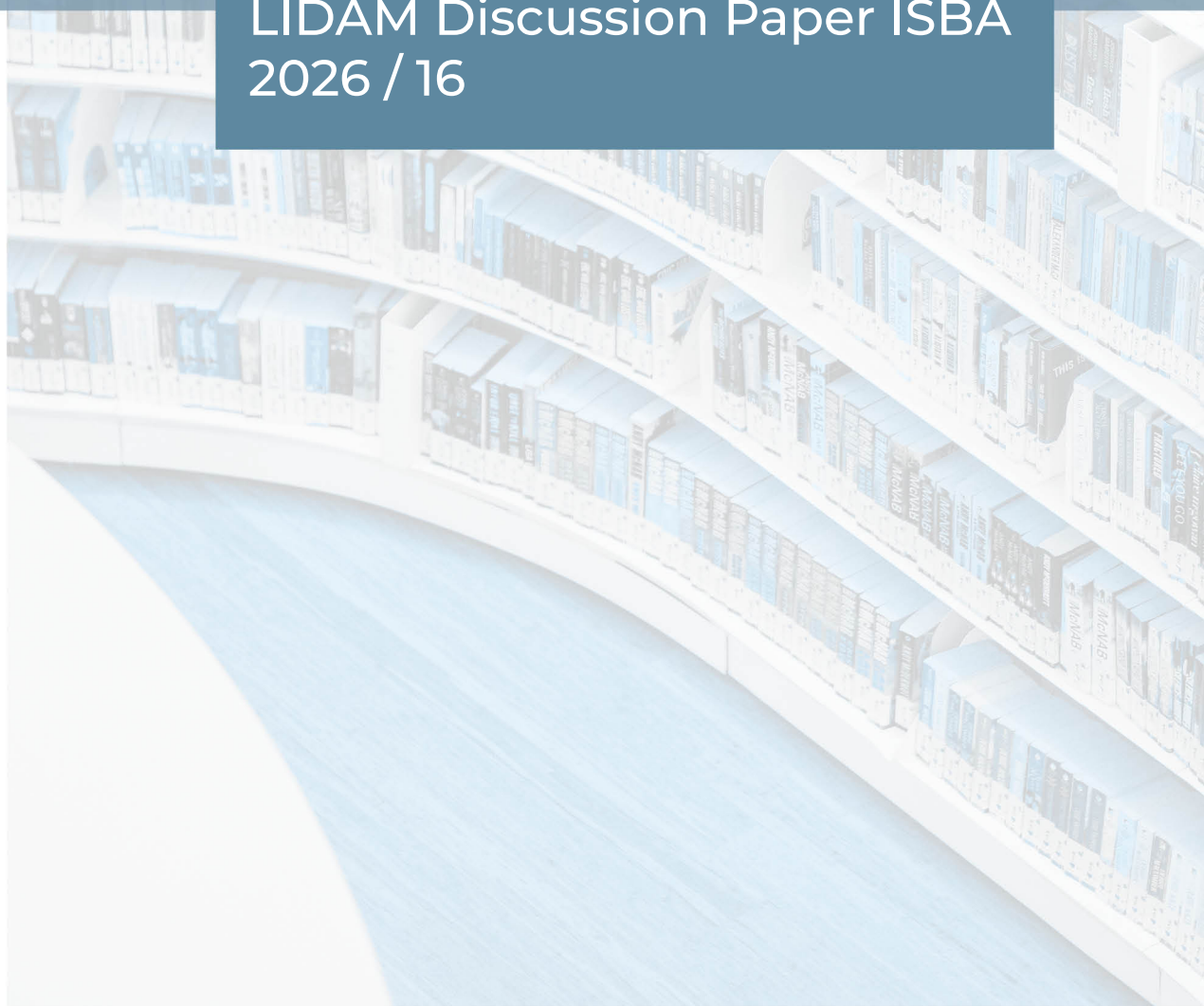
ZERO-COUPLINGS OF INFINITE MEASURES WITH CYCLICALLY MONOTONE SUPPORT AND MULTIVARIATE REGULAR VARIATION

Alexandre Reber, Anne Sabourin, Johan Segers, Cees de Valk

LIDAM Discussion Paper ISBA
2026 / 16



UCLouvain

ISBA | LIDAM
Louvain Institute of Data Analysis and Modeling in economics and statistics

ISBA

Voie du Roman Pays 20 - L1.04.01

B-1348 Louvain-la-Neuve

Email : lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/isba/publication

arXiv:2605.10491v1 [math.PR] 11 May 2026

Submitted to Bernoulli

Zero-couplings of infinite measures with cyclically monotone support and multivariate regular variation

ALEXANDRE REBER[1,a], ANNE SABOURIN[2,b], JOHAN SEGERS[3,c] and CEES DE VALK[4,d]

[1]*Aarhus University, Department of Mathematics, Ny Munkegade 118, 8000 Aarhus C, Denmark,* [a]*alexandre.reber@math.au.dk*
[2]*Université Paris Cité, Université Paris Saclay, ENS Paris Saclay, CNRS, SSA, INSERM, Centre Borelli, F-75006, Paris, France,* [b]*anne.sabourin@u-paris.fr*
[3]*KU Leuven, Department of Mathematics, Celestijnenlaan 200B, 3001 Heverlee, Belgium, and UCLouvain, LIDAM/ISBA,* [c]*jjjsegers@kuleuven.be*
[4]*Koninklijk Nederlands Meteorologisch Instituut, PO Box 201, NL-3730 AE De Bilt, Netherlands,* [d]*cees.de.valk@knmi.nl*

We study cyclically monotone transport plans between measures in $\mathrm{M}_0(\mathbb{R}^d)$, the class of Borel measures on $\mathbb{R}^d \setminus \{0\}$ that are finite on sets bounded away from the origin but may have infinite total mass. We avoid moment assumptions and allow the transport cost to be infinite. This framework naturally arises for exponent measures in multivariate regular variation and includes other examples such as Lévy measures.

We introduce the notion of a zero-coupling and establish existence of cyclically monotone zero-couplings for arbitrary pairs of measures in $\mathrm{M}_0(\mathbb{R}^d)$. Under a Hausdorff-dimension condition on the first measure and when at least one of the two measures has infinite mass, we prove uniqueness of the cyclically monotone zero-coupling, yielding an analogue of the Brenier–McCann theorem in this infinite-measure setting. We further derive a representation of such couplings through gradients of closed convex functions and identify conditions under which the zero-coupling is proper in the sense that the second measure is equal to the restriction to the punctured space of the push-forward of the first measure by a cyclically monotone transport map.

Finally, we apply these results to regularly varying probability measures. We show that a cyclically monotone coupling between two such distributions admits a tail limit that coincides with the unique proper cyclically monotone zero-coupling between the corresponding exponent measures.

Keywords: Cyclical monotonicity; Fell topology; Multivariate extremes; Optimal transport; Regular variation

1. Introduction

1.1. General motivation

We study measure transportation by cyclically monotone mappings between nonnegative Borel measures on Euclidean space with possibly infinite total mass that are finite on complements of neighborhoods of the origin. Such measures arise in the definition of multivariate regular variation; Lévy measures provide another example. Especially in the first setting, which appears naturally in extreme value theory, the usual moment assumptions that are key to classical results in optimal transport theory cannot be taken for granted.

Our goal is to strengthen the theoretical foundation for the use of cyclically monotone mappings in statistics, where such maps are of intrinsic interest—even when they cannot be interpreted as optimal

transport plans due to infinite transport costs. A prominent example is the theory of center-outward multivariate quantiles [7,12,18,19], which provides a notion of rank statistics for multivariate observations. Our results are intended to lay the groundwork for extensions to multivariate extreme quantile contours. A difference to the existing literature is that a natural reference distribution in an extreme value context is not spherical uniform, but would typically factorize into a product of a reference angular distribution and a heavy-tailed radial component. In another direction, monotone maps and optimal transport play a central role in generative modeling [1,24]. We anticipate that our results may help extend these methods to the generation of synthetic extreme samples.

