_{\mathbb{H}} |\mathbf{f}(x, \mu)|^2 d\mu(x) \right) d\mathbf{M}_1(\mu) &= \int_{\mathcal{P}_2(\mathbb{H})} m_2^2(\mathcal{F}(\mu)) d\mathbf{M}_1(\mu) = \mathbf{M}_2^2(\mathbf{M}_2) < \infty. \end{aligned} \quad (6.5)$$

It is then convenient to represent \mathbf{f} as a \mathbb{H} -valued L^2 map of the unfolded measure

$$\bar{\mathbf{M}}_1 := \int (\mu \otimes \delta_\mu) d\mathbf{M}_1(\mu) \in \mathcal{P}_2(\mathbb{H} \times \mathcal{P}_2(\mathbb{H})), \quad (6.6)$$

which satisfies $\pi_{\sharp}^2 \bar{\mathbf{M}}_1 = \mathbf{M}_1$. In fact if $\mathbf{f} \in L^2(\bar{\mathbf{M}}_1; \mathbb{H})$ then

$$\|\mathbf{f}\|_{L^2(\bar{\mathbf{M}}_1; \mathbb{H})}^2 = \int_{\mathcal{P}_2(\mathbb{H})} \int_{\mathbb{H}} |\mathbf{f}(x, \mu)|^2 d\mu d\mathbf{M}_1(\mu) < \infty \quad (6.7)$$

and we can represent the corresponding unfolded measure $\bar{\mathbf{M}}_2$ as

$$\bar{\mathbf{M}}_2 = (\mathbf{f}, \mathcal{F})_{\sharp} \bar{\mathbf{M}}_1, \quad \bar{\mathbf{M}}_i = \int (\mu \otimes \delta_\mu) d\mathbf{M}_i(\mu), \quad (6.8)$$

since

$$\begin{aligned} \int \xi(y, \nu) d\bar{\mathbf{M}}_2(y, \nu) &= \int \left(\int \xi(y, \nu) d\nu(y) \right) d\mathbf{M}_2(\nu) \\ &= \int \left(\int \xi(y, \mathcal{F}(\mu)) d(\mathcal{F}(\mu))(y) \right) d\mathbf{M}_1(\mu) \\ &= \int \left(\int \xi(\mathbf{f}(x, \mu), \mathcal{F}(\mu)) d\mu(x) \right) d\mathbf{M}_1(\mu) \\ &= \int \xi(\mathbf{f}(x, \mu), \mathcal{F}(\mu)) d\bar{\mathbf{M}}_1(x, \mu). \end{aligned}$$

The above remarks justify the following definition.

Definition 6.1 (Fully deterministic random couplings). *We say that a random coupling law $\mathbb{P} \in \mathfrak{P}_2(\mathbb{H} \times \mathbb{H})$ is fully deterministic if*

$$\pi_{\sharp}^1 \text{ is } \mathbb{P}\text{-essentially injective and } \mathbb{P} \text{ is concentrated on } \mathfrak{P}_2^{\text{det}}(\mathbb{H} \times \mathbb{H}). \quad (6.9)$$

We denote by $\mathfrak{P}_2^{\text{det}}(\mathbb{H} \times \mathbb{H})$ the set of fully deterministic random couplings.

Lemma 6.2 (Representation of fully deterministic random couplings). *A random coupling law $\mathbb{P} \in \mathfrak{P}_2(\mathbb{H} \times \mathbb{H})$ with first random marginal $\mathbb{M}_1 = \pi_{\sharp\sharp}^1 \mathbb{P}$ is fully deterministic if and only if there exists a Borel map $\mathbf{f} \in L^2(\bar{\mathbb{M}}_1; \mathbb{H})$ such that*

$$\mathbb{P} = \int \delta_{(\mathbf{i} \times \mathbf{f}(\cdot, \mu))_{\sharp} \mu} d\mathbb{M}_1(\mu) = \mathcal{G}_{\sharp} \mathbb{M}_1, \quad \mathcal{G}(\mu) := (\mathbf{i} \times \mathbf{f}(\cdot, \mu))_{\sharp} \mu. \quad (6.10)$$

In this case setting $\mathcal{F}(\mu) := \mathbf{f}(\cdot, \mu)_{\sharp} \mu$ we have

$$\begin{aligned} \mathbb{M}_2 &= \pi_{\sharp\sharp}^2 \mathbb{P} = \mathcal{F}_{\sharp} \mathbb{M}_1, \quad \bar{\mathbb{M}}_2 = (\mathbf{f}, \mathcal{F})_{\sharp} \bar{\mathbb{M}}_1 \\ \int_{\mathfrak{P}_2(\mathbb{H} \times \mathbb{H})} \left(\int_{\mathbb{H} \times \mathbb{H}} |x - y|^2 d\gamma \right) d\mathbb{P}(\gamma) &= \int_{\mathbb{H} \times \mathfrak{P}_2(\mathbb{H})} |\mathbf{f}(x, \mu) - x|^2 d\mu(x) d\mathbb{M}_1(\mu). \end{aligned} \quad (6.11)$$

Using the unfolding $\bar{\mathbb{P}}$ of \mathbb{P} we can also express (6.11) as

$$\bar{\mathbb{P}} = (\mathbf{i} \times \mathbf{f}, \mathcal{G})_{\sharp} \bar{\mathbb{M}}_1, \quad \int |x - y|^2 d\bar{\mathbb{P}}(x, y, \gamma) = \int |\mathbf{f}(x, \mu) - x|^2 d\bar{\mathbb{M}}_1(x, \mu). \quad (6.12)$$

Proof. Because of the previous digression, we have just to show the converse direction: given a Borel map $\mathbf{f} \in L^2(\bar{\mathbb{M}}_1; \mathbb{H})$ the coupling \mathbb{P} given by (6.10) is well defined, i.e. the map $\mathcal{G} : \mathfrak{P}(\mathbb{H}) \rightarrow \mathfrak{P}(\mathbb{H} \times \mathbb{H})$ is Borel. This property follows by standard argument, see e.g. [PS25, Lemma D.2 and Corollary D.7]; here is a self-contained discussion.

Let us first consider for a bounded Borel real function $z : \mathbb{H} \times \mathfrak{P}(\mathbb{H}) \rightarrow \mathbb{R}$ the functional $\mathcal{Z} : \mathfrak{P}(\mathbb{H}) \rightarrow \mathbb{R}$,

$$\mathcal{Z}(\mu) := \int z(x, \mu) d\mu \quad \mu \in \mathfrak{P}(\mathbb{H}), \quad (6.13)$$

and let us call \mathcal{B} the class of functions z for which \mathcal{Z} is Borel.

Clearly \mathcal{B} contains all the bounded and continuous functions $z \in C_b(\mathbb{H} \times \mathfrak{P}(\mathbb{H}))$ (for which \mathcal{Z} is continuous). It is also easy to check that \mathcal{B} is closed with respect to uniform and monotone limits. By the functional monotone class Theorem [Bog07, Theorem 2.12.9] we deduce that \mathcal{B} contains all the bounded Borel functions.

It follows that for every bounded continuous (or even Borel) map $\zeta : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{R}$ the map

$$\mu \mapsto \int_{\mathbb{H} \times \mathbb{H}} \zeta d\mathcal{G}(\mu) = \int \zeta(x, \mathbf{f}(x, \mu)) d\mu \quad (6.14)$$

is Borel, so the map \mathcal{G} is Borel as well, since the functionals $\gamma \mapsto \int \zeta, d\gamma$, $\zeta \in C_b(\mathbb{H} \times \mathbb{H})$ generates the weak (Polish) topology of $\mathfrak{P}(\mathbb{H} \times \mathbb{H})$.

Now, it is immediate to conclude that \mathbb{P} is fully deterministic, since \mathcal{G} is injective and maps $\mathfrak{P}_2(\mathbb{H})$ to $\mathfrak{P}_2^{\text{det}}(\mathbb{H} \times \mathbb{H})$. \square

We end up with the following stronger Monge formulation of OT problem between (laws of) random measures.

Problem 6.3 (Strict Monge formulation). *Given $\mathbb{M}_1, \mathbb{M}_2 \in \mathfrak{P}_2(\mathbb{H})$ find a Borel map $\mathbf{f} \in L^2(\bar{\mathbb{M}}_1; \mathbb{H})$ such that setting $\mathcal{F}(\mu) := \mathbf{f}(\cdot, \mu)_{\sharp} \mu$ we have*

$$\mathcal{F}_{\sharp} \mathbb{M}_1 = \mathbb{M}_2 \quad \text{and} \quad W_2^2(\mathbb{M}_1, \mathbb{M}_2) = \int_{\mathbb{H} \times \mathfrak{P}_2(\mathbb{H})} |\mathbf{f}(x, \mu) - x|^2 d\bar{\mathbb{M}}_1(x, \mu). \quad (6.15)$$

Notice that the transformation $\mathbf{f} \times \mathcal{F} : (x, \mu) \rightarrow (\mathbf{f}(x, \mu), \mathbf{f}(\cdot, \mu)_{\#}\mu)$ maps $\mathbb{H} \times \mathcal{P}_2(\mathbb{H})$ to itself and satisfies

$$(\mathbf{f}, \mathcal{F})_{\#}\bar{\mathbf{M}}_1 = \bar{\mathbf{M}}_2 \quad (6.16)$$

where, as usual, $\mathcal{F}(\mu) = \mathbf{f}(\cdot, \mu)_{\#}\mu$ and $\bar{\mathbf{M}}_i = \int \mu \otimes \delta_{\mu} d\mathbf{M}_i(\mu)$.

It could seem that the strict Monge formulation is considerably more demanding than the usual Monge formulation expressed by (6.1). For example in the extreme case when $\mathbf{M}_i = \delta_{\mu_i}$ are Dirac masses concentrated in two measures $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{H})$ clearly there is just one solution to the Monge formulation (6.1) but there could be many (or even no) solutions to the strict Monge formulation, which reduces to the usual Optimal Transport problem between μ_1 and μ_2 in $\mathcal{P}_2(\mathbb{H})$.

However, when we look for conditions on \mathbf{M}_1 which guarantee that (6.1) is solvable for every target measures \mathbf{M}_2 then the two formulations are equivalent and force uniqueness of solutions, as the following result shows.

Theorem 6.4 (Monge vs strict Monge). *Given $\mathbf{M}_1 \in \mathfrak{P}_2(\mathbb{H})$, the following two properties are equivalent:*

1. for every $\mathbf{M}_2 \in \mathfrak{P}_2(\mathbb{H})$ there exists a Borel map $\mathcal{F} = \mathcal{F}_{\mathbf{M}_2} : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathcal{P}_2(\mathbb{H})$ (depending on \mathbf{M}_2) solving the OT problem in Monge form (6.1);
2. for every $\mathbf{M}_2 \in \mathfrak{P}_2(\mathbb{H})$ there exists a Borel map $\mathbf{f} = \mathbf{f}_{\mathbf{M}_2} \in L^2(\bar{\mathbf{M}}_1; \mathbb{H})$ (depending on \mathbf{M}_2) solving the OT problem in the strict Monge form (6.15).

In both cases, for every $\mathbf{M}_2 \in \mathfrak{P}_2(\mathbb{H})$ the set of optimal couplings $\Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$ contains the unique element $\Pi = \Pi_{\mathbf{M}_2} = (\text{Id} \times \mathcal{F}_{\mathbf{M}_2})_{\#}\mathbf{M}_1$ and, for Π -a.e. (μ_1, μ_2) , the set $\Gamma_o(\mu_1, \mu_2)$ contains the unique deterministic coupling $\gamma = (\mathbf{i} \times \mathbf{f}_{\mathbf{M}_2}(\cdot, \mu_1))_{\#}\mu_1$ so that \mathbf{f} corresponds to the (unique) solution of the strict Monge formulation given in Problem 6.3. We also have $\mathbf{f} = \nabla_W \phi$ for every optimal Kantorovich potential ϕ (recall (4.23)).

Finally, every Lipschitz totally convex function ϕ is W -differentiable at \mathbf{M}_1 -a.e. μ , according to (4.24).