Our focus is on transport maps and an appropriate notion of coupling between measures in the set $\mathrm{M}_0(\mathbb{R}^d)$, the set of Borel measures on $\mathbb{R}^d_{\not\to 0} = \mathbb{R}^d \setminus \{0\}$ that are finite on sets bounded away from the origin but may have infinite total mass [22,23]. The space $\mathrm{M}_0(\mathbb{R}^d)$ is equipped with a notion of convergence in duality with the set $C_0(\mathbb{R}^d)$ of real-valued bounded continuous functions whose support is bounded away from the origin. A typical assumption in multivariate extreme value theory is that observations are drawn from a probability distribution Q on $\mathbb{R}^d$ that is regularly varying: there exists a function b on the positive reals and a limit measure $\nu \in \mathrm{M}_0(\mathbb{R}^d)$ such that

$$t\,Q(b(t)\,\cdot\,) \to \nu, \qquad t \to \infty, \tag{1}$$

in $\mathrm{M}_0(\mathbb{R}^d)$. In practice, ν can be used to approximate the distribution Q in regions of $\mathbb{R}^d$ far from the origin, where few or no observations are available.

1.2. Related work, zero-couplings, proper couplings

Our interest lies in transport maps between measures $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ with arbitrarily heavy tails. In such cases, neither the Monge problem for the L^2 cost,

$$\min_{T : T_\# \mu = \nu} \int \|x - T(x)\|^2 \,\mathrm{d}\mu(x), \tag{2}$$

nor the Kantorovich problem,

$$\min_{\gamma \in \Pi(\mu,\nu)} \int \|x - y\|^2 \,\mathrm{d}\gamma(x,y), \tag{3}$$

may admit solutions with finite objective, since μ and ν need not have finite second moments themselves. Here, $T_\#\mu$ is the push-forward measure of μ induced by T, see Section 1.5 below, while $\Pi(\mu,\nu)$ denotes the set of coupling measures between μ and ν, i.e., measures on $\mathbb{R}^d \times \mathbb{R}^d$ with margins μ and ν. As we will explain below, we will actually use a slightly weakened notion of couplings, called zero-couplings, to deal with the possible singularity at the origin.

The special case where μ, ν are probability measures with finite second moments falls within the classical framework of optimal transport, for which there are many excellent and extensive textbook introductions [e.g. 4,34,37]. If μ places no mass on sets of Hausdorff dimension less than $d-1$, Brenier's Theorem [9,10] (attributed also to [33]; see [37], Chapter 9, and the bibliographic comments on page 212 on many other seminal papers) guarantees that the Kantorovich problem admits a unique solution π, which can be written as $\pi = (\mathrm{Id} \times T)_\# \mu$, where the map T is defined μ-almost everywhere as $T = \nabla\psi$ for some convex function ψ and which, up to a μ-null set, also solves the Monge problem (2). A key property of T—or, equivalently, of the support $\mathrm{spt}\,\pi$ of π—is cyclical monotonicity, the definition of which is recalled in Eq. (6) below. Conversely, if a coupling measure π has cyclically monotone support, then it is necessarily optimal in the sense that it solves the Kantorovich problem (3), see [6,14,27,35]. This is why

we focus on cyclically monotone transport plans (maps and couplings) even when the expected cost may be infinite. For simplicity, we focus on the notion of cyclical monotonicity with respect to the quadratic cost function $c(x, y) = \|x - y\|^2$, deferring the study of general cost functions $c(x, y)$ to future research.

Brenier's result has been extended in two distinct directions:

(i) Positive and potentially infinite Borel measures are considered in several articles, though typically under moment assumptions that allow for the study of the Kantorovich problem. For instance, [17] introduces a transportation metric between Lévy measures in the context of viscosity equations, a metric that [11] later uses to construct an index of dependence between random measures. In [16], solutions of the Kantorovich problem are linked to gradient flows that solve the heat equation with Dirichlet boundary conditions. A common feature of these references is the presence of a specific subset of the domain—the origin in [11,17] or the boundary in [16]—which acts as a potential "sink" or "reservoir" of mass in the transport problem. These works also introduce extended notions of coupling, called "admissible couplings" in [16,17] and "extended couplings" in [11], in which the first and second marginals are required to coincide with μ and ν only on the complement of the reservoir or sink. In our paper, we adopt the terminology "zero-couplings" to describe such couplings, reflecting the special role of the origin in $\mathrm{M}_0(\mathbb{R}^d)$.

(ii) In [25], the condition that second moments exist is relaxed, and existence and uniqueness of a cyclically monotone coupling and the associated transport map are established under assumptions similar to those in [9]. Let $\mathcal{P}(\mathbb{R}^d)$ denote the set of probability measures on $\mathbb{R}^d$ and $\Pi_{\mathrm{cm}}(P, Q)$ the set of coupling measures of $P, Q \in \mathcal{P}(\mathbb{R}^d)$ with cyclically monotone support. The subdifferential of a convex function $\psi : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ is denoted by $\partial\psi$ (see Section A.2 for preliminaries).

Theorem 1.1 (McCann, [25]). *Let $P, Q \in \mathcal{P}(\mathbb{R}^d)$, with P vanishing on sets of Hausdorff dimension at most $d - 1$. The set $\Pi_{\mathrm{cm}}(P, Q)$ contains a single element π. For any closed convex function ψ such that $\operatorname{spt}\pi \subset \partial\psi$, the gradient $\nabla\psi$ is a transport map, $(\nabla\psi)_{\#}P = Q$. This transport map is unique P-a.e.*

In this context, the optimality of a transport map is not applicable if the Kantorovich problem has no solutions for which the objective is finite. Still, the focus on cyclical monotonicity is justified by Brenier's theorem and the earlier cited results on the sufficiency of cyclical monotonicity for optimality in case second moments do exist.

1.3. Problem statement and main results

To our knowledge, the following questions have not been treated in the literature for infinite measures without finite-moment assumptions:

- (Q1) **Existence, Uniqueness:** Given two potentially infinite measures $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$, is there a (unique) cyclically monotone zero-coupling (see Definition 2.2) of μ, ν?
- (Q2) **Transport map:** Given such a zero-coupling γ, and a convex function ψ such that $\operatorname{spt}\gamma \subset \partial\psi$, is it necessarily the case that the map $\nabla\psi$ pushes μ forward to ν?
- (Q3) **Stability:** In case μ, ν are limit measures of P, Q, respectively, in the definition in Eq. (1) of regular variation, and if π is a cyclically monotone coupling between P and Q, what are the connections between $(P, Q, \pi, \nabla\psi)$ and $(\mu, \nu, \gamma, \nabla\bar{\psi})$, where $\nabla\psi$ and $\nabla\bar{\psi}$ are transport maps from P to Q and from μ to ν, respectively?

The answer to (Q1), and in particular the uniqueness part, involves substantially more technicality than in [25], due to the potentially infinite masses of sets. A cyclically monotone zero-coupling always exists (Theorem 2.8) and a sufficient condition for it to be unique is that one of the two measures vanishes on

sets of Hausdorff dimension at most $d-1$ and, moreover, one of the two measures has infinite mass (Theorem 4.1). We give a counter-example to the uniqueness if both measures are absolutely continuous and have finite mass (Remark 4.2).

Regarding (Q2), if P, Q in Theorem 1.1 are replaced by infinite measures $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$, the conclusion is no longer valid: even if the cyclically monotone zero-coupling between μ and ν exists and is unique, it may not be possible to write ν as $(\nabla\psi)_{\#}\mu$ for a closed convex function ψ, as the following counterexample shows.

Example 1.2. Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R})$ and suppose that $\mu(\mathbb{R}_{>0}) = \nu(\mathbb{R}_{<0}) = \infty$ while $\mu(\mathbb{R}_{<0}) = \nu(\mathbb{R}_{>0}) = 0$. The measure $\gamma = \mu \otimes \delta_0 + \delta_0 \otimes \nu$ is a zero-coupling of μ and ν [see Definition 2.2: we have $\gamma(A \times \mathbb{R}) = \mu(A)$ and $\gamma(\mathbb{R} \times A) = \nu(A)$ for every Borel set $A \subset \mathbb{R} \setminus \{0\}$], and its support is contained in $\big(\{0\} \times \mathbb{R}_{\leqslant 0}\big) \cup \big(\mathbb{R}_{\geqslant 0} \times \{0\}\big)$ and therefore (cyclically) monotone. Still, we claim that there cannot exist a non-decreasing function $T : \mathbb{R} \to \mathbb{R}$ such that $\nu = T_{\#}\mu$. Indeed, the existence of such a map T would lead to a contradiction, as we show next. Let $x > 0$. If $T(x) < 0$, then, since $\mu(\mathbb{R}_{>0}) = \infty$ but $\mu((x, \infty)) < \infty$ (as $\mu \in \mathrm{M}_0(\mathbb{R})$), we would have

$$\nu((-\infty, T(x)]) = \mu(\{z \in \mathbb{R} \setminus \{0\} \mid T(z) \leqslant T(x)\}) \geqslant \mu((0, x]) = \infty,$$

in contradiction to the assumption that $\nu \in \mathrm{M}_0(\mathbb{R})$. Hence, we have $T(x) \geqslant 0$ for all $x > 0$. But then we obtain another contradiction, since

$$\infty = \nu(\mathbb{R}_{<0}) = \mu(\{z \in \mathbb{R} \setminus \{0\} \mid T(z) < 0\}) \leqslant \mu(\mathbb{R}_{<0}) = 0.$$

Therefore, a map T with the stated properties cannot exist, proving the claim. □

In an earlier version of this paper, de Valk and Segers [15], ordinary couplings rather than zero-couplings between infinite measures were considered instead. This eventually led to an error: Example 1.2 above is in contradiction to Theorem 4.4 in [15]. The crucial point is that the measure γ in the example is not a coupling between μ and ν: we have $\gamma(\{0\} \times \mathbb{R}) = \infty$ while $\mu(\{0\})$ is either undefined or zero, and similarly $\gamma(\mathbb{R} \times \{0\}) = \infty$ is not equal to $\nu(\{0\})$. The mistake can be traced back to Lemma A.2 in [15]: the M_0-limit γ of a sequence of coupling measures γ_n of measures μ_n and ν_n (see Section 2.2 for the definition of M_0-convergence) is not necessarily a coupling itself between the M_0-limits μ and ν, but only a zero-coupling.

As will appear clearly at later stages of the paper, the answer to (Q2) in Theorem 2.9 depends on whether the origin plays the role of a reservoir. A zero-coupling provides a cyclically monotone transport map $\nabla\psi$ if and only if the coupling is what we call *proper*, that is, $\gamma(\{0\} \times \mathbb{R}^d_{\smallsetminus 0}) = 0$: the origin $0 \in \mathbb{R}^d$ does not play the role of reservoir of mass. Sufficient conditions for cyclically monotone couplings to be proper are given in Section 3.

Our main results in relation to (Q3) are summarized in Fig. 1. A sufficient condition for them to hold is that (a) both P and μ vanish on sets of Hausdorff dimension at most $d-1$, referred to as "small sets", and (b) $\operatorname{spt}\nu \subset \operatorname{spt}\mu$. Weaker assumptions for (b) are asserted in Theorem 5.1. In the commutative diagram in Fig. 1, the notation RV denotes the operator that associates a probability measure to the limit measure in the definition in Eq. (1) of regular variation (or an extension thereof in the case of the coupling measure). There are two ways to travel from probability measures P, Q in the top-left of the diagram to the unique cyclically monotone zero-coupling γ between μ, ν in the bottom-right:

- first take the limit measures μ and ν in the definition of regular variation of P and Q respectively, and then consider the unique cyclically monotone zero-coupling γ between them;
- first consider the unique cyclically monotone coupling π between P and Q, and then take its limit measure γ in an extension of the definition of regular variation.

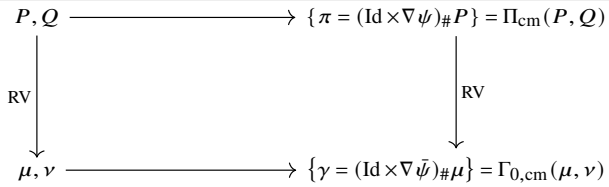


Figure 1. Probability measures P and Q on $\mathbb{R}^d$, with P vanishing on small sets; the coupling π of P and Q with cyclically monotone support contained in the subdifferential of a convex function ψ; the limit measures μ and ν on $\mathbb{R}^d_{\diagdown 0}$ of P and Q, respectively, in the definition of regular variation, with μ vanishing on small sets; the zero-coupling γ of μ and ν with cyclically monotone support contained in the subdifferential of a convex function $\bar{\psi}$, this subdifferential being the limit in the Fell topology of the one of the subdifferential of ψ. Under an additional condition on the support of μ, the zero-coupling γ is proper, so that γ and ν can be written as push-forward measures of μ by $\mathrm{Id}\times\nabla\bar{\psi}$ and $\nabla\bar{\psi}$, respectively.

Moreover, the transport map $\nabla\bar{\psi}$ associated with the zero-coupling γ of μ and ν is the tail limit of the transport map $\nabla\psi$ associated with the coupling π of P and Q.

1.4. Outline of the paper

In Section 2, zero-marginals and zero-couplings are introduced, and sequences of zero-couplings and sets containing their supports are studied. The existence of zero-couplings whose supports are cyclically monotone is proved and a representation of such zero-couplings using gradients of closed convex functions is derived.

Proper zero-couplings are the focus of Section 3. We provide conditions on the supports of the two measures to be coupled for the zero-coupling between them with (cyclically) monotone support to be proper. The latter property is crucial for the representation in Section 2 to yield a push-forward formula for one measure in terms of the other, precluding situations as in the counterexample in Example 1.2.

In Section 4, an extension of Brenier's Theorem to measures in $\mathrm{M}_0(\mathbb{R}^d)$ with potentially infinite total mass is provided. When at least one of the zero-marginals has infinite mass, conditions are given for the cyclically monotone zero-coupling to be unique and for a cyclically monotone transport map to exist.

Section 5 investigates the connections linking couplings between multivariate regularly varying probability distributions on the one hand and zero-couplings between their limit measures in the definition of regular variation on the other hand. A criterion is given for a cyclically monotone transport map to exist; it only concerns the support of the first zero-marginal. An extended form of regular variation of the cyclically monotone coupling is proven, together with the other assertions in the commutative diagram in Fig. 1. A short discussion in Section 6 concludes the paper.

The proofs of the main results are given at the end of each section. The appendix contains background material on M_0-convergence and convex analysis (Appendix A), together with auxiliary results on zero-couplings (Appendix B).

1.5. Notation

On $\mathbb{R}^k$, we always work with the Euclidean inner product $\langle\cdot,\cdot\rangle$ and the Euclidean norm $\|\cdot\|$. The interior, closure and boundary of a set A of a topological space are denoted by $\operatorname{int} A$, $\operatorname{cl} A$ and $\operatorname{bnd} A$, respectively. The open ball in $\mathbb{R}^k$ with center $x \in \mathbb{R}^k$ and radius $r > 0$ is denoted by $\mathbb{B}_{x,r} = \mathbb{B}(x, r) =$

$\mathbb{B}_k(x,r) = \{y \in \mathbb{R}^k \mid \|y-x\| < r\}$. The σ-field of Borel subsets on a topological space X is denoted by $\mathcal{B}(X)$. We frequently use the projections $\mathrm{proj}_1, \mathrm{proj}_2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ defined by $\mathrm{proj}_1(x,y) = x$ and $\mathrm{proj}_2(x,y) = y$.

The effective domain, or just domain in short, of a convex function $\psi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is the set $\mathrm{dom}\,\psi = \{x \in \mathbb{R}^d \mid \psi(x) < \infty\}$, while its subgradient is the set

$$\partial\psi = \left\{(x,v) \in \mathrm{dom}\,\psi \times \mathbb{R}^d \mid \forall y \in \mathbb{R}^d : \psi(y) \geqslant \psi(x) + \langle y-x, v\rangle\right\}. \tag{4}$$

For $x \in \mathbb{R}^d$, we write $\partial\psi(x) = \left\{v \in \mathbb{R}^d \mid (x,v) \in \partial\psi\right\}$; note that $\partial\psi(x) = \varnothing$ if $x \notin \mathrm{dom}\,\psi$. We refer to Section A.2 for more details and useful results on convex analysis.