Proof. Since $2. \Rightarrow 1.$, it is sufficient to prove that 1. implies 2. and all the further properties stated by the Theorem.

We thus fix \mathbf{M}_2 , an optimal Kantorovich potential ϕ for the pair $\mathbf{M}_1, \mathbf{M}_2$ with $\mathbf{F} = \partial_t \phi$ and we denote by $\mathbf{f}^\circ : \mathbb{H} \times \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{H}$ a Borel version of the minimal section of \mathbf{F} (as in Proposition 4.2) and we set $\mathcal{F}^\circ(\mu) := \mathbf{f}^\circ(\cdot, \mu)_{\#}\mu$.

We then select an optimal random coupling law $\mathbf{P} \in \text{R}\Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$ with $\Pi = (\pi_{\#}^1, \pi_{\#}^2)_{\#}\mathbf{P} \in \Gamma_o(\mathbf{M}_1, \mathbf{M}_2)$; since \mathbf{P} is optimal, $\text{supp } \mathbf{P} \subset \mathbf{F}$.

By the minimality of \mathbf{f}° we have $\mathbf{f}^\circ \in L^2(\bar{\mathbf{M}}_1; \mathbb{H})$ since

$$\begin{aligned} W_2^2(\mathbf{M}_1, \mathbf{M}_2) &= \int \int |x_2 - x_1|^2 d\gamma(x_1, x_2) d\mathbf{P}(\gamma) \geq \int \int |\mathbf{f}^\circ(x_1, \pi_{\#}^1 \gamma) - x_1|^2 d\gamma(x_1, x_2) d\mathbf{P}(\gamma) \\ &= \int \int |\mathbf{f}^\circ(x_1, \mu) - x_1|^2 d\bar{\mathbf{M}}_1(x, \mu) \end{aligned}$$

We then introduce

$$\mathbf{P}' := \frac{1}{2}\mathbf{P} + \frac{1}{2}\mathcal{G}_{\#}\mathbf{M}_1, \quad \mathcal{G}^\circ(\mu) := (\mathbf{i} \times \mathbf{f}^\circ(\cdot, \mu))_{\#}\mu, \quad \mathbf{M}'_2 := \pi_{\#}^2 \mathbf{P}'.$$

Since $\text{supp } \mathbf{P}' \subset \mathbf{F}$ we have $\mathbf{P}' \in \text{R}\Gamma_o(\mathbf{M}_1, \mathbf{M}'_2)$ and its disintegration with respect to $\pi_{\#}^1$ can be expressed via the corresponding disintegration \mathbf{P}'_μ of \mathbf{P}' by

$$\mathbf{P}'_\mu = \frac{1}{2}\mathbf{P}_\mu + \frac{1}{2}\delta_{\mathcal{G}^\circ(\mu)}.$$

We then select the midpoint $\tilde{M} := (\pi_t^{1 \rightarrow 2})_{\#} P'$ between M_1 and M'_2 induced by P' corresponding to $t = 1/2$. By Theorem 5.7 we know that $\tilde{P} := (\pi^1, \pi_t^{1 \rightarrow 2})_{\#} P'$ is the unique element of $\text{R}\Gamma_o(M_1, \tilde{M})$, and $\tilde{II} = (\pi_{\#}^1, \pi_{\#}^2) \tilde{P}$ is the unique element of $\Gamma_o(M_1, \tilde{M})$ so that by assumption $\tilde{II} = (\text{Id} \times \tilde{\mathcal{F}})_{\#} M_1$ for $\tilde{\mathcal{F}} = \mathcal{F}_{\tilde{M}}$. On the other hand setting

$$P_{\mu,t} = (\pi^1, \pi_t^{1 \rightarrow 2})_{\#} P_{\mu}, \quad II_{\mu,t} = \pi_{\#}^2 P_{\mu,t}, \quad \mathcal{G}_t^{\circ}(\mu) = (\pi^1, \pi_t^{1 \rightarrow 2})_{\#} \mathcal{G}^{\circ}(\mu), \quad \mathcal{F}_t^{\circ}(\mu) = (\pi_t^{1 \rightarrow 2})_{\#} \mathcal{G}^{\circ}(\mu) = \pi_{\#}^2 \mathcal{G}_t^{\circ}(\mu)$$

we have

$$\tilde{P}_{\mu} = \frac{1}{2} P_{\mu,t} + \frac{1}{2} \delta_{\mathcal{G}_t^{\circ}(\mu)}, \quad \tilde{II}_{\mu} = \frac{1}{2} II_{\mu,t} + \frac{1}{2} \delta_{\mathcal{F}_t^{\circ}(\mu)} \quad M_1\text{-a.e.}$$

Since $\tilde{II}_{\mu} = \delta_{\tilde{\mathcal{F}}(\mu)}$, we deduce that $II_{\mu,t} = \mathcal{F}_t^{\circ}(\mu)$ for M_1 -a.e. μ , so that \tilde{M} is the middle point also between M_1 and M_2 and between M_1 and $M_2^{\circ} = (\mathcal{F}^{\circ})_{\#} M_1$ (with respect to the same optimal set \mathbf{F}). By the non-branching property of Theorem 5.7 we deduce that $M_2 = M_2^{\circ}$ and by the strict minimality of the section $\mathbf{f}^{\circ} P = \mathcal{G}_{\#}^{\circ} M_1$ is a strict Monge solution and is the unique element of $\text{R}\Gamma_o(M_1, M_2)$; similarly $II = (\text{Id} \times \mathcal{F}^{\circ})_{\#} M_1$ is deterministic and is the unique element of $\Gamma_o(M_1, M_2)$.

Let us eventually check the last statement. We take a Lipschitz totally convex function ϕ : if there exists a Borel set $B \subset \mathcal{P}_2(\mathbb{H})$ where $\partial_t \phi$ is not a singleton with $M_1(B) > 0$, by a standard measurable selection we can construct an optimal random coupling law P with first random marginal M_1 and second random marginal $M_2 = \pi_{\#}^2 P \in \mathcal{P}_2(\mathbb{H})$ (thanks to the fact that ϕ is Lipschitz) which is not concentrated on the map $\nabla_W \phi$, contradicting the above result. \square

By the above result if we want to solve the Monge problem for arbitrary target M_2 it seems natural to start from measures M_1 concentrated on $\mathcal{P}_2^r(\mathbb{H})$. The next simpler Proposition shows that in this case there is also a one-to-one correspondence between optimal couplings of (laws of) random measures and (laws of) random optimal couplings and for every fixed target M_2 the Monge and the strict Monge formulations of the L^2 -Optimal Transport problem are equivalent as well.

Proposition 6.5. *Let $M_1, M_2 \in \mathcal{P}_2(\mathbb{H})$ and let us assume that M_1 is concentrated on $\mathcal{P}_2^r(\mathbb{H})$, i.e. $\mu \in \mathcal{P}_2^r(\mathbb{H})$ for M_1 -a.e. μ . Then the restriction of the map $(\pi_{\#}^1, \pi_{\#}^2)_{\#}$ on $\text{R}\Gamma_o(M_1, M_2)$ is injective and every $P \in \text{R}\Gamma_o(M_1, M_2)$ is concentrated on $\mathcal{P}_2^{\text{det}}(\mathbb{H} \times \mathbb{H})$. Moreover, if $P \in \text{R}\Gamma_o(M_1, M_2)$ and $II = (\pi_{\#}^1, \pi_{\#}^2)_{\#} P$ is deterministic then P is fully deterministic.*

Proof. Since $P \in \text{R}\Gamma_o(M_1, M_2)$ and M_1 is concentrated on $\mathcal{P}_2^r(\mathbb{H})$, thanks to Theorem 2.5 we can find a Borel set $B \subset \mathcal{P}_2(\mathbb{H} \times \mathbb{H})$ of full P -measure such that $B \subset \mathcal{P}_{2,o}^{\text{det}}(\mathbb{H} \times \mathbb{H})$ and $\pi_{\#}^1(B) \subset \mathcal{P}_2^r(\mathbb{H})$ (when \mathbb{H} has finite dimension we just take $B = (\pi_{\#}^1)^{-1}(\mathcal{P}_2^r(\mathbb{H})) \cap \mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H})$, see Proposition 2.6).

The restriction of $(\pi_{\#}^1, \pi_{\#}^2)$ to B is injective, since, given $(\mu_1, \mu_2) \in (\pi_{\#}^1, \pi_{\#}^2)(B)$, the set $\Gamma_o(\mu_1, \mu_2)$ contains a unique element and it is deterministic. It follows that $(\pi_{\#}^1, \pi_{\#}^2)_{\#}$ is injective as well on $\text{R}\Gamma_o(M_1, M_2)$. If moreover II is deterministic and induced by the Borel map \mathcal{F} , we see that for P -a.e. γ , $\gamma \in \Gamma_o(\pi_{\#}^1 \gamma, \mathcal{F}(\pi_{\#}^1 \gamma))$ so that $\pi_{\#}^1$ is P -essentially injective. \square

6.2 Regular and super-regular measures in $\mathcal{P}_2(\mathbb{H})$ and solution to the Monge problem

We first observe that the definitions of σ -d.c. hypersurfaces and Gaussian null sets given in Section 2.2 also apply to the infinite dimensional Hilbert space \mathcal{H} ; we keep the notation $\mathcal{P}_2^r(\mathcal{H})$, $\mathcal{P}_2^{gr}(\mathcal{H})$ to denote the corresponding class of regular measures, thus vanishing on all d.c. hypersurfaces and Gaussian-null Borel subsets of \mathcal{H} respectively.

Definition 6.6 (Random exceptional and Gaussian null sets, regular and super-regular measures).

- We say that a Borel set $B \subset \mathcal{P}_2(\mathbb{H})$ is a random exceptional (resp. G -null) set if $\iota^{-1}(B)$ is contained in a σ -d.c. hypersurface of \mathcal{H} (resp. Gaussian-null in \mathcal{H}).

- We denote by $\mathcal{P}_2^r(\mathcal{P}_2(\mathbb{H}))$ (resp. $\mathcal{P}_2^{gr}(\mathcal{P}_2(\mathbb{H}))$) the set of regular (resp. G-regular) measures $\mathbb{M} \in \mathcal{P}_2(\mathbb{H})$ such that $\mathbb{M}(B) = 0$ for every random exceptional (resp. G-null) Borel set $B \subset \mathcal{P}_2(\mathbb{H})$.
- The set of super-regular measures $\mathfrak{P}_2^{rr}(\mathbb{H}) = \mathcal{P}_2^r(\mathcal{P}_2^r(\mathbb{H}))$ (resp. super-G-regular measures $\mathfrak{P}_2^{grr}(\mathbb{H}) = \mathcal{P}_2^{gr}(\mathcal{P}_2^r(\mathbb{H}))$) is the set of regular (resp. of G-regular) measures concentrated on $\mathcal{P}_2^r(\mathbb{H})$.

It is clear that

$$\mathcal{P}_2^{gr}(\mathcal{P}_2(\mathbb{H})) \subset \mathcal{P}_2^r(\mathcal{P}_2(\mathbb{H})), \quad \mathfrak{P}_2^{grr}(\mathbb{H}) \subset \mathfrak{P}_2^{rr}(\mathbb{H}). \quad (6.17)$$

Let us make a few comments on the previous definitions.

Remark 6.7 (LGGRM measures). *We can say that a measure $\mathbb{G} \in \mathcal{P}_2(\mathbb{H})$ is a Law of Gaussian Generated Random Measures (LGGRM) if $\mathbb{G} = \iota_{\sharp} \mathfrak{g}$ for some nondegenerate Gaussian measure \mathfrak{g} in \mathcal{H} . An equivalent way to say that a Borel set in $\mathcal{P}_2(\mathbb{H})$ is a random G-null set is*

$$\mathbb{G}(B) = 0 \quad \text{for every LGGRM } \mathbb{G}. \quad (6.18)$$

Remark 6.8 (Random exceptional and G-null sets are independent of the choice of $(\mathbb{Q}, \mathcal{F}_{\mathbb{Q}}, \mathbb{M})$). It is not difficult to see that the above definitions of random exceptional and G-null sets (and the corresponding classes of regular and super-regular measures) are independent of the choice of the nonatomic standard Borel space $(\mathbb{Q}, \mathcal{F}_{\mathbb{Q}}, \mathbb{M})$. In fact, if $(\mathbb{Q}', \mathcal{F}_{\mathbb{Q}'}, \mathbb{M}')$ is another standard Borel measure space endowed with a nonatomic measure \mathbb{M}' , we can find a measure preserving isomorphism $\mathfrak{h} : \mathbb{Q}' \rightarrow \mathbb{Q}$ such that $\mathfrak{h}_{\sharp} \mathbb{M}' = \mathbb{M}$ and $\mathfrak{h}_{\sharp}^{-1} \mathbb{M} = \mathbb{M}'$. \mathfrak{h} induces a linear isometry of \mathcal{H} onto $\mathcal{H}' = L^2(\mathbb{Q}', \mathbb{M}'; \mathbb{H})$ defined by $\mathfrak{h}^* X := X \circ \mathfrak{h}$, with $\iota' \circ \mathfrak{h}^* = \iota$, since

$$X' = X \circ \mathfrak{h}, \quad \iota'(X') = (X \circ \mathfrak{h})_{\sharp} \mathbb{M}' = X_{\sharp} \mathfrak{h}_{\sharp} \mathbb{M}' = X_{\sharp} \mathbb{M} = \iota(X). \quad (6.19)$$

Since the d.c. hypersurfaces are preserved by isometric isomorphisms between Hilbert spaces, using (6.19) it is immediate to check that a random exceptional set w.r.t. \mathbb{Q}' is also exceptional w.r.t. \mathbb{Q} .

In a similar way, if B' is a random Gaussian-null set with respect to \mathbb{Q}', \mathbb{M}' , i.e. $(\iota')^{-1}(B')$ is Gaussian-null in \mathcal{H}' and let \mathfrak{g} be an arbitrary nondegenerate Gaussian measure in \mathcal{H} . We introduce $\mathfrak{g}' = (\mathfrak{h}^*)_{\sharp} \mathfrak{g}$ and we observe that \mathfrak{g}' is a nondegenerate Gaussian measure in \mathcal{H}' (since \mathfrak{h}^* is a linear surjective isometry). By definition $\mathfrak{g}'((\iota')^{-1}B') = 0$ and therefore

$$\mathfrak{g}(\iota^{-1}B) = \mathfrak{g}((\iota' \circ \mathfrak{h}^*)^{-1}B) = \mathfrak{g}((\mathfrak{h}^*)^{-1}(\iota')^{-1}B) = \mathfrak{g}'((\iota')^{-1}B) = 0,$$

so that $\iota^{-1}B$ is Gaussian null in \mathcal{H} .

Remark 6.9 (Stability in the class of mutually absolutely continuous measures). *The super-regularity condition is stable with respect to multiplication by an integrable factor:*

$$\mathbb{M} \in \mathfrak{P}_2^{rr}(\mathbb{H}), \quad \mathbb{M}' \ll \mathbb{M} \quad \Rightarrow \quad \mathbb{M}' \in \mathfrak{P}_2^{rr}(\mathbb{H}). \quad (6.20)$$

Remark 6.10. *If \mathbb{M} is concentrated on $\mathcal{P}_2^r(\mathbb{H})$ then it is super-regular if it vanishes on all exceptional subsets of $\mathcal{P}_2^r(\mathbb{H})$, i.e. it is sufficient to check that*

$$B \subset \mathcal{P}_2^r(\mathbb{H}), \quad \iota^{-1}(B) \text{ exceptional in } \mathcal{H} \quad \Rightarrow \quad \mathbb{M}(B) = 0. \quad (6.21)$$

Similarly, if

$$B \subset \mathcal{P}_2^{gr}(\mathbb{H}), \quad \iota^{-1}(B) \text{ Gaussian null in } \mathcal{H} \quad \Rightarrow \quad \mathbb{M}(B) = 0 \quad (6.22)$$

then \mathbb{M} is super-G-regular.

There is a simple way to generate super-regular measures.

Lemma 6.11. *Let \mathfrak{m} be a regular measure in $\mathcal{P}_2^r(\mathcal{H})$ (respectively G -regular in $\mathcal{P}_2^{gr}(\mathcal{H})$) such that*

$$\iota(X) = X_{\#}\mathbb{M} \in \mathcal{P}_2^r(\mathbb{H}) \quad \text{for } \mathfrak{m}\text{-a.e. } X \in \mathcal{H}. \quad (6.23)$$

Then $\mathbb{M} := \iota_{\#}\mathfrak{m}$ is a super-regular measure in $\mathcal{P}_2^{rr}(\mathbb{H})$ (resp. super- G -regular in $\mathcal{P}_2^{grr}(\mathbb{H})$).

Proof. It is immediate to see that $\mathbb{M} := \iota_{\#}\mathfrak{m} \in \mathcal{P}_2^r(\mathcal{P}_2(\mathbb{H}))$ is regular: in fact, if B is a random exceptional set

$$\mathbb{M}(B) = \mathfrak{m}(\iota^{-1}B) = 0$$

since $\iota^{-1}(B)$ is a σ -d.c. hypersurface and \mathfrak{m} is regular. Condition (6.23) also shows that \mathbb{M} is concentrated on $\mathcal{P}_2^r(\mathbb{H})$. \square

It is also possible to change the reference measure \mathbb{M} : we show two simple cases.

Lemma 6.12. *Let $\mathbb{M}', \mathbb{M}''$ be atomless Borel probability measures on the standard Borel space $(\mathbb{Q}, \mathcal{F}_{\mathbb{Q}})$ with $\mathbb{M}'' \leq a\mathbb{M}'$ for some $a > 0$ (so that the corresponding Hilbert spaces $\mathcal{H}', \mathcal{H}''$ satisfy $\mathcal{H}' \subset \mathcal{H}''$ with continuous inclusion). Denote by $\iota' : X \rightarrow X_{\#}\mathbb{M}'$, $\iota'' : X \rightarrow X_{\#}\mathbb{M}''$ the corresponding law maps.*

If $\mathfrak{m} \in \mathcal{P}_2^r(\mathcal{H}')$ and $\mathbb{M}' = \iota'_{\#}\mathfrak{m}$ is super-regular then also $\mathbb{M}'' = \iota''_{\#}\mathfrak{m}$ is super-regular.

A similar result holds if $\mathbb{M}'' \ll \mathbb{M}'$, and $\mathfrak{m} \in \mathcal{P}_2^r(\mathcal{B})$, for some separable Banach space $\mathcal{B} \subset \mathcal{H}' \cap \mathcal{H}''$.

Proof. Let $N \subset \mathcal{H}'$ be a \mathfrak{m} -negligible Borel set such that $\iota'(X) \in \mathcal{P}_2^r(\mathbb{H})$ for every $X \in \mathcal{H}' \setminus N$. Since $\iota''(X) \leq a\iota'(X)$ for every $X \in \mathcal{H}'$ we deduce that $\iota''(X) \in \mathcal{P}_2^r(\mathbb{H})$ for every $X \in \mathcal{H}' \setminus N$ as well, so that \mathbb{M}'' is concentrated in $\mathcal{P}_2^r(\mathbb{H})$.

If B is a Borel exceptional set of $\mathcal{P}_2(\mathbb{H})$ then $(\iota'')^{-1}(B)$ is contained in a σ -d.c. hypersurface S of \mathcal{H}'' ; since $S \cap \mathcal{H}'$ is a σ -d.c. hypersurface as well, we deduce that $\mathbb{M}''(B) \leq \mathfrak{m}(S) = \mathfrak{m}(S \cap \mathcal{H}') = 0$ since \mathfrak{m} is regular. We conclude that \mathbb{M}'' is super-regular.

A similar argument applies to the second statement. \square

The relation between super-regular measures and differentiability of Lipschitz totally displacement convex functions is clarified by the next two results.

Theorem 6.13. *If $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R}$ is a totally displacement convex and Lipschitz function then the singular set $\text{Sing}^r(\phi) := \{\mu \in \mathcal{P}_2^r(\mathbb{H}) : \#\partial_t\phi[\mu] > 1\}$ of regular measures where the total subdifferential of ϕ is not reduced to a singleton (the minimal section) is exceptional.*

Proof. Since every measure $\mu \in \text{Sing}^r(\phi)$ is regular and $\partial_t\phi[\mu] \subset \mathcal{P}_{2,o}(\mathbb{H} \times \mathbb{H})$ by Corollary 4.10, all the elements of $\partial_t\phi[\mu]$ are deterministic couplings. If $\partial\phi[\mu]$ contains at least two different elements, they are associated with two different fields $\mathbf{f}_1(\cdot, \mu), \mathbf{f}_2(\cdot, \mu)$ in $L^2(\mathbb{H}, \mu; \mathbb{H})$.

If $X \in \iota^{-1}(\mu)$, setting $Y_i = \mathbf{f}_i(X, \mu)$ we deduce that $(X, Y_i)_{\#}\mathbb{M} \in \partial_t\phi[\mu]$ and therefore $Y_i \in \partial\hat{\phi}(X)$, where $\hat{\phi} = \phi \circ \iota$. We conclude that $\partial\hat{\phi}(X)$ contains two different elements and therefore $\hat{\phi}$ is not Gateaux-differentiable at X . This arguments shows that $\iota^{-1}(\text{Sing}^r(\phi))$ is a subset where $\hat{\phi}$ is not Gateaux-differentiable and therefore is a σ -d.c. hypersurface in \mathcal{H} since $\hat{\phi}$ is a convex Lipschitz function. \square

Combining Proposition 4.7, Remark 6.10 and the above Theorem we immediately get:

Corollary 6.14. *If $\mathbb{M} \in \mathcal{P}_2^{rr}(\mathbb{H})$ and $\phi : \mathcal{P}_2(\mathbb{H}) \rightarrow \mathbb{R}$ is a Lipschitz totally displacement convex function, then for \mathbb{M} -a.e. μ we have:*

1. $\partial_t\phi[\mu] = \partial_t^{\circ}\phi[\mu]$ is reduced to a single deterministic coupling in $\mathcal{P}_{2,o}^{\text{det}}(\mathbb{H}^2)$ of the form $(\mathbf{i} \times \nabla_W\phi(\cdot, \mu))_{\#}\mu$;
2. $\partial^-\phi(\mu) = \nabla_W\phi(\cdot, \mu)_{\#}\mu$.

Theorem 6.15 (Solutions to the strict Monge problem for super-regular measures). *If $\mathbf{M} \in \mathfrak{P}_2^{str}(\mathbb{H})$ and $\mathbf{N} \in \mathfrak{P}_2(\mathbb{H})$, then $\text{R}\Gamma_o(\mathbf{M}, \mathbf{N})$ and $\Gamma_o(\mathbf{M}, \mathbf{N})$ contain a unique element \mathbf{P} and $\Pi = (\pi_{\sharp}^1, \pi_{\sharp}^2)_{\sharp} \mathbf{P}$ respectively.*

\mathbf{P} is fully deterministic and there exists a unique Borel map $\mathbf{f} \in L^2(\bar{\mathbb{M}}; \mathbb{H})$ solving the Monge OT problem 6.3. \mathbf{f} is essentially totally cyclically monotone and coincides with the minimal section $\nabla_W \phi$ of an optimal Kantorovich potential.

Proof. By Proposition 6.5 and Theorem 6.4 it is sufficient to prove that $\Gamma_o(\mathbf{M}, \mathbf{N})$ contains a deterministic coupling Π .