For two Borel measures μ and ν on $\mathbb{R}^d$, let $\Pi(\mu,\nu)$ denote the set of coupling measures, i.e., the Borel measures γ on $\mathbb{R}^d \times \mathbb{R}^d$ such that $\gamma(B \times \mathbb{R}^d) = \mu(B)$ and $\gamma(\mathbb{R}^d \times B) = \nu(B)$ for all $B \in \mathcal{B}(\mathbb{R}^d)$. Furthermore, $\Pi_{\mathrm{cm}}(\mu,\nu)$ is the subset of such couplings γ whose support is cyclically monotone in the sense of Eq. (6) below. The set of Borel probability measures on $\mathbb{R}^d$ is denoted by $\mathcal{P}(\mathbb{R}^d)$.

Let the map $T : \mathrm{dom}\,T \subset \mathbb{R}^k \to \mathbb{R}^l$ be Borel measurable and let μ be a Borel measure on a $S \subset \mathbb{R}^k$, where both $\mathrm{dom}\,T$ and S are Borel sets. The push-forward measure $T_\#\mu$ of μ by T is the Borel measure on $\mathbb{R}^l$ defined by

$$\forall B \in \mathcal{B}(\mathbb{R}^l): \qquad T_\#\mu(B) = \mu\big(T^{-1}(B)\big) = \mu\left(\{x \in S \cap \mathrm{dom}\,T \mid T(x) \in B\}\right).$$

2. Zero-couplings between infinite measures

First we introduce the notion of zero-couplings in Section 2.1, and then discuss their stability under M_0-convergence of measures in Section 2.2. Finally, we discuss their existence and their representation in terms of push-forward measures plus a potential remainder term in Section 2.3.

2.1. Zero-couplings

A subset B of $\mathbb{R}^d$ is bounded away from $0 \in \mathbb{R}^d$ if $0 \notin \mathrm{cl}(B)$. Let $\mathcal{B}_0(\mathbb{R}^d)$ denote the collection of Borel subsets of $\mathbb{R}^d_{\setminus 0} = \mathbb{R}^d \setminus \{0\}$ that are bounded away from 0. Let $\mathrm{M}_0(\mathbb{R}^d)$ denote the set of nonnegative Borel measures μ on $\mathbb{R}^d_{\setminus 0}$ such that $\mu(B) < \infty$ for all $B \in \mathcal{B}_0(\mathbb{R}^d)$. For $\mu \in \mathrm{M}_0(\mathbb{R}^d)$, let $[\mu]$ denote the set of nonnegative Borel measures $\bar\mu$ on $\mathbb{R}^d$ whose restriction $\mathrm{res}\,\bar\mu$ to $\mathbb{R}^d_{\setminus 0}$ is equal to μ, i.e., $\bar\mu(B) = \mu(B)$ for any $B \in \mathcal{B}(\mathbb{R}^d_{\setminus 0})$. Further, let $\tilde\mu \in [\mu]$ be the unique member of $[\mu]$ such that $\tilde\mu(\{0\}) = 0$, i.e., $\tilde\mu(B) = \mu(B \setminus \{0\})$ for $B \in \mathcal{B}(\mathbb{R}^d)$. For Borel measures $\bar\mu$ on $\mathbb{R}^d$, we use the notation $[\bar\mu]$ in the same way.

It will be convenient to define the support of $\mu \in \mathrm{M}_0(\mathbb{R}^d)$ as the set of points $x \in \mathbb{R}^d$ such that $\tilde\mu(U) > 0$ for every open set $U \subset \mathbb{R}^d$ such that $x \in U$; that is, the support of μ is defined as the support (in the traditional sense) of $\tilde\mu$. This means that 0 can belong to the support of μ, even though μ is a measure on $\mathbb{R}^d_{\setminus 0}$ only. In particular, $0 \in \mathrm{spt}\,\mu$ if $\mu(U \setminus \{0\}) > 0$ for every open set $U \in \mathbb{R}^d$ containing 0. Note that $\mathrm{spt}\,\mu$ cannot be equal to $\{0\}$.

Definition 2.1 (Zero-marginals). *The left and right zero-marginals of a nonnegative Borel measure γ on $(\mathbb{R}^d \times \mathbb{R}^d) \setminus \{0\}$ are the nonnegative Borel measures $\varphi_1\gamma, \varphi_2\gamma$ on $\mathbb{R}^d_{\setminus 0}$, respectively, with*

$$\varphi_1\gamma(A) = \gamma(A \times \mathbb{R}^d) \quad \text{and} \quad \varphi_2\gamma(A) = \gamma(\mathbb{R}^d \times A) \quad \text{for } A \in \mathcal{B}(\mathbb{R}^d_{\setminus 0}).$$

If $A \subset \mathbb{R}^d_{\not\ni 0}$ is bounded away from $0 \in \mathbb{R}^d$, then $A \times \mathbb{R}^d$ and $\mathbb{R}^d \times A$ are bounded away from $0 \in \mathbb{R}^d \times \mathbb{R}^d$. It follows that $\gamma \in \mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$ implies $\varphi_1\gamma, \varphi_2\gamma \in \mathrm{M}_0(\mathbb{R}^d)$.

A subset T of $\mathbb{R}^d \times \mathbb{R}^d$ is *monotone* if

$$\forall (x_1, y_1), (x_2, y_2) \in T : \qquad \langle x_1 - x_2, y_1 - y_2 \rangle \geqslant 0. \tag{5}$$

In the following, we need a stronger notion: a set $T \subset \mathbb{R}^d \times \mathbb{R}^d$ is *cyclically monotone* if for every integer $n \geqslant 2$ and all $(x_1, y_1), \ldots, (x_n, y_n) \in T$ with $y_{n+1} := y_1$, we have

$$\sum_{i=1}^{n} \langle x_i, y_i \rangle \geqslant \sum_{i=1}^{n} \langle x_i, y_{i+1} \rangle. \tag{6}$$

The case $n = 2$ in Eq. (6) leads back to Eq. (5), so that a cyclically monotone set is also monotone.

Definition 2.2 (Zero-couplings). *Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$. The set $\Gamma_0(\mu, \nu)$ of zero-couplings between μ and ν consists of all nonnegative Borel measures γ on $(\mathbb{R}^d \times \mathbb{R}^d) \setminus \{0\}$ such that*

$$\varphi_1\gamma = \mu \quad \text{and} \quad \varphi_2\gamma = \nu.$$

Let $\Gamma_{0,\mathrm{m}}(\mu, \nu)$ and $\Gamma_{0,\mathrm{cm}}(\mu, \nu)$ denote the sets of zero-couplings γ between μ and ν such that $\operatorname{spt}\gamma$ *is monotone and cyclically monotone, respectively.*

In Definition 2.2, we implicitly use the property that $\Gamma_0(\mu, \nu) \subset \mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$, which is justified by Lemma 2.4. Clearly, $\Gamma_{0,\mathrm{cm}}(\mu, \nu) \subset \Gamma_{0,\mathrm{m}}(\mu, \nu) \subset \Gamma_0(\mu, \nu)$ for $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$.

Remark 2.3. Theorems 2.1 and 2.2 are similar to the ones of the usual marginals and couplings of measures. Indeed, for a measure π in the set $\mathrm{M}^+(\mathbb{R}^d \times \mathbb{R}^d)$ of nonnegative Borel measures on $\mathbb{R}^d$, the marginals $\phi_1\pi, \phi_2\pi$ are defined like the zero-marginals with the difference that the set A is allowed to contain the origin, while the collection of (cyclically monotone) couplings $\Pi_{(\mathrm{cm})}(p, q) \subset \mathrm{M}^+(\mathbb{R}^d \times \mathbb{R}^d)$ between $p, q \in \mathrm{M}^+(\mathbb{R}^d)$ is defined like the one of zero-couplings by substituting φ_1, φ_2 with ϕ_1, ϕ_2.

The following lemma ensures $\Gamma_0(\mu, \nu) \subset \mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$, so that for any zero-coupling γ, the expressions $\operatorname{spt}\gamma$, $\tilde{\gamma}$ and $[\gamma]$ are well-defined.

Lemma 2.4. *For all $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$, we have $\Gamma_0(\mu, \nu) \subset \mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$.*

Proof. To avoid ambiguities, we indicate the dimension of a ball by a subscript. Let $r > 0$. For $0 < r' \leqslant r/\sqrt{2}$, we have $\mathbb{B}_d(0, r') \times \mathbb{B}_d(0, r') \subset \mathbb{B}_{2d}(0, r)$ and thus

$$(\mathbb{R}^d \times \mathbb{R}^d) \setminus \mathbb{B}_{2d}(0, r) \subset \left(\left(\mathbb{R}^d \setminus \mathbb{B}_d(0, r') \right) \times \mathbb{R}^d \right) \cup \left(\mathbb{R}^d \times \left(\mathbb{R}^d \setminus \mathbb{B}_d(0, r') \right) \right).$$

For any $\gamma \in \Gamma_0(\mu, \nu)$, we have therefore

$$\gamma\left((\mathbb{R}^d \times \mathbb{R}^d) \setminus \mathbb{B}_{2d}(0, r) \right) \leqslant \mu(\mathbb{R}^d \setminus \mathbb{B}_d(0, r')) + \nu(\mathbb{R}^d \setminus \mathbb{B}_d(0, r')) < \infty. \qquad \square$$

We now define a special set of zero-couplings that will be of major interest in the remainder of this paper.

Definition 2.5 (Proper zero-couplings). *A zero-coupling $\gamma \in \Gamma_0(\mu,\nu)$ is called* proper *if*

$$\gamma\left(\{0\}\times\mathbb{R}^d_{\setminus 0}\right) = 0.$$

Let $\Gamma_0^{\mathrm{p}}(\mu,\nu)$ denote the set of proper zero-couplings and let $\Gamma_{0,\mathrm{m}}^{\mathrm{p}}(\mu,\nu) = \Gamma_0^{\mathrm{p}}(\mu,\nu)\cap\Gamma_{0,\mathrm{m}}(\mu,\nu)$ and $\Gamma_{0,\mathrm{cm}}^{\mathrm{p}}(\mu,\nu) = \Gamma_0^{\mathrm{p}}(\mu,\nu)\cap\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ denote the sets of proper zero-couplings with monotone and cyclically monotone support, respectively.

Theorem 2.5 is crucial because for applications involving transport maps between measures, only proper zero-couplings are relevant. Indeed, if $\gamma(\{0\}\times\mathbb{R}^d_{\setminus 0})$ is non-zero, the origin $\{0\}$ serves as an additional reservoir of mass when transporting μ onto ν. As shall become clear in Theorem 2.9, this precludes the construction of a map T that pushes μ forward to ν.

Note that 'being proper' is a property which concerns equivalence classes of zero-couplings: If the zero-couplings γ_1 and γ_2 belong to the same equivalence class $[\gamma]$, then γ_1 is proper if and only if γ_2 is proper. Additional facts about zero-couplings and their supports are collected in Appendix B.

2.2. Stability

To deal with transport problems between infinite measures, a common approach is to consider transport problems between finite approximations for which there are already results in the literature. We refer to Huesmann [20],Huesmann and Sturm [21] and Catalano et al. [11] where the authors are interested in vague convergence and weak* convergence respectively. This approach is all the more natural in our setting since we aim at studying transport problems between infinite measures in $\mathrm{M}_0(\mathbb{R}^d)$ appearing as limits of rescaled probability measures in the definition of regular variation (Section 5).

Let $C_0(\mathbb{R}^d)$ denote the set of bounded and real-valued functions f on $\mathbb{R}^d_{\setminus 0}$ that vanish on a neighbourhood of the origin, i.e., for which there exists $r>0$ such that $f(x)=0$ whenever $\|x\|\leqslant r$. The topology on $\mathrm{M}_0(\mathbb{R}^d)$ is the weakest one for which the evaluation functionals $\mu\mapsto\mu(f)=\int f\,\mathrm{d}\mu$ are continuous for all $f\in C_0(\mathbb{R}^d)$. The space $\mathrm{M}_0(\mathbb{R}^d)$ equipped with this topology is Polish [22, Theorem 2.3]. For a sequence $(\mu_n)_n$ in $\mathrm{M}_0(\mathbb{R}^d)$, the convergence $\mu_n \xrightarrow{0}\mu$ in $\mathrm{M}_0(\mathbb{R}^d)$ as $n\to\infty$ is equivalent to the convergence $\mu_n(f)\to\mu(f)$ as $n\to\infty$ for all $f\in C_0(\mathbb{R}^d)$. For convergence of measures in $\mathrm{M}_0(\mathbb{R}^d)$, a version of the Portmanteau theorem and of Prohorov's theorem are available in [22] and are quoted in Section A.1.

Consider two sequences $\mu_n,\nu_n\in\mathrm{M}_0(\mathbb{R}^d)$, $n\in\mathbb{N}$, converging to $\mu,\nu\in\mathrm{M}_0(\mathbb{R}^d)$, respectively, and a sequence of zero-couplings $\gamma_n\in\Gamma_0(\mu_n,\nu_n)$ and closed sets $T_n\subset\mathbb{R}^d\times\mathbb{R}^d$ satisfying $\operatorname{spt}\gamma_n\subset T_n$. The asymptotic behaviour of γ_n and T_n is described below in $\mathrm{M}_0(\mathbb{R}^d\times\mathbb{R}^d)$ and in the Fell topology, respectively. Theorem 2.6 is inspired by results from Segers [36, Theorem 1.1], treating weakly converging sequences of probability measures and their couplings.

We follow [26] for the description and properties of the Fell topology. For an open subset $\mathbb{E}$ of $\mathbb{R}^d$, let $\mathcal{F}(\mathbb{E})$ denote the set of all subsets of $\mathbb{E}$ which are closed in the trace topology on $\mathbb{E}$. For a subset A of $\mathbb{E}$, we consider the collections $\mathcal{F}_A=\{F\in\mathcal{F}(\mathbb{E})\mid F\cap A\neq\varnothing\}$ and $\mathcal{F}^A=\{F\in\mathcal{F}(\mathbb{E})\mid F\cap A=\varnothing\}$ of sets that respectively hit and miss A. The Fell hit-and-miss topology is defined by the sub-base consisting of $\mathcal{F}_G$ for all open subsets G of $\mathbb{E}$ and $\mathcal{F}^K$ for all compact subsets K of $\mathbb{E}$. Let $F,F_n\in\mathcal{F}(\mathbb{E})$, $n\in\mathbb{N}$. The sequence F_n $\mathcal{F}$-converges to F as $n\to\infty$ in the Fell topology—we write $F_n\xrightarrow{\mathcal{F}}F$—if for every open $G\subset\mathbb{E}$ such that $F\in\mathcal{F}_G$ we have $F_n\in\mathcal{F}_G$ for all large n, and for every compact $K\subset\mathbb{E}$ satisfying $F\in\mathcal{F}^K$ we have $F_n\in\mathcal{F}^K$ for all large n.

We write $\mathcal{F}_{\mathrm{cm}}$ for the collection of closed subsets of $\mathbb{R}^d\times\mathbb{R}^d$ satisfying the cyclical monotonicity property in Eq. (6) and $\mathcal{F}_{\mathrm{mcm}}$ for the collection of maximal cyclically monotone sets, which are the

cyclically monotone sets that are not strictly included in any other cyclically monotone set—obviously any maximal cyclically monotone set is closed. Likewise, we write $\mathcal{F}_{\mathrm{mm}}$ for the collection of maximal monotone sets, that is, monotone sets as in Eq. (5) that are not strictly included in any other monotone set.