Let us first suppose that \mathbf{N} is concentrated on the closed ball

$$\mathcal{B}(R) := \left\{ \nu \in \mathcal{P}_2(\mathbb{H}) : \mathfrak{m}_2(\nu) \leq R \right\}. \quad (6.24)$$

In this case, by Corollary 5.6 there is an optimal Kantorovich potential ϕ which is a R -Lipschitz totally convex function ϕ with $S := \text{supp}(\Pi) \subset \partial^- \phi$ and Corollary 6.14 shows that there exists a (unique) strictly Monge solution which is given by

$$\mathbf{f}(\cdot, \mu) = \nabla_W \phi(x, \mu). \quad (6.25)$$

In the general case, we argue as in the proof of Theorem 6.2.10 [AGS08]. For every $R > 0$ we set

$$\mathcal{B}'(R) := \mathcal{P}_2(\mathbb{H}) \times \mathcal{B}(R)$$

and for an optimal coupling $\Pi \in \Gamma_o(\mathbf{M}, \mathbf{N})$, we set for sufficiently large n

$$Z_n := \Pi(\mathcal{B}'(n)), \quad \Pi_n := Z_n^{-1} \Pi \llcorner \mathcal{B}'(n), \quad \mathbf{M}_n := \pi_{\sharp}^1 \Pi_n, \quad \mathbf{N}_n := \pi_{\sharp}^2 \Pi_n. \quad (6.26)$$

Since optimality is preserved by restriction, $\Pi_n \in \Gamma_o(\mathbf{M}_n, \mathbf{N}_n)$; since $\mathbf{M}_n \ll \mathbf{M}$ and \mathbf{N}_n is concentrated on $\mathcal{B}(n)$ the previous argument shows that there exists a unique map \mathcal{F}_n such that $\Pi_n = (\text{Id} \times \mathcal{F}_n)_{\sharp} \mathbf{M}_n$.

Moreover, for $m > n$ we easily get $\mathcal{F}_m = \mathcal{F}_n$ \mathbf{M}_n -a.e., so that there exists a map \mathcal{F} such that $\mathcal{F} = \mathcal{F}_n$ \mathbf{M}_n -a.e. for every $n \in \mathbb{N}$. Passing to the limit in the identity $\Pi_n = (\text{Id} \times \mathcal{F})_{\sharp} \mathbf{M}_n$ we obtain $\Pi = (\text{Id} \times \mathcal{F})_{\sharp} \mathbf{M}$. It follows that Π is deterministic. \square

Anticipating some of the results of the next section, it is easy to see that $\mathfrak{P}_2^{str}(\mathbb{H})$ is dense in $\mathfrak{P}_2(\mathbb{H})$.

Proposition 6.16. *If \mathbb{H} has finite dimension then $\mathfrak{P}_2^{gtr}(\mathbb{H})$ is dense in $\mathfrak{P}_2(\mathbb{H})$. In particular, the class of initial measures for which the OT problem has a unique solution in (strict) Monge form is dense.*

Proof. Thanks to (6.20), if $\mathbf{M} \in \mathfrak{P}_2^{gtr}(\mathbb{H})$ and $\mathbf{M}' \ll \mathbf{M}$ then also $\mathbf{M}' \in \mathfrak{P}_2^{gtr}(\mathbb{H})$. If \mathbb{H} has finite dimension, we will see in the next section (see Theorems 6.19 and 6.25) that there exists a reference measure $\mathbf{G} \in \mathfrak{P}_2^{gtr}(\mathbb{H})$ with full support. It is then sufficient to observe that the set

$$\left\{ \mathbf{M}' \in \mathfrak{P}_2(\mathbb{H}) : \mathbf{M}' \ll \mathbf{G} \right\} \quad \text{is dense in } \mathfrak{P}_2(\mathbb{H}),$$

since its closure contains all the finite combination of Dirac masses in $\mathfrak{P}_2(\mathbb{H})$. In fact, if $\mathbf{M} = \sum_{k=1}^n a_k \delta_{\mu_k}$ for distinct points $\mu_1, \dots, \mu_n \in \mathfrak{P}_2(\mathbb{H})$, we can choose $r > 0$ so small that the balls $B_{r,k} = B_r(\mu_k)$ are disjoint. Since \mathbf{G} has full support, $Z_{r,k} := G(B_{r,k}) > 0$ so that

$$\mathbf{M}_r := \left(\sum_{k=1}^n \frac{a_k}{Z_{r,k}} \chi_{B_{r,k}} \right) G \in \mathfrak{P}_2^{gtr}(\mathbb{H})$$

and $\mathbf{M}_r \rightarrow \mathbf{M}$ in $\mathfrak{P}_2(\mathbb{H})$ as $r \downarrow 0$. \square

6.3 Examples of super-regular measures for finite dimensional \mathbb{H}

In this last section we will exhibit many examples of super-regular measures. We will focus on the relevant case of measures induced by nondegenerate Gaussian measures \mathfrak{g} on \mathcal{H} as in Example 2.10, when \mathbb{H} has finite dimension: by using Lemma 6.11 it will be sufficient to check that (6.23) holds, i.e. $\iota_{\sharp}\mathfrak{g}(\mathcal{P}_2(\mathbb{H}) \setminus \mathcal{P}_2^r(\mathbb{H})) = 0$.

We start from the 1-dimensional case, where we will prove a very general result.

The 1-dimensional case $\mathbb{H} = \mathbb{R}$

Example 6.17 (The random occupation measure associated with Brownian motion in $[0, 1]$). Let $Q = [0, 1]$ endowed with the usual Lebesgue measure \mathbb{M} and let \mathfrak{w} be the standard Wiener measure concentrated on $C([0, 1])$. We claim that

$W := \iota_{\sharp}\mathfrak{w}$ is super- G -regular with full support.

We can apply Lemma 6.11. Since \mathfrak{w} is a Gaussian non-degenerate measure in $\mathcal{H} = L^2([0, 1], \mathbb{M})$, it is clearly a regular measure in $\mathcal{P}_2^r(\mathcal{H})$, so it is sufficient to check that

$$X_{\sharp}\mathbb{M} \text{ is nonatomic for } \mathfrak{w}\text{-a.e. path } X \in \mathcal{H}, \quad (6.27)$$

i.e.

$$\mathbb{M}\left(\left\{t \in [0, 1] : X_t = y\right\}\right) = 0 \quad \text{for } \mathfrak{w}\text{-a.e. } X \in C([0, 1]). \quad (6.28)$$

In fact we have the much stronger result that the so-called occupation measure $\iota(X) = X_{\sharp}\mathbb{M}$ is absolutely continuous with respect to the Lebesgue measure in \mathbb{R} , see e.g. [MP10, Theorem 3.26].

The previous example is in fact a particular case of a general result: we will show that *any* measure $G = \iota_{\sharp}\mathfrak{g}$ obtained as the push forward of an arbitrary nondegenerate Gaussian measure \mathfrak{g} in \mathcal{H} is super- G -regular.

Before discussing this result, let us show a simple criterion ensuring that a measure $\mathbb{M} \in \mathcal{P}_2(\mathbb{R})$ is concentrated on $\mathcal{P}_2^r(\mathbb{R})$. We set

$$\chi_0(r) := \begin{cases} 1 & \text{if } r = 0, \\ 0 & \text{if } r \neq 0. \end{cases} \quad (6.29)$$

Lemma 6.18. *Let $\mathbb{M} = \iota_{\sharp}\mathfrak{m}$ for $\mathfrak{m} \in \mathcal{P}_2(\mathcal{H})$ and let us suppose that \mathfrak{m} admits the representation (2.29), (2.30) discussed in Section 2.4. If*

$$\int_{\Omega} \chi_0(\Xi(\omega, q_1) - \Xi(\omega, q_2)) \, d\mathbb{P}(\omega) = 0 \quad \text{for } \mathbb{M} \otimes \mathbb{M}\text{-a.e. } (q_1, q_2) \quad (6.30)$$

then \mathbb{M} is concentrated on $\mathcal{P}_2^r(\mathbb{R})$.

Proof. Let $D := \{(x, x) : x \in \mathbb{R}\}$ be the diagonal in \mathbb{R}^2 , whose characteristic function is given by $(x, y) \mapsto \chi_0(x - y)$. We recall that a measure $\mu \in \mathcal{P}_2(\mathbb{R})$ is atomless (and thus belongs to $\mathcal{P}_2^r(\mathbb{R})$) if and only if

$$\mu \otimes \mu(D) = \int_{\mathbb{R}^2} \chi_0(x - y) \, d\mu(x) \, d\mu(y) = 0. \quad (6.31)$$

Recalling the definition (2.34) of the k -projection of \mathbb{M} , we deduce that \mathbb{M} is concentrated on $\mathcal{P}_2^r(\mathbb{R})$ if and only if

$$\text{pr}^2[\mathbb{M}](D) = \int \chi_0(x - y) \, d\text{pr}^2[\mathbb{M}](x, y) = 0.$$

Applying formula (2.36) we thus express the above integral as

$$\int_{\Omega} \left(\int \chi_0(\Xi(\omega, q_1) - \Xi(\omega, q_2)) \, d\mathbb{M}^{\otimes 2}(q_1, q_2) \right) \, d\mathbb{P}(\omega) = 0. \quad (6.32)$$

An application of Fubini's Theorem yields (6.30). \square

As an application of the above Lemma we have the following general result.

Theorem 6.19 (Push forward of nondegenerate Gaussian measures are superregular). *If $\mathbb{H} = \mathbb{R}$ and \mathfrak{g} is a nondegenerate Gaussian measure in \mathcal{H} , then $\mathbb{G} := \nu_{\sharp}\mathfrak{g} \in \mathcal{P}_2^{grr}(\mathbb{H})$ is super- G -regular.*

Proof. Since \mathbb{Q} is a standard Borel space, we can find a bounded metric $d_{\mathbb{Q}}$ in \mathbb{Q} such that $(\mathbb{Q}, d_{\mathbb{Q}})$ is a complete and separable metric space and $\mathcal{F}_{\mathbb{Q}}$ coincides with the Borel σ -algebra induced by $d_{\mathbb{Q}}$. We can also define $d_{\mathbb{Q}^2}((q_1, q_2), (q'_1, q'_2)) := \max[d_{\mathbb{Q}}(q_1, q'_1), d_{\mathbb{Q}}(q_2, q'_2)]$ and the swap isometric map $S : \mathbb{Q}^2 \rightarrow \mathbb{Q}^2$, $S(q_1, q_2) := (q_2, q_1)$.

We adopt the notation of Example (2.10). Since \mathfrak{g} is nondegenerate, we can also assume that $\lambda_n > 0$ for every $n \in \mathbb{N}_+$. By Lemma 6.11, it is sufficient to prove that \mathfrak{g} satisfies condition (6.23). Using the representation (2.38) and (2.39), we can then apply Lemma (6.18): our thesis follows if we prove (6.30).

Thanks to (2.39), we have

$$D(\omega; q_1, q_2) := \Xi(\omega, q_1) - \Xi(\omega, q_2) = \sum_n \xi_n(\omega) \left(\mathbf{E}_n(q_1) - \mathbf{E}_n(q_2) \right) \quad (6.33)$$

which is a series of independent random Gaussian variables. We know that for $\mathbb{M}^{\otimes 2}$ -a.e. (q_1, q_2)

- the series defining D converges in $L^2(\Omega, \mathbb{P})$ and also \mathbb{P} -a.e.,
- its law $\nu_{q_1, q_2} := D(\cdot; q_1, q_2)_{\sharp}\mathbb{P}$ is a Gaussian measure
- $\nu_{q_1, q_2} = N(0, \lambda^2(q_1, q_2))$ where

$$\lambda^2(q_1, q_2) = \sum_n \lambda_n^2 \left(\mathbf{E}_n(q_1) - \mathbf{E}_n(q_2) \right)^2. \quad (6.34)$$

We can observe that the integral in (6.30) is just $\nu_{q_1, q_2}(\{0\})$, so that (6.30) holds if ν_{q_1, q_2} is nondegenerate, i.e.

$$\lambda^2(q_1, q_2) > 0 \quad \text{for } \mathbb{M}^{\otimes 2}\text{-a.e. } (q_1, q_2). \quad (6.35)$$

Let us denote by $A \subset \mathbb{Q}^2$ the set where λ vanishes and let $D_{\mathbb{Q}} := \{(q, q) : q \in \mathbb{Q}\}$ be the diagonal in \mathbb{Q}^2 . Since \mathbb{M} is diffuse, $\mathbb{M}^{\otimes 2}(D_{\mathbb{Q}}) = 0$ so that we have to prove that $\mathbb{M}^{\otimes 2}(A') = 0$ where $A' := A \setminus D_{\mathbb{Q}}$. Since $\lambda_n > 0$ for every n , we immediately see that

$$(q_1, q_2) \in A \quad \Leftrightarrow \quad \mathbf{E}_n(q_1) = \mathbf{E}_n(q_2) \quad \text{for every } n \in \mathbb{N}_+. \quad (6.36)$$

We argue by contradiction and we suppose that $\mathbb{M}^{\otimes 2}(A') > 0$. We can thus find $(\bar{q}_1, \bar{q}_2) \in A' \cap \text{supp}(\mathbb{M})$ and a sufficiently small ball $B = B_r(\bar{q}_1, \bar{q}_2)$ such that $S(B) \cap B = \emptyset$ and $\mathbb{M}^{\otimes 2}(A' \cap B) > 0$. Since A' is symmetric and $S_{\sharp}\mathbb{M}^{\otimes 2} = \mathbb{M}^{\otimes 2}$, we have $\mathbb{M}^{\otimes 2}(A' \cap B) = \mathbb{M}^{\otimes 2}(A' \cap S(B)) > 0$.

We set $B' := A' \cap B$ and we eventually consider the bounded Borel function

$$f(q_1, q_2) := \chi_{B'}(q_1, q_2) - \chi_{S(B')}(q_1, q_2) = \chi_{B'}(q_1, q_2) - \chi_{B'}(q_2, q_1).$$

We can expand f as a orthogonal series in $L^2(\mathbb{Q}^2, \mathbb{M}^{\otimes 2})$ with respect to the complete orthonormal system $\mathbf{E}_{m,n}(q_1, q_2) := \mathbf{E}_m(q_1)\mathbf{E}_n(q_2)$ obtaining

$$f = \sum_{m,n} \hat{f}_{m,n} \mathbf{E}_{m,n} \quad \text{converging in } L^2(\mathbb{Q}^2, \mathbb{M}^{\otimes 2}) \quad (6.37)$$

where

$$\hat{f}_{m,n} := \int_{\mathbb{Q}^2} f(q_1, q_2) \mathbf{E}_m(q_1) \mathbf{E}_n(q_2) d\mathbb{M}^{\otimes 2}(q_1, q_2). \quad (6.38)$$

Since $f \equiv 0$ if $(q_1, q_2) \notin A$, the integral in (6.38) can in fact be restricted to A . Since (6.36) implies in particular

$$\mathbf{E}_{m,n}(q_1, q_2) = \mathbf{E}_m(q_1)\mathbf{E}_n(q_2) = \mathbf{E}_n(q_1)\mathbf{E}_m(q_2) = \mathbf{E}_{n,m}(q_1, q_2) \quad \mathbb{M}^{\otimes 2}\text{-a.e. in } A$$

we immediately get $\hat{f}_{m,n} = \hat{f}_{n,m}$. On the other hand, inverting the order of q_1, q_2 in (6.38), using the invariance of $\mathbb{M}^{\otimes 2}$ and the anti-symmetry of f , i.e. $f(q_2, q_1) = -f(q_1, q_2)$ we also get

$$\begin{aligned} \hat{f}_{m,n} &= \int_{\mathbb{Q}^2} f(q_2, q_1)\mathbf{E}_m(q_2)\mathbf{E}_n(q_1) \, d\mathbb{M}^{\otimes 2}(q_1, q_2) \\ &= - \int_{\mathbb{Q}^2} f(q_1, q_2)\mathbf{E}_m(q_2)\mathbf{E}_n(q_1) \, d\mathbb{M}^{\otimes 2}(q_1, q_2) = -\hat{f}_{n,m} \end{aligned}$$

We deduce that $\hat{f}_{m,n} = 0$ for every pair of indexes, a contradiction since f is not identically 0. \square

The range of application of the previous Theorem can be considerably extended thanks to the following simple results.

Corollary 6.20. *Let $\mathcal{N} := \mathcal{P}_2(\mathbb{R}) \setminus \mathcal{P}_2^r(\mathbb{R})$. Then the (Borel) set $\iota^{-1}(\mathcal{N})$ is Gaussian-null in \mathcal{H} .*

Proof. Theorem 6.19 shows that $\mathbf{g}(\iota^{-1}(\mathcal{N})) = 0$ for every non-degenerate Gaussian \mathbf{g} , so that $\iota^{-1}(\mathcal{N})$ is Gaussian-null by definition. \square

Corollary 6.21 (Push forward of G-regular measures are super-G-regular). *If $\mathbb{H} = \mathbb{R}$ and \mathbf{r} is a G-regular measure in $\mathcal{P}_2^{gr}(\mathcal{H})$, then $\mathbf{R} := \iota_{\sharp}\mathbf{r}$ is super-G-regular.*

Proof. Recall that as a G-regular measure in \mathcal{H} \mathbf{r} satisfies

$$\mathbf{r}(B) = 0 \quad \text{for every Gaussian null Borel set in } \mathcal{H}, \quad (6.39)$$

in particular $\mathbf{f}(\iota^{-1}(\mathcal{N})) = 0$, so that $\iota(X) \in \mathcal{P}_2^r(\mathbb{R})$ for \mathbf{r} -a.e. $X \in \mathcal{H}$. We conclude by Lemma 6.11. \square

Let us see two simple examples.

Example 6.22 (Sum of Gaussians with random signs). Let $\mathbb{Q} := \{-1, +1\}^{\mathbb{N}}$ be the Cantor set endowed with the uniform product measure $\mathbb{M} := (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes \mathbb{N}}$. Every element $q \in \mathbb{Q}$ is a vector $(q_i)_{i \in \mathbb{N}}$ of signs ± 1 indexed by $i \in \mathbb{N}$. We denote by $\varepsilon_i : \mathbb{Q} \rightarrow \mathbb{R}$ the i -th coordinate, and for every finite subset $I \subset \mathbb{N}$ we consider the Walsh function

$$W_I(q) := \prod_{i \in I} \varepsilon_i, \quad W_{\emptyset} \equiv 1. \quad (6.40)$$

If \mathcal{J} denotes the (countable) collection of all the finite parts of \mathbb{N} , the Walsh system $(W_I)_{I \in \mathcal{J}}$ is a complete orthonormal system in $\mathcal{H} = L^2(\mathbb{Q}, \mathbb{M})$.

We select a family of independent Gaussian random variables $\xi_I \sim N(0, \lambda_I^2)$ indexed by $I \in \mathcal{J}$ and coefficients $\lambda_I > 0$ for every $I \in \mathcal{J}$ such that $\Lambda := \sum_{I \in \mathcal{J}} \lambda_I^2 < \infty$; we form the random vector $\boldsymbol{\xi} = \sum_{I \in \mathcal{J}} \xi_I W_I \in \mathcal{H}$. Denoting by \mathbf{g}_W the law of $\boldsymbol{\xi}$ in \mathcal{H} we obtain a nondegenerate Gaussian measure $\mathbf{g}_W \in \mathcal{P}_2^r(\mathcal{H})$. Applying Theorem 6.19 we immediately get

$G_W := \iota_{\sharp}\mathbf{g}_W$ is super-regular.

Example 6.23 (The law of random Fourier series). Let us now select $\mathbb{Q} := (0, \pi)$ with the (normalized) Lebesgue measure \mathbb{M} . We consider the complete orthonormal system in $\mathcal{H} := L^2(0, \pi)$ given by the usual Fourier basis

$$S_n(q) := \sqrt{2} \sin(nq), \quad n \in \mathbb{N}_+. \quad (6.41)$$

As for the previous example, we select a sequence of independent Gaussian random variables $\xi_n \sim N(0, \lambda_n^2)$ and coefficients $\lambda_n > 0$ for every $n \in \mathbb{N}_+$ with $\Lambda = \sum_{n=1}^{\infty} \lambda_n^2 < \infty$. Denoting by \mathbf{g}_F the law of the random vector $\boldsymbol{\xi} = \sum_{n=1}^{\infty} \xi_n S_n$ we obtain a nondegenerate Gaussian measure $\mathbf{g}_F \in \mathcal{P}_2^r(\mathcal{H})$.

$G_F := \iota_{\sharp}\mathbf{g}_F$ is super-regular.

Notice that the case $\lambda_n := \frac{1}{\pi n}$ corresponds to the (centered) Brownian bridge.

The finite dimensional case $H = \mathbb{R}^d$, $d > 1$

Let us now discuss the case when $H = \mathbb{R}^d$, $d > 1$. We will still focus on the construction of suitable Gaussian measures \mathfrak{g} on \mathcal{H} such that $G = \iota_{\sharp}\mathfrak{g}$ is super-regular and we will consider two different approaches. Unlike the previous 1-d case, we will impose further properties on \mathfrak{g} .

Example 6.24 (Gaussian measures concentrated on C^1 maps). In this first example, we select $Q := (0, 1)^d$ (or any smooth domain in \mathbb{R}^d) endowed with the d -dimensional Lebesgue measure \mathbb{M} .

Theorem 6.25. *If \mathfrak{g} is a non-degenerate Gaussian measure on the Banach space $\mathcal{B} = C^1(\overline{Q}; \mathbb{R}^d) \subset \mathcal{H}$, then $G = \iota_{\sharp}\mathfrak{g}$ is super-G-regular with full support.*

Proof. As before, we apply Lemma 6.11. Since \mathfrak{g} is a Gaussian non-degenerate measure in \mathcal{B} and \mathcal{B} is dense in $\mathcal{H} = L^2(Q, \mathbb{M}; \mathbb{R}^d)$, \mathfrak{g} is clearly a Gaussian nondegenerate measure in $\mathcal{P}_2^{gr}(\mathcal{H})$, so it is sufficient to check that

$$X_{\sharp}\mathbb{M} \text{ is absolutely continuous for } \mathfrak{g}\text{-a.e. } X \in \mathcal{B}. \quad (6.42)$$

By the area and co-area formulae (see e.g. [GH80, Thm. 2.3]), the push forward $\mu_X = X_{\sharp}\mathbb{M}$ of a map $X \in C^1(\overline{Q}; \mathbb{R}^d)$ is absolutely continuous if

$$\mathbb{M}(\{q \in Q : \det DX(q) = 0\}) = 0, \quad (6.43)$$

so that (6.42) is true if we show that (6.43) holds for \mathfrak{g} -a.e. $X \in \mathcal{B}$. We can prove this property by Fubini's Theorem. We consider the product measure $\tilde{\mathfrak{g}} := \mathfrak{g} \otimes \mathbb{M}$ concentrated on $\mathcal{B} \times Q$, we denote by $D : \mathcal{B} \times Q \rightarrow \mathbb{R}^{d \times d}$ the ‘‘differential’’ evaluation map $D(X, q) := DX(q)$, and we introduce the closed set

$$A := \{(X, q) \in \mathcal{B} \times Q : DX(q) \in S\}, \quad S := \{D \in \mathbb{R}^{d \times d} : \det D = 0\}. \quad (6.44)$$

We know that for every $q \in Q$ the map $X \mapsto D(X, q)$ is linear, continuous, and surjective from \mathcal{B} to $\mathbb{R}^{d \times d}$. We thus deduce that $D(\cdot, q)_{\sharp}\mathfrak{g}$ is a nondegenerate Gaussian measure in $\mathbb{R}^{d \times d}$, so that

$$\mathfrak{g}(\{X \in \mathcal{B} : D(X, q) \in S\}) = 0 \quad \text{for every } q \in Q.$$

Integrating in Q we get

$$\tilde{\mathfrak{g}}(A) = \int \chi_A(X, q) d\tilde{\mathfrak{g}} = \int_Q \mathfrak{g}(X \in \mathcal{B} : (X, q) \in A) d\mathbb{M}(q) = 0. \quad (6.45)$$

Applying Fubini's Theorem we thus deduce that

$$\int_{\mathcal{B}} \mathbb{M}(q \in Q : (X, q) \in A) d\mathfrak{g}(X) = 0$$

which yields (6.43) for \mathfrak{g} -a.e. $X \in \mathcal{B}$. □

The next example is a natural generalization of Example 6.17.