Theorem 2.6 (Stability). *Let $\mu_n, \nu_n \in \mathrm{M}_0(\mathbb{R}^d)$ and $\gamma_n \in \Gamma_0(\mu_n, \nu_n)$, $n \in \mathbb{N}$, be such that $\mu_n \xrightarrow{0} \mu$ and $\nu_n \xrightarrow{0} \nu$. Let $T_n \in \mathcal{F}(\mathbb{R}^d \times \mathbb{R}^d)$ be a sequence of closed sets in $\mathbb{R}^d \times \mathbb{R}^d$ such that $\operatorname{spt}\gamma_n \subset T_n$ for each $n \in \mathbb{N}$. Then there exists an infinite subset N of $\mathbb{N}$ such that*

(a) *$\gamma_n \xrightarrow{0} \gamma$ as $n \to \infty$ in N for some $\gamma \in \Gamma_0(\mu, \nu)$,*
(b) *$T_n \xrightarrow{\mathcal{F}} T$ as $n \to \infty$ in N for some $T \in \mathcal{F}(\mathbb{R}^d \times \mathbb{R}^d)$,*
(c) *$\operatorname{spt}\gamma \subset T$.*

If we further assume that $\gamma_n \in \Gamma_{0,\mathrm{cm}}(\mu_n, \nu_n)$ and $T_n \in \mathcal{F}_{\mathrm{mcm}}$ for all $n \in \mathbb{N}$, then all limit points of γ_n belong to $\Gamma_{0,\mathrm{cm}}(\mu, \nu)$ while all limit points of T_n lie in $\mathcal{F}_{\mathrm{mcm}}$.

Anticipating the uniqueness of the cyclically monotone zero-coupling asserted in Theorem 4.1, we prove a stronger result in the cyclically monotone case. It will be of major interest in Section 5 to derive the convergences announced in the commutative diagram in Fig. 1. For $V \subset \mathbb{R}^d$ and $T \subset \mathbb{R}^d \times \mathbb{R}^d$, we write $T \llcorner V = T \cap (V \times \mathbb{R}^d)$, to be thought of as the restriction of T (seen as a multi-valued map) to V.

Corollary 2.7. *Besides the assumptions of Theorem 2.6, assume that $\gamma_n \in \Gamma_{0,\mathrm{cm}}(\mu_n, \nu_n)$ and $T_n \in \mathcal{F}_{\mathrm{mcm}}$ for every $n \in \mathbb{N}$ and that there exist $\gamma \in \mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$ and $T \in \mathcal{F}_{\mathrm{mcm}}$ such that $\Gamma_{0,\mathrm{cm}}(\mu, \nu) = \{\gamma\}$ and $\operatorname{spt}\gamma \subset T$. Then, writing $V = \operatorname{int}(\operatorname{spt}\mu)$, we have*

$$\gamma_n \xrightarrow{0} \gamma \qquad \text{and} \qquad T_n \llcorner V \xrightarrow{\mathcal{F}} T \llcorner V \ \text{ in } \mathcal{F}(V \times \mathbb{R}^d) \qquad \text{as } n \to \infty.$$

Proofs

Proof of Theorem 2.6. We deal with the assertions (a) and (b,c) separately.

Proof of (a) — If μ_n, ν_n M_0-converge to $\mu, \nu \in \mathrm{M}_0$ respectively, then the sets $K = \{\mu_n : n \geqslant 1\} \cup \{\mu\}$ and $L = \{\nu_n : n \geqslant 1\} \cup \{\nu\}$ are both compact. By Theorem B.3(c), the set $Z = \{(a, b, \zeta) \mid a \in K, b \in L, \zeta \in \Gamma_0(a, b)\}$ is then compact too. Since the triplets (μ_n, ν_n, γ_n) belong to Z, there is an infinite subset $\mathrm{N} \subset \mathbb{N}$ and a limit point $(\mu', \nu', \gamma) \in Z$ such that $(\mu_n, \nu_n, \gamma_n) \to (\mu', \nu', \gamma)$ along N, in $\mathrm{M}_0(\mathbb{R}^d) \times \mathrm{M}_0(\mathbb{R}^d) \times \mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$. In particular, $\gamma_n \xrightarrow{0} \gamma$ and, by definition of Z, also $\gamma \in \Gamma_0(\mu', \nu')$. Since, by assumption, μ_n and ν_n converge to μ and ν, respectively, we obtain $\mu' = \mu$ and $\nu' = \nu$. The assertion follows.

We turn to the additional assumption that $\gamma_n \in \Gamma_{0,\mathrm{cm}}(\mu_n, \nu_n)$. The set

$$\tilde{Z} = \big\{(a, b, \zeta) \mid a \in K, b \in L, \zeta \in \Gamma_{0,\mathrm{cm}}(a, b)\big\}$$

can be written as $\tilde{Z} = Z \cap \{(a, b, \zeta) \mid a \in \mathrm{M}_0(\mathbb{R}^d), b \in \mathrm{M}_0(\mathbb{R}^d), \zeta \in \Gamma_{0,\mathrm{cm}}(a, b)\}$. The first set in the intersection is compact while, from Theorem B.3(b), the second set is closed. Thus $\tilde{Z}$ is closed. Since $(\mu_n, \nu_n, \gamma_n) \in \tilde{Z}$ for all n, so does the limit point (μ, ν, γ), whence $\gamma \in \Gamma_{0,\mathrm{cm}}(\mu, \nu)$.

Proof of (b) and (c) — Since $\mathcal{F}(\mathbb{R}^d \times \mathbb{R}^d)$ is compact, there exist an infinite subset N' of N in the proof of (a) together with some T in $\mathcal{F}(\mathbb{R}^d \times \mathbb{R}^d)$ such that T_n $\mathcal{F}$-converges to T as $n \to \infty$ in N'. Thanks to Lemma 3.2 in [36], $\mathcal{F}_{\mathrm{mcm}}(\mathbb{R}^d \times \mathbb{R}^d) \cup \{\varnothing\}$ is a closed subset of $\mathcal{F}(\mathbb{R}^d \times \mathbb{R}^d)$. As a consequence, $T_n \in \mathcal{F}_{\mathrm{mcm}}(\mathbb{R}^d \times \mathbb{R}^d), n \geqslant 1$ implies $T \in \mathcal{F}_{\mathrm{mcm}}(\mathbb{R}^d \times \mathbb{R}^d) \cup \{\varnothing\}$. Since μ and ν are nonzero measures

and since $\gamma \in \Gamma_{0,\mathrm{cm}}(\mu,\nu)$, we have $\operatorname{spt}\gamma \neq \varnothing$. By Theorem B.2, we have $\operatorname{spt}\gamma \subset T$, so that T is non empty, and T must lie in $\mathcal{F}_{\mathrm{mcm}}(\mathbb{R}^d \times \mathbb{R}^d)$. □

Proof of Corollary 2.7. The assertion $\gamma_n \xrightarrow{0} \gamma$ is a direct consequence of Theorem 2.6: if every subsequence of $(\gamma_n)_n$ contains a convergent subsequence and if all limits points of such subsequences are equal to γ, then γ_n must converge to γ in $\mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$ along the whole sequence.

Since $\gamma_n \in \Gamma_{0,\mathrm{cm}}(\mu_n,\nu_n)$ for all $n \in \mathbb{N}$, Theorem 2.6 implies that any limit point of $(T_n)_n$ belongs to $\mathcal{F}_{\mathrm{mcm}}(\mathbb{R}^d \times \mathbb{R}^d)$ and thus to $\mathcal{F}_{\mathrm{mm}}(\mathbb{R}^d \times \mathbb{R}^d)$, because any maximal cyclically monotone set is also maximal monotone [32, Theorems 12.17 and 12.25]. Let $T_\infty \in \mathcal{F}_{\mathrm{mm}}$ be a limit point of $(T_n)_n$ as $n \to \infty$ for $n \in \mathrm{N}$, with N an infinite subset of $\mathbb{N}$. By point (c) of Theorem 2.6, we have $\operatorname{spt}\gamma \subset T_\infty$. For $T \in \mathcal{F}_{\mathrm{mcm}} \subset \mathcal{F}_{\mathrm{mm}}$ in the statement of the corollory, Theorem B.6 implies $T \llcorner V = T_\infty \llcorner V$. Since $V \subset \mathbb{R}^d$ is open, the restriction map $\cdot \llcorner V : \mathcal{F}(\mathbb{R}^d \times \mathbb{R}^d) \to \mathcal{F}(V \times \mathbb{R}^d)$ is continuous [36, Proposition D.12]. It follows that along the subsequence N, the restrictions $T_n \llcorner V$ converge in $\mathcal{F}(V \times \mathbb{R}^d)$ to $T_\infty \llcorner V = T \llcorner V$. Since this is true for an arbitrary converging subsequence and since any subsequence contains a converging sequence (Theorem 2.6), we conclude that $T_n \llcorner V \to T \llcorner V$ in $\mathcal{F}(V \times \mathbb{R}^d)$ along the whole sequence, as required. □

2.3. Existence and representation

When a moment assumption is made on both zero-marginals μ and ν, the existence of a cyclically monotone zero-coupling between μ and ν is a direct application of Theorems A.5 and Theorem A.13 in Guillen et al. [17]. Relying on Theorem 2.6, which provides stability results and the existence of limit points for the sequence of couplings, we derive the same conclusion without making any moment assumption.

Theorem 2.8 (Existence). *For all $\mu,\nu \in \mathrm{M}_0(\mathbb{R}^d)$ we have $\Gamma_{0,\mathrm{cm}}(\mu,\nu) \neq \varnothing$.*

Recall the notion of the subdifferential of a convex function in Eq. (4). Relying on Rockafellar's Theorem A.4, which establishes a connection between cyclical monotonicity and subdifferentials of closed convex functions, we derive a representation result for cyclically monotone zero-couplings similar to Proposition 10 in McCann [25]. Recall Theorem 2.5 of proper zero-couplings and recall that for $\mu \in \mathrm{M}_0(\mathbb{R}^d)$, the measure $\tilde{\mu}$ is the extension of μ to $\mathbb{R}^d$ determined by $\tilde{\mu}(\{0\}) = 0$, while $\operatorname{res} m$ denotes the restriction to $\mathbb{R}^d_{\setminus 0}$ of a measure m defined on $\mathbb{R}^d$. Let Id be the identity map on $\mathbb{R}^d$, and for a map f defined on some subset of $\mathbb{R}^d$, put $(\mathrm{Id} \times f)(x) = (x, f(x))$ for $x \in \operatorname{dom} f$. Here and throughout, $\operatorname{dom} f$ denotes the subset of $\mathbb{R}^d$ where f is defined (except for convex functions ψ from $\mathbb{R}^d$ into $\mathbb{R} \cup \{\infty\}$, where $\operatorname{dom}\psi = \{x \in \mathbb{R}^d \mid \psi(x) < \infty\}$). For a measure m on $\mathbb{R}^d$ and a measurable map f from some subset of $\mathbb{R}^d$ to $\mathbb{R}^d$, the measure $(\mathrm{Id} \times f)_\# m$ on $\mathbb{R}^d \times \mathbb{R}^d$ is determined for $A, B \in \mathcal{B}(\mathbb{R}^d)$ by

$$(\mathrm{Id} \times f)_\# m(A \times B) = m\left(A \cap f^{-1}(B)\right) = m\left(\{x \in A \cap \operatorname{dom} f : f(x) \in B\}\right).$$

If m is a measure defined on some subset M of $\mathbb{R}^d$ and if f is a measurable function on some subset $\operatorname{dom} f$ of $\mathbb{R}^d$, we use $f_\# m$ to mean the measure on $\mathbb{R}^d$ defined as

$$f_\# m(A) = m(M \cap f^{-1}(A)) = m(\{x \in M \cap \operatorname{dom} f : f(x) \in A\}),$$

for Borel sets A. Also, we extend common terminology and call the measure $f_\# M$ the push-forward measure of m by the mapping f, and if $\lambda = f_\# m$ we shall say that f pushes m forward to λ. For instance,

the measure $(\nabla\psi)_{\#}\mu$ in Eq. (7) below is defined for $A \in \mathcal{B}(\mathbb{R}^d)$ by

$$(\nabla\psi)_{\#}\mu(A) = \mu\left((\nabla\psi)^{-1}(A) \setminus \{0\}\right) \text{ with } (\nabla\psi)^{-1}(A) = \{x \in \operatorname{dom}\nabla\psi \mid \nabla\psi(x) \in A\}\,.$$

By Rockafellar's Theorem A.4, any cyclically monotone set such as $\operatorname{spt}\gamma$ in the statement below is contained in the subdifferential $\partial\psi$ of some closed convex function ψ.

Theorem 2.9 (Representation). *Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ and $\gamma \in \Gamma_{0,\mathrm{cm}}(\mu,\nu)$, and assume that μ vanishes on sets of Hausdorff dimension at most $d-1$. Then for any closed convex function $\psi : \mathbb{R}^d \to \mathbb{R}\cup\{\infty\}$ satisfying $\operatorname{spt}\gamma \subset \partial\psi$, we have $\mu(\mathbb{R}^d_{\diagdown 0} \setminus \operatorname{dom}\nabla\psi) = 0$. The measure ν can be written as the push-forward of μ by $\nabla\psi$, up to a residual term: for every Borel set $B \subset \mathbb{R}^d_{\diagdown 0}$,*

$$\nu(B) = (\nabla\psi)_{\#}\mu(B) + \gamma(\{0\}\times B).$$

In addition, γ is uniquely determined by $\nabla\psi, \mu, \nu$ through the identities

$$\gamma(A\times B) = (\mathrm{Id}\times\nabla\psi)_{\#}\mu(A\times B) = \mu\left(A \cap (\nabla\psi)^{-1}(B)\right), \qquad A \in \mathcal{B}(\mathbb{R}^d_{\diagdown 0}),\ B \in \mathcal{B}(\mathbb{R}^d),$$

$$\gamma(\{0\}\times B) = \nu(B) - (\nabla\psi)_{\#}\mu(B), \qquad B \in \mathcal{B}(\mathbb{R}^d),\ 0 \notin \mathrm{cl}(B).$$

Finally, if γ is proper, i.e., $\gamma \in \Gamma^{\mathrm{p}}_{0,\mathrm{cm}}(\mu,\nu)$, then we have the proper representation

$$\gamma = \operatorname{res}\left((\mathrm{Id}\times\nabla\psi)_{\#}\mu\right) \quad \textit{and} \quad \nu = \operatorname{res}\left((\nabla\psi)_{\#}\mu\right). \tag{7}$$

Rephrasing the last part of the theorem, if a proper cyclically monotone zero-coupling γ of $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ exists, then there is a closed convex function ψ whose gradient $\nabla\psi$ pushes μ forward to a measure $\bar{\nu}$ on $\mathbb{R}^d$ of which the restriction to $\mathbb{R}^d_{\diagdown 0}$ is equal to ν.

Contrary to McCann's result for cyclically monotone couplings between finite measures with equal mass in Theorem 1.1, there is in Theorem 2.9 an extra term $\gamma(\{0\}\times B)$. It can be interpreted as assigning mass to $\{0\}\times\mathbb{R}^d_{\diagdown 0}$ when we need to "extract" mass from the origin for the cyclically monotone zero-coupling to exist. Example 1.2 presents a case where this term is non-zero and the zero-coupling is not proper.

Proofs

Proof of Theorem 2.8. If μ and ν are both the zero measure on $\mathbb{R}^d_{\diagdown 0}$, then the zero measure on $(\mathbb{R}^d\times\mathbb{R}^d)\setminus\{0\}$ is a zero-coupling of μ and ν. So assume at least one of μ and ν is not the zero measure.