Example 6.26 (The occupation measure of the fractional Brownian motion). Let $Q = [0, 1]$ endowed with the Lebesgue measure; we fix a Hurst parameter $H < 1/d$ and we consider the d -dimensional fractional Brownian motion $(\Xi_t^H)_{t \in Q}$ [Bia+08]. Since Ξ^H has local time (or, equivalently, its occupation measure is absolutely continuous with square integrable density [Pit78; GH80], [Bia+08, Thm. 10.2.3]), its law \mathfrak{w}^H in $\mathcal{B} := C([0, 1]; \mathbb{R}^d)$ is a nondegenerate Gaussian measures satisfying (6.23), so that $W^H = \iota_{\sharp}\mathfrak{w}^H \in \mathcal{P}_2(H)$ is super-G-regular, according to Lemma 6.11.

The above example is an application of a general technique due to Berman [Ber69] and involving the Fourier transform. In order to explain the main idea in a general case, let us denote by $\mu_X \in \mathcal{P}_2(\mathbb{R}^d)$ the law $\iota(X) = X_{\#}\mathbb{M}$ of a generic element $X \in \mathcal{H}$. The Fourier transform of μ_X is the continuous function $\hat{\mu}_X : \mathbb{R}^d \rightarrow \mathbb{C}$ defined by

$$\hat{\mu}_X(u) := \int_{\mathbb{R}^d} e^{i u \cdot x} d\mu_X(x) = \int_{\mathbb{Q}} e^{i u \cdot X(q)} d\mathbb{M}(q) \quad u \in \mathbb{R}^d. \quad (6.46)$$

Notice that $(X, u) \mapsto \hat{\mu}_X(u)$ is jointly continuous in $\mathcal{H} \times \mathbb{R}^d$. By Plancherel theorem, μ_X is absolutely continuous w.r.t. the d -dimensional Lebesgue measure in \mathbb{R}^d with a density $\varrho_X \in L^2(\mathbb{R}^d)$ if and only if $\hat{\mu}_X \in L^2(\mathbb{R}^d)$ and moreover

$$\int_{\mathbb{R}^d} \varrho_X^2 dx = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |\hat{\mu}_X(u)|^2 du \quad (6.47)$$

It follows that if the measure $\mathbf{g} \in \mathcal{P}_2(\mathcal{H})$ satisfies

$$L^2 := \int_{\mathcal{H}} \left(\int_{\mathbb{R}^d} |\hat{\mu}_X(u)|^2 du \right) d\mathbf{g}(X) = \int_{\mathbb{R}^d} \left(\int_{\mathcal{H}} |\hat{\mu}_X(u)|^2 d\mathbf{g}(X) \right) du < \infty \quad (6.48)$$

we get $\mu_X \in \mathcal{P}_2^r(\mathbb{R}^d)$ for \mathbf{g} -a.e. X . We can rewrite (6.48) by using the representation (2.29) by the process Ξ , so that $X(\cdot) = \Xi(\omega, \cdot)$ and therefore we can set $\mu_\omega = \mu_{\Xi(\omega, \cdot)}$. We have

$$\begin{aligned} |\hat{\mu}_\omega(u)|^2 &= \hat{\mu}_\omega(u) \cdot \overline{\hat{\mu}_\omega(u)} = \left(\int_{\mathbb{Q}} e^{i u \cdot \Xi(\omega, q_1)} d\mathbb{M}(q_1) \right) \left(\int_{\mathbb{Q}} e^{-i u \cdot \Xi(\omega, q_2)} d\mathbb{M}(q_2) \right) \\ &= \int_{\mathbb{Q}^2} e^{i u \cdot \Xi(\omega, q_1)} e^{-i u \cdot \Xi(\omega, q_2)} d\mathbb{M}^{\otimes 2}(q_1, q_2) \\ &= \int_{\mathbb{Q}^2} e^{i u \cdot (\Xi(\omega, q_1) - \Xi(\omega, q_2))} d\mathbb{M}^{\otimes 2}(q_1, q_2). \end{aligned} \quad (6.49)$$

Taking now the expectation w.r.t. \mathbb{P} and integrating in \mathbb{R}^d w.r.t. u we end up with Berman condition [GH80, Thm. 21.9]

$$L^2 = \int_{\mathbb{Q}^2} \left(\int_{\mathbb{R}^d} \mathbb{E}_{\mathbb{P}} \left[e^{i u \cdot (\Xi(\omega, q_1) - \Xi(\omega, q_2))} \right] du \right) d\mathbb{M}^{\otimes 2}(q_1, q_2) < \infty \quad (6.50)$$

In the particular case when $\Xi = (\Xi^1, \dots, \Xi^d)$ is a Gaussian process and the determinant $\Delta(q_1, q_2)$ of the covariance matrix of $\Xi(\cdot, q_1) - \Xi(\cdot, q_2)$ is positive for a.e. q_1, q_2 , we end up with the sufficient condition for the validity of (6.50) [GH80, Thm. 22.1]

$$\int_{\mathbb{Q}^2} \frac{1}{\Delta(q_1, q_2)^d} d\mathbb{M}^{\otimes 2}(q_1, q_2) < \infty. \quad (6.51)$$

In the case of the fractional Brownian motion of Example 6.26 we thus recover the condition $Hd < 1$.

Example 6.27 (Berman condition for Karhunen-Loève expansions). We slightly modify the above argument, by considering an example inspired to the general framework discussed in Example 2.10 and based on a complete orthonormal system \mathbf{E}'_n of the Hilbert space $\mathcal{H}' := L^2(\mathbb{Q}, \mathbb{M}; \mathbb{R})$ of scalar valued square summable functions. If e_1, \dots, e_d is an orthogonal basis of \mathbb{R}^d (e.g. the canonical one), we can then form the complete orthonormal system $\mathbf{E}_{n,k} := e_k \mathbf{E}'_n$ of $\mathcal{H} := L^2(\mathbb{Q}, \mathbb{M}; \mathbb{R}^d)$.

We assign a sequence Σ_n , $n \in \mathbb{N}_+$, of symmetric and positive definite matrices in $\mathbb{R}^{d \times d}$ satisfying the boundedness and coercivity condition

$$0 < \alpha_n^2 \leq \Sigma_n \mathbf{v} \cdot \mathbf{v} \leq \beta_n^2 \quad \text{for every } \mathbf{v} \in \mathbb{R}^d, |\mathbf{v}| = 1; \quad \sum_{n=1}^{\infty} \beta_n^2 < \infty. \quad (6.52)$$

We assign a sequence of centered independent Gaussian random variables in \mathbb{R}^d $\xi_n \sim N(0, \Sigma_n)$; since $\sum_n \beta_n^2 < \infty$, we can form the random vector

$$\boldsymbol{\xi} = \sum_n \xi_n \mathbf{E}'_n \quad (6.53)$$

corresponding to the the measurable process

$$\Xi(\omega, q) = \sum_n \xi_n(\omega) \mathbf{E}'_n(q).$$

It is clear that $\mathbf{g} = \boldsymbol{\xi}_\# \mathbb{P}$ is a nondegenerate Gaussian in \mathcal{H} . As in (6.34) we consider the functions

$$\alpha^2(q_1, q_2) = \sum_n \alpha_n^2 \left(\mathbf{E}'_n(q_1) - \mathbf{E}'_n(q_2) \right)^2, \quad \beta^2(q_1, q_2) = \sum_n \beta_n^2 \left(\mathbf{E}'_n(q_1) - \mathbf{E}'_n(q_2) \right)^2 \quad (6.54)$$

formed with the orthonormal system of the scalar-valued L^2 - space \mathcal{H}' . We have already seen as a particular consequence of the calculations of Theorem 6.19 that $\alpha^2(q_1, q_2) > 0$ a.e. if $q_1 \neq q_2$.

Theorem 6.28. *If*

$$\int_{\mathbb{Q}^2} \frac{1}{\alpha^d(q_1, q_2)} d\mathbb{M}(q_1) d\mathbb{M}(q_2) < \infty \quad (6.55)$$

then $\mathbf{G} = \iota_\# \mathbf{g}$ is super-G-regular.

Notice that the integral in (6.55) can be restricted to the complement of the diagonal $D_{\mathbb{Q}}$ in \mathbb{Q}^2 , since \mathbb{M} is atomless.

Proof. We want to prove that (6.50) holds. We first integrate (6.49) with respect to \mathbb{P} , obtaining

$$\mathbb{E}_{\mathbb{P}} |\hat{\mu}_\omega(u)|^2 = \int_{\Omega} \left(\int_{\mathbb{Q}^2} e^{i u \cdot (\Xi(\omega, q_1) - \Xi(\omega, q_2))} d\mathbb{M}^{\otimes 2}(q_1, q_2) \right) d\mathbb{P}(\omega) \quad (6.56)$$

$$= \int_{\mathbb{Q}^2} \left(\int_{\Omega} e^{i u \cdot (\Xi(\omega, q_1) - \Xi(\omega, q_2))} d\mathbb{P}(\omega) \right) d\mathbb{M}^{\otimes 2}(q_1, q_2) \quad (6.57)$$

$$= \int_{\mathbb{Q}^2} \mathbb{E}_{\mathbb{P}} \left[e^{i u \cdot (\Xi(\cdot, q_1) - \Xi(\cdot, q_2))} \right] d\mathbb{M}^{\otimes 2}(q_1, q_2). \quad (6.58)$$

We now observe that for $\mathbb{M}^{\otimes 2}$ -a.e. (q_1, q_2) $\beta^2(q_1, q_2)$ is finite so that the expression

$$D(\omega; q_1, q_2) := \Xi(\omega, q_1) - \Xi(\omega, q_2) = \sum_n \xi_n(\omega) [\mathbf{E}'_n(q_1) - \mathbf{E}'_n(q_2)] \quad (6.59)$$

is a series of independent Gaussian variables pointwise converging \mathbb{P} -a.e. Its sum is a Gaussian random variable with covariance matrix

$$\Sigma(q_1, q_2) := \sum_{n=1}^{\infty} \Sigma_n [\mathbf{E}'_n(q_1) - \mathbf{E}'_n(q_2)]^2, \quad (6.60)$$

satisfying

$$\alpha^2(q_1, q_2) \leq \Sigma(q_1, q_2) \mathbf{v} \cdot \mathbf{v} \leq \beta^2(q_1, q_2) \quad \text{for every } \mathbf{v} \in \mathbb{R}^d, |\mathbf{v}| = 1, \quad (6.61)$$

so that

$$\mathbb{E}_{\mathbb{P}} \left[e^{i u \cdot (\Xi(\cdot, q_1) - \Xi(\cdot, q_2))} \right] = \exp \left(-\frac{1}{2} \Sigma(q_1, q_2) u \cdot u \right). \quad (6.62)$$

Combining (6.62) with (6.57) we get

$$\int_{\mathcal{H}} |\hat{\mu}_X(u)|^2 d\mathbf{g}(X) = \mathbb{E}_{\mathbb{P}} |\hat{\mu}_\omega(u)|^2 = \int_{\mathbb{Q}^2} \exp \left(-\frac{1}{2} \Sigma(q_1, q_2) u \cdot u \right) d\mathbb{M}^{\otimes 2}(q_1, q_2). \quad (6.63)$$

We can plug (6.63) in (6.48) obtaining after a further application of Fubini's Theorem

$$\begin{aligned} L^2 &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{Q}^2} \exp \left(-\frac{1}{2} \Sigma(q_1, q_2) u \cdot u \right) d\mathbb{M}^{\otimes 2}(q_1, q_2) \right) du \\ &= \int_{\mathbb{Q}^2} \left(\int_{\mathbb{R}^d} \exp \left(-\frac{1}{2} \Sigma(q_1, q_2) u \cdot u \right) du \right) d\mathbb{M}^{\otimes 2}(q_1, q_2) \end{aligned}$$

Since the inner integral is

$$\int_{\mathbb{R}^d} \exp \left(-\frac{1}{2} \Sigma(q_1, q_2) u \cdot u \right) du \leq \int_{\mathbb{R}^d} \exp \left(-\frac{1}{2} \alpha^2(q_1, q_2) |u|^2 \right) du \leq \frac{(2\pi)^{d/2}}{\alpha^d(q_1, q_2)}$$

we conclude that

$$L^2 \leq (2\pi)^{d/2} \int_{\mathbb{Q}^2} \frac{1}{\alpha^d(q_1, q_2)} d\mathbb{M}^{\otimes 2}(q_1, q_2),$$

so that L^2 is finite if and only if (6.55) holds. \square

We apply the above result to the d -dimensional version of Example (6.22).

Example 6.29 (Sum of d -dimensional Gaussians with random signs). Let \mathbb{Q}, \mathbb{M} , and the Walsh system W_I as in Example (6.22): they form a complete orthonormal system for the “scalar” Hilbert space $\mathcal{H}' = L^2(\mathbb{Q}, \mathbb{M}; \mathbb{R})$. We now select a family of independent \mathbb{R}^d -Gaussian random variables $\xi_I \sim N(0, \Sigma_I)$ as in the previous discussion with

$$0 < \alpha_I^2 \leq \Sigma_I \mathbf{v} \cdot \mathbf{v} \leq \beta_I^2 \quad \text{for every } \mathbf{v} \in \mathbb{R}^d, \quad B^2 = \sum_I \beta_I^2 < \infty. \quad (6.64)$$

As in (6.53) we consider the random vector

$$\boldsymbol{\xi} = \sum_{I \in \mathcal{J}} \xi_I W_I \quad \text{with } \mathbf{g} = \boldsymbol{\xi}_{\#} \mathbb{P}. \quad (6.65)$$

We decompose the set \mathcal{J} of finite parts of \mathbb{N}_+ in the disjoint union of \mathcal{J}_n , $n \in \mathbb{N}$, with

$$\mathcal{J}_0 = \{\emptyset\}, \quad \mathcal{J}_n := \{I \in \mathcal{J} : \max I = n\}, \quad n > 0, \quad (6.66)$$

i.e.

$$\mathcal{J}_1 = \{\{1\}\}, \quad \mathcal{J}_2 = \{\{2\}, \{1, 2\}\}, \quad \mathcal{J}_3 = \{\{3\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\}, \dots$$

Notice that for every $I \in \mathcal{J}_n$ the corresponding Walsh function can be factorized as

$$W_I = \varepsilon_n W_{I'} \quad \text{for some } I' \subset \{1, \dots, n-1\}. \quad (6.67)$$

For each \mathcal{J}_n , $n \in \mathbb{N}$, we compute the contribution of α_I^2 to the total sum

$$A_n^2 := \sum_{I \in \mathcal{J}_n} \alpha_I^2, \quad \text{so that } A^2 = \sum_{n=0}^{\infty} A_n^2 = \sum_I \alpha_I^2 \leq B^2. \quad (6.68)$$

The next result shows that $G = \iota_{\#} \mathbf{g}$ is super-G-regular if A_n does not decay too fast.

Theorem 6.30. *If*

$$\sum_{n=1}^{\infty} \frac{1}{2^n A_n^d} < \infty \quad (6.69)$$

then $\mathbf{G} = \iota_{\#} \mathbf{g}$ is super-G-regular.

Notice that in the simplest case when $\Sigma_I = \lambda_I I_{d \times d}$ and $\alpha_I = \beta_I = \lambda_I$, asymptotic behaviours as $A_n \sim n^{-\theta}$ with $\theta > 1/2$ or $A_n \sim a^{-n}$ with $1 < a < 2^{1/d}$ comply with (6.69) and the summability of $n \mapsto A_n^2$ (corresponding to $\sum_I \beta_I^2 < \infty$).

Proof. We are in the setting of Theorem (6.28), so it is sufficient to check that (6.55) holds, where in our case

$$\alpha^2(q_1, q_2) = \sum_{I \in \mathcal{J}} \alpha_I^2 \left(W_I(q_1) - W_I(q_2) \right)^2 = \sum_{n \in \mathbb{N}} \sum_{I \in \mathcal{J}_n} \alpha_I^2 \left(W_I(q_1) - W_I(q_2) \right)^2 \quad (6.70)$$

For every pair $(q_1, q_2) \in \mathbb{Q}^2$ with $q_1 \neq q_2$ let us denote by $N(q_1, q_2)$ the first integer $n \in \mathbb{N}_+$ such that $\varepsilon_n(q_1) \neq \varepsilon_n(q_2)$:

$$N(q_1, q_2) := \min \left\{ n \in \mathbb{N}_+ : \varepsilon_n(q_1) \neq \varepsilon_n(q_2) \right\}. \quad (6.71)$$

Since $q_1 \neq q_2$ the set in (6.71) is not empty, so that $N(q_1, q_2)$ is well defined. Since $\varepsilon_k(q_1) = \varepsilon_k(q_2)$ for every $k < N(q_1, q_2)$ and therefore $W_{I'}(q_1) = W_{I'}(q_2)$ for every $I' \subset \{1, \dots, N(q_1, q_2) - 1\}$. Therefore the factorization (6.67) shows that

$$N = N(q_1, q_2), \quad I \in \mathcal{J}_N \quad \Rightarrow \quad W_I(q_1) = \varepsilon_N(q_1) W_{I'(q_1)} \neq W_I(q_2) = \varepsilon_N(q_2) W_{I'(q_2)} \quad (6.72)$$

so that

$$\alpha^2(q_1, q_2) \geq \sum_{I \in \mathcal{J}_N} \alpha_I^2 \left(W_I(q_1) - W_I(q_2) \right)^2 = 4 \sum_{I \in \mathcal{J}_N} \alpha_I^2 = 4A_N^2 \quad (6.73)$$

so that

$$\int_{\mathbb{Q}^2} \frac{1}{\alpha^d} d\mathbb{M}^{\otimes 2} \leq \sum_{n=1}^{\infty} \frac{1}{A_n^d} \mathbb{M}^{\otimes 2} \left[\{(q_1, q_2) \in \mathbb{Q}^2 : N(q_1, q_2) = n\} \right]. \quad (6.74)$$

Recall now that ε_n are independent and $\mathbb{M}^{\otimes 2}[\varepsilon_k(q_1) \neq \varepsilon_k(q_2)] = \mathbb{M}^{\otimes 2}[\varepsilon_k(q_1) = \varepsilon_k(q_2)] = 1/2$ for all $k \in \mathbb{N}_+$. We thus obtain

$$\mathbb{M}^{\otimes 2} \left[\{(q_1, q_2) \in \mathbb{Q}^2 : N(q_1, q_2) = n\} \right] = \frac{1}{2^n}$$

and inserting this expression in (6.74) we eventually get

$$\int_{\mathbb{Q}^2} \frac{1}{\alpha^d} d\mathbb{M}^{\otimes 2} \leq \sum_{n=1}^{\infty} \frac{1}{2^n A_n^d} < +\infty$$

thanks to (6.69). □

References

- [Acc+25] Beatrice Acciaio et al. *Absolutely Continuous Curves of Stochastic Processes*. 2025. arXiv: [2506.13634](https://arxiv.org/abs/2506.13634) [[math.PR](#)]. URL: <https://arxiv.org/abs/2506.13634>.
- [AGS08] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Second. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008, pp. x+334. ISBN: 978-3-7643-8721-1.
- [Aro76] N. Aronszajn. “Differentiability of Lipschitzian mappings between Banach spaces”. In: *Studia Math.* 57.2 (1976), pp. 147–190. ISSN: 0039-3223,1730-6337. DOI: [10.4064/sm-57-2-147-190](https://doi.org/10.4064/sm-57-2-147-190). URL: <https://doi.org/10.4064/sm-57-2-147-190>.
- [Ber69] Simeon M. Berman. “Local times and sample function properties of stationary Gaussian processes”. In: *Transactions of the American Mathematical Society* 137.0 (1969), pp. 277–299. ISSN: 0002-9947. DOI: [10.1090/s0002-9947-1969-0239652-5](https://doi.org/10.1090/s0002-9947-1969-0239652-5).
- [Bia+08] Francesca Biagini et al. *Stochastic calculus for fractional Brownian motion and applications*. Probability and its Applications (New York). Springer-Verlag London, Ltd., London, 2008, pp. xii+329. ISBN: 978-1-85233-996-8. DOI: [10.1007/978-1-84628-797-8](https://doi.org/10.1007/978-1-84628-797-8). URL: <https://doi.org/10.1007/978-1-84628-797-8>.

- [BL00] Yoav Benyamini and Joram Lindenstrauss. *Geometric nonlinear functional analysis. Vol. 1*. Vol. 48. American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 2000, pp. xii+488. ISBN: 0-8218-0835-4. DOI: [10.1090/coll/048](https://doi.org/10.1090/coll/048). URL: <https://doi.org/10.1090/coll/048>.
- [Bog07] V. I. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007, Vol. I: xviii+500 pp., Vol. II: xiv+575. DOI: [10.1007/978-3-540-34514-5](https://doi.org/10.1007/978-3-540-34514-5). URL: <https://doi.org/10.1007/978-3-540-34514-5>.
- [Bog18] V. I. Bogachev. “Negligible sets in infinite-dimensional spaces”. In: *Anal. Math.* 44.3 (2018), pp. 299–323. ISSN: 0133-3852,1588-273X. DOI: [10.1007/s10476-018-0503-7](https://doi.org/10.1007/s10476-018-0503-7). URL: <https://doi.org/10.1007/s10476-018-0503-7>.
- [Bog84] V. I. Bogachev. “Negligible sets in locally convex spaces”. In: *Mat. Zametki* 36.1 (1984), pp. 51–64. ISSN: 0025-567X.
- [Bog98] Vladimir I. Bogachev. *Gaussian measures*. Vol. 62. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1998, pp. xii+433. ISBN: 0-8218-1054-5. DOI: [10.1090/surv/062](https://doi.org/10.1090/surv/062). URL: <https://doi.org/10.1090/surv/062>.
- [Bre20] Yann Brenier. *Examples of Hidden Convexity in Nonlinear PDEs*. 2020. eprint: [hal-02928398](https://hal.science/hal-02928398). URL: <https://hal.science/hal-02928398/document>.
- [Bre91] Yann Brenier. “Polar factorization and monotone rearrangement of vector-valued functions”. In: *Comm. Pure Appl. Math.* 44.4 (1991), pp. 375–417. ISSN: 0010-3640,1097-0312. DOI: [10.1002/cpa.3160440402](https://doi.org/10.1002/cpa.3160440402). URL: <https://doi.org/10.1002/cpa.3160440402>.
- [BVK25] Clément Bonet, Christophe Vauthier, and Anna Korba. “Flowing Datasets with Wasserstein over Wasserstein Gradient Flows”. In: *arXiv* (2025). DOI: [10.48550/arxiv.2506.07534](https://arxiv.org/abs/2506.07534). eprint: [2506.07534](https://arxiv.org/abs/2506.07534).
- [Car13] Pierre Cardaliaguet. *Notes on mean field games. From P. L. Lions’ lectures at College de France*. 2013. eprint: [preprint](https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf). URL: <https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf>.
- [CD18] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications. I*. Vol. 83. Probability Theory and Stochastic Modelling. Mean field FBSDEs, control, and games. Springer, Cham, 2018, pp. xxv+713.
- [CL24] Marta Catalano and Hugo Lavenant. “Hierarchical Integral Probability Metrics: A distance on random probability measures with low sample complexity”. In: *ICML’24: Proceedings of the 41st International Conf. on Machine Learning* (2024).
- [CM89] Juan Antonio Cuesta and Carlos Matrán. “Notes on the Wasserstein metric in Hilbert spaces”. In: *Ann. Probab.* 17.3 (1989), pp. 1264–1276. ISSN: 0091-1798,2168-894X. URL: [http://links.jstor.org/sici?sici=0091-1798\(198907\)17:3%3C1264:NOTWMI%3E2.0.CO;2-J&origin=MSN](http://links.jstor.org/sici?sici=0091-1798(198907)17:3%3C1264:NOTWMI%3E2.0.CO;2-J&origin=MSN).
- [Csö99] Marianna Csörnyei. “Aronszajn null and Gaussian null sets coincide”. In: *Israel J. Math.* 111 (1999), pp. 191–201. ISSN: 0021-2172,1565-8511. DOI: [10.1007/BF02810684](https://doi.org/10.1007/BF02810684). URL: <https://doi.org/10.1007/BF02810684>.
- [CSS23a] Giulia Cavagnari, Giuseppe Savaré, and Giacomo Enrico Sodini. *A Lagrangian approach to totally dissipative evolutions in Wasserstein spaces*. 2023. DOI: [10.48550/arXiv.2305.05211](https://arxiv.org/abs/2305.05211). arXiv: [2305.05211](https://arxiv.org/abs/2305.05211) [math.FA].
- [CSS23b] Giulia Cavagnari, Giuseppe Savaré, and Giacomo Enrico Sodini. “Dissipative probability vector fields and generation of evolution semigroups in Wasserstein spaces”. In: *Probab. Theory Related Fields* 185.3-4 (2023), pp. 1087–1182. ISSN: 0178-8051,1432-2064. DOI: [10.1007/s00440-022-01148-7](https://doi.org/10.1007/s00440-022-01148-7). URL: <https://doi.org/10.1007/s00440-022-01148-7>.