We construct two sequences $\mu_n, \nu_n \in \mathrm{M}_0(\mathbb{R}^d)$, $n \in \mathbb{N}$, of finite and equal mass converging to μ, ν in $\mathrm{M}_0(\mathbb{R}^d)$, respectively, together with a sequence $\tilde{\gamma}_n \in \Pi_{cm}(\tilde{\mu}_n, \tilde{\nu}_n)$, $n \in \mathbb{N}$, of cyclically monotone couplings between $\tilde{\mu}_n$ and $\tilde{\nu}_n$ whose restrictions $\gamma_n = \operatorname{res}\tilde{\gamma}_n$ to $(\mathbb{R}^d\times\mathbb{R}^d)\setminus\{0\}$ are zero-couplings between $\mu_n = \operatorname{res}\tilde{\mu}_n$ and $\nu_n = \operatorname{res}\tilde{\nu}_n$; recall that $\tilde{\mu}_n$ is the extension of μ_n from $\mathbb{R}^d_{\diagdown 0}$ to $\mathbb{R}^d$ by $\tilde{\mu}_n(\{0\}) = 0$, and similarly for the other measures. The result then follows by Theorem 2.6 upon taking the limit γ of a converging subsequence of γ_n.

For $n \in \mathbb{N}$, let B_n be the closed ball in $\mathbb{R}^d$ with center 0 and radius $1/n$. Define $\tilde{\mu}_n = \tilde{\mu}\left(\cdot \cap B_n^c\right)$ and $\tilde{\nu}_n = \tilde{\nu}\left(\cdot \cap B_n^c\right)$. These are measures on $\mathbb{R}^d$ that coincide with μ and ν, respectively, on the complement of B_n. We have $\mu_n = \operatorname{res}\tilde{\mu}_n \xrightarrow{0} \mu$ as $n \to \infty$, since for every $f \in C_0(\mathbb{R}^d)$, we have $\tilde{\mu}_n(f) = \mu(f)$ as soon as $n \in \mathbb{N}$ is sufficiently large such that f vanishes on B_n. Likewise, we have $\nu_n = \operatorname{res}\tilde{\nu}_n \xrightarrow{0} \nu$.

It is clear that both $\tilde{\mu}_n$ and $\tilde{\nu}_n$ have finite total mass since they originate from measures lying in $\mathrm{M}_0(\mathbb{R}^d)$, restricted to a set bounded away from the origin. If $\tilde{\mu}_n(\mathbb{R}^d)$ and $\tilde{\nu}_n(\mathbb{R}^d)$ differ, we modify the one with the smallest total mass. We write the reasoning for the case $\tilde{\mu}_n(\mathbb{R}^d) < \tilde{\nu}_n(\mathbb{R}^d)$; the same argument can be used in the other case. Let $\delta_n = \tilde{\nu}_n(\mathbb{R}^d) - \tilde{\mu}_n(\mathbb{R}^d) \in (0, \infty)$ and write $\tilde{\kappa}_n$ for the Lebesgue-uniform probability distribution on B_n. Replace $\tilde{\mu}_n$ by $\tilde{\mu}_n + \delta_n \cdot \tilde{\kappa}_n$. The same argument as above then still gives $\operatorname{res} \tilde{\mu}_n \xrightarrow{0} \mu$ as $n \to \infty$, but this time we have $a_n = \tilde{\mu}_n(\mathbb{R}^d) = \tilde{\nu}_n(\mathbb{R}^d)$. Since ν is not the zero measure, we have $a_n > 0$ for all sufficiently large n.

We can now apply Theorem 6 in [25] to the probability measures $\hat{\mu}_n = a_n^{-1}\tilde{\mu}_n$ and $\hat{\nu}_n = a_n^{-1}\tilde{\nu}_n$. It yields the existence of some $\hat{\pi}_n \in \Pi(\hat{\mu}_n, \hat{\nu}_n)$ with cyclically monotone support. The measure $\tilde{\gamma}_n = a_n\hat{\pi}_n$ then belongs to $\Pi(\tilde{\mu}_n, \tilde{\nu}_n)$ and has the same cyclically monotone support as $\hat{\pi}_n$. By Rockafellar's Theorem A.4, there exists a closed convex function ψ_n on $\mathbb{R}^d$ such that the set $\operatorname{spt}\gamma_n = \operatorname{spt}\tilde{\gamma}_n$ is contained in $T_n = \partial\psi_n \in \mathcal{F}_{\mathrm{mcm}}(\mathbb{R}^d \times \mathbb{R}^d)$. One may notice that $\tilde{\gamma}_n$ is a true coupling between $\tilde{\mu}_n$ and $\tilde{\nu}_n$ seen as Borel measures on $\mathbb{R}^d$ putting no mass on $\{0\}$; indeed, by construction, $\tilde{\gamma}_n$ puts no mass on $(\{0\} \times \mathbb{R}^d) \cup (\mathbb{R}^d \times \{0\})$. Its restriction $\gamma_n = \operatorname{res}\tilde{\gamma}_n$ to $(\mathbb{R}^d \times \mathbb{R}^d) \setminus \{0\}$ lies in $\Gamma_{0,\mathrm{cm}}(\mu_n, \nu_n)$ and we have $\operatorname{spt}\gamma_n = \operatorname{spt}\tilde{\gamma}_n$.

By Corollary 2.6, there exist an infinite subset N of $\mathbb{N}$ and a measure γ in $\Gamma_{0,\mathrm{cm}}(\mu, \nu)$ such that γ_n M_0-converges to γ as $n \to \infty$ in N. □

Proof of Theorem 2.9. The proof is largely inspired by the one of Proposition 10 in [25]. We split $(\mathbb{R}^d \times \mathbb{R}^d) \setminus \{0\}$ into $\mathbb{R}^d_{\setminus 0} \times \mathbb{R}^d$ and $\{0\} \times \mathbb{R}^d_{\setminus 0}$.

Restriction of γ to $\mathbb{R}^d_{\setminus 0} \times \mathbb{R}^d$. — We first consider γ on subsets of $\mathbb{R}^d_{\setminus 0} \times \mathbb{R}^d$. Let

$$\begin{aligned} S &= \left\{(x, \nabla\psi(x)) : x \in \operatorname{dom}\nabla\psi \cap \mathbb{R}^d_{\setminus 0}\right\} \\ &= \left((\operatorname{dom}\nabla\psi \cap \mathbb{R}^d_{\setminus 0}) \times \mathbb{R}^d\right) \cap \partial\psi. \end{aligned}$$

Following the argument from the proof of Proposition 10 in [25], the measurability of $\nabla\psi$ is manifest since it coincides with the pointwise limit of a sequence of continuous approximants

$$\langle \nabla\psi(x), z\rangle = \lim_{n\to\infty} n\big(\psi(x + z/n) - \psi(x)\big),$$

so that $\operatorname{dom}\nabla\psi$ belongs to $\mathcal{B}(\mathbb{R}^d_{\setminus 0})$. Since $\partial\psi$ is closed and contains $\operatorname{spt}\gamma$, the set S lies in $\mathcal{B}((\mathbb{R}^d \times \mathbb{R}^d) \setminus \{0\})$ and contains the set

$$\tilde{S} = \left(\left(\operatorname{dom}\nabla\psi \cap \mathbb{R}^d_{\setminus 0}\right) \times \mathbb{R}^d\right) \cap \operatorname{spt}\gamma.$$

The latter set satisfies

$$\gamma\big(\tilde{S}^c \cap \left(\mathbb{R}^d_{\setminus 0} \times \mathbb{R}^d\right)\big) = \tilde{\mu}\left((\operatorname{dom}\nabla\psi)^c\right) = 0. \tag{8}$$

Indeed, by Theorem B.4 and Eq. (47), $\operatorname{spt}\mu \subset \operatorname{cl}(\operatorname{dom}\partial\psi) = \operatorname{cl}(\operatorname{dom}\psi)$. The latter closure is convex and its boundary has therefore Hausdorff dimension at most $d - 1$. Moreover, a convex function is differentiable in the interior of its domain except perhaps at a set of Hausdorff dimension at most $d - 1$, see Anderson and Klee [5, (F) and (3.2)]. It follows that $\tilde{\mu}((\operatorname{dom}\nabla\psi)^c) = \tilde{\mu}(\operatorname{cl}(\operatorname{dom}\psi) \setminus \operatorname{dom}\nabla\psi) = 0$, confirming Eq. (8).

Thus, for every $A \in \mathcal{B}(\mathbb{R}^d_{\not 0})$ and $B \in \mathcal{B}(\mathbb{R}^d)$ we have

$$\begin{aligned}\gamma(A \times B) = \gamma