- [CSS25] Giulia Cavagnari, Giuseppe Savaré, and Giacomo Enrico Sodini. “Extension of monotone operators and Lipschitz maps invariant for a group of isometries”. In: *Canad. J. Math.* 77.1 (2025), pp. 149–186. ISSN: 0008-414X,1496-4279. DOI: [10.4153/S0008414X23000846](https://doi.org/10.4153/S0008414X23000846). URL: <https://doi.org/10.4153/S0008414X23000846>.
- [Del20] Lorenzo Dello Schiavo. “A Rademacher-type theorem on L^2 -Wasserstein spaces over closed Riemannian manifolds”. In: *J. Funct. Anal.* 278.6 (2020), pp. 108397, 57. ISSN: 0022-1236,1096-0783. DOI: [10.1016/j.jfa.2019.108397](https://doi.org/10.1016/j.jfa.2019.108397). URL: <https://doi.org/10.1016/j.jfa.2019.108397>.
- [DM82] Claude Dellacherie and Paul-André Meyer. *Probabilities and potential. B*. Vol. 72. North-Holland Mathematics Studies. Theory of martingales, Translated from the French by J. P. Wilson. North-Holland Publishing Co., Amsterdam, 1982, pp. xvii+463. ISBN: 0-444-86526-8.
- [EP25] Pedram Emami and Brendan Pass. “Optimal transport with optimal transport cost: the Monge-Kantorovich problem on Wasserstein spaces”. In: *Calc. Var. Partial Differential Equations* 64.2 (2025), Paper No. 43, 11. ISSN: 0944-2669,1432-0835. DOI: [10.1007/s00526-024-02905-3](https://doi.org/10.1007/s00526-024-02905-3). URL: <https://doi.org/10.1007/s00526-024-02905-3>.
- [ET76] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*. Vol. Vol. 1. Studies in Mathematics and its Applications. Translated from the French. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1976, pp. ix+402.
- [FHS23] Massimo Fornasier, Pascal Heid, and Giacomo Enrico Sodini. “Approximation Theory, Computing, and Deep Learning on the Wasserstein Space”. In: *arXiv* (2023). DOI: [10.48550/arxiv.2310.19548](https://doi.org/10.48550/arxiv.2310.19548). eprint: [2310.19548](https://arxiv.org/abs/2310.19548).
- [FSS23] Massimo Fornasier, Giuseppe Savaré, and Giacomo Enrico Sodini. “Density of subalgebras of Lipschitz functions in metric Sobolev spaces and applications to Wasserstein Sobolev spaces”. In: *J. Funct. Anal.* 285.11 (2023), Paper No. 110153, 76. ISSN: 0022-1236,1096-0783. DOI: [10.1016/j.jfa.2023.110153](https://doi.org/10.1016/j.jfa.2023.110153). URL: <https://doi.org/10.1016/j.jfa.2023.110153>.
- [GH80] Donald Geman and Joseph Horowitz. “Occupation densities”. In: *Ann. Probab.* 8.1 (1980), pp. 1–67. ISSN: 0091-1798,2168-894X. URL: [http://links.jstor.org/sici?sici=0091-1798\(198002\)8:1%3C1:OD%3E2.0.CO;2-M&origin=MSN](http://links.jstor.org/sici?sici=0091-1798(198002)8:1%3C1:OD%3E2.0.CO;2-M&origin=MSN).
- [GT19] Wilfrid Gangbo and Adrian Tudorascu. “On differentiability in the Wasserstein space and well-posedness for Hamilton-Jacobi equations”. In: *J. Math. Pures Appl. (9)* 125 (2019), pp. 119–174. ISSN: 0021-7824,1776-3371. DOI: [10.1016/j.matpur.2018.09.003](https://doi.org/10.1016/j.matpur.2018.09.003). URL: <https://doi.org/10.1016/j.matpur.2018.09.003>.
- [JKO98] Richard Jordan, David Kinderlehrer, and Felix Otto. “The variational formulation of the Fokker-Planck equation”. In: *SIAM J. Math. Anal.* 29.1 (1998), pp. 1–17. ISSN: 0036-1410,1095-7154. DOI: [10.1137/S0036141096303359](https://doi.org/10.1137/S0036141096303359). URL: <https://doi.org/10.1137/S0036141096303359>.
- [Kal17] Olav Kallenberg. *Random measures, theory and applications*. Vol. 77. Probability Theory and Stochastic Modelling. Springer, Cham, 2017, pp. xiii+694. DOI: [10.1007/978-3-319-41598-7](https://doi.org/10.1007/978-3-319-41598-7). URL: <https://doi.org/10.1007/978-3-319-41598-7>.
- [KR24] Vitalii Konarovskyi and Max-K. von Renesse. “Reversible coalescing-fragmentating Wasserstein dynamics on the real line”. In: *J. Funct. Anal.* 286.8 (2024), Paper No. 110342, 60. ISSN: 0022-1236,1096-0783. DOI: [10.1016/j.jfa.2024.110342](https://doi.org/10.1016/j.jfa.2024.110342). URL: <https://doi.org/10.1016/j.jfa.2024.110342>.

- [KS84] M. Knott and C. S. Smith. “On the optimal mapping of distributions”. In: *J. Optim. Theory Appl.* 43.1 (1984), pp. 39–49. ISSN: 0022-3239,1573-2878. DOI: [10.1007/BF00934745](https://doi.org/10.1007/BF00934745). URL: <https://doi.org/10.1007/BF00934745>.
- [LS25] Hugo Lavenant and Giuseppe Savaré. “Continuous transformations of probability measures and push-forward maps”. In: *In preparation* (2025).
- [McC97] Robert J. McCann. “A convexity principle for interacting gases”. In: *Adv. Math.* 128.1 (1997), pp. 153–179. ISSN: 0001-8708,1090-2082. DOI: [10.1006/aima.1997.1634](https://doi.org/10.1006/aima.1997.1634). URL: <https://doi.org/10.1006/aima.1997.1634>.
- [MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Vol. 30. Cambridge Series in Statistical and Probabilistic Mathematics. With an appendix by Oded Schramm and Wendelin Werner. Cambridge University Press, Cambridge, 2010, pp. xii+403. ISBN: 978-0-521-76018-8. DOI: [10.1017/CB09780511750489](https://doi.org/10.1017/CB09780511750489). URL: <https://doi.org/10.1017/CB09780511750489>.
- [MS20] Matteo Muratori and Giuseppe Savaré. “Gradient flows and evolution variational inequalities in metric spaces. I: Structural properties”. In: *J. Funct. Anal.* 278.4 (2020), pp. 108347, 67. ISSN: 0022-1236. DOI: [10.1016/j.jfa.2019.108347](https://doi.org/10.1016/j.jfa.2019.108347). URL: <https://doi.org/10.1016/j.jfa.2019.108347>.
- [NS21] Emanuele Naldi and Giuseppe Savaré. “Weak topology and Opial property in Wasserstein spaces, with applications to gradient flows and proximal point algorithms of geodesically convex functionals”. In: *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* 32.4 (2021), pp. 725–750. ISSN: 1120-6330. DOI: [10.4171/rlm/955](https://doi.org/10.4171/rlm/955). URL: <https://doi.org/10.4171/rlm/955>.
- [Phe78] R. R. Phelps. “Gaussian null sets and differentiability of Lipschitz map on Banach spaces”. In: *Pacific J. Math.* 77.2 (1978), pp. 523–531. ISSN: 0030-8730,1945-5844. URL: <http://projecteuclid.org/euclid.pjm/1102806464>.
- [Pic19] Benedetto Piccoli. “Measure differential equations”. In: *Arch. Ration. Mech. Anal.* 233.3 (2019), pp. 1289–1317. ISSN: 0003-9527,1432-0673. DOI: [10.1007/s00205-019-01379-4](https://doi.org/10.1007/s00205-019-01379-4). URL: <https://doi.org/10.1007/s00205-019-01379-4>.
- [Pit78] Loren D. Pitt. “Local times for Gaussian vector fields”. In: *Indiana Univ. Math. J.* 27.2 (1978), pp. 309–330. ISSN: 0022-2518,1943-5258. DOI: [10.1512/iumj.1978.27.27024](https://doi.org/10.1512/iumj.1978.27.27024). URL: <https://doi.org/10.1512/iumj.1978.27.27024>.
- [PS25] Alessandro Pinzi and Giuseppe Savaré. *Nested superposition principle for random measures and the geometry of the Wasserstein on Wasserstein space*. 2025.
- [Roc66] R. T. Rockafellar. “Characterization of the subdifferentials of convex functions”. In: *Pacific J. Math.* 17 (1966), pp. 497–510. ISSN: 0030-8730,1945-5844. URL: <http://projecteuclid.org/euclid.pjm/1102994514>.
- [RR90] L. Rüschendorf and S. T. Rachev. “A characterization of random variables with minimum L^2 -distance”. In: *J. Multivariate Anal.* 32.1 (1990), pp. 48–54. ISSN: 0047-259X,1095-7243. DOI: [10.1016/0047-259X\(90\)90070-X](https://doi.org/10.1016/0047-259X(90)90070-X). URL: [https://doi.org/10.1016/0047-259X\(90\)90070-X](https://doi.org/10.1016/0047-259X(90)90070-X).
- [RR98] Svetlozar T. Rachev and Ludger Rüschendorf. *Mass transportation problems. Vol. I, II. Probability and its Applications* (New York). Theory. Springer-Verlag, New York, 1998.
- [RS09] Max-K. von Renesse and Karl-Theodor Sturm. “Entropic measure and Wasserstein diffusion”. In: *Ann. Probab.* 37.3 (2009), pp. 1114–1191. ISSN: 0091-1798,2168-894X. DOI: [10.1214/08-AOP430](https://doi.org/10.1214/08-AOP430). URL: <https://doi.org/10.1214/08-AOP430>.

- [Sch73] Laurent Schwartz. *Radon measures on arbitrary topological spaces and cylindrical measures*. Vol. No. 6. Tata Institute of Fundamental Research Studies in Mathematics. Tata Institute of Fundamental Research, Bombay; by Oxford University Press, London, 1973, pp. xii+393.
- [Stu11] Karl-Theodor Sturm. “Entropic measure on multidimensional spaces”. In: *Seminar on Stochastic Analysis, Random Fields and Applications VI*. Vol. 63. Progr. Probab. Birkhäuser/Springer Basel AG, Basel, 2011, pp. 261–277. ISBN: 978-3-0348-0020-4. DOI: [10.1007/978-3-0348-0021-1_17](https://doi.org/10.1007/978-3-0348-0021-1_17). URL: https://doi.org/10.1007/978-3-0348-0021-1_17.
- [Vil09] Cédric Villani. *Optimal transport*. Vol. 338. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Old and new. Springer-Verlag, Berlin, 2009, pp. xxii+973. ISBN: 978-3-540-71049-3. DOI: [10.1007/978-3-540-71050-9](https://doi.org/10.1007/978-3-540-71050-9). URL: <https://doi.org/10.1007/978-3-540-71050-9>.
- [Zaj79] Luděk Zajíček. “On the differentiation of convex functions in finite and infinite dimensional spaces”. In: *Czechoslovak Math. J.* 29(104).3 (1979), pp. 340–348. ISSN: 0011-4642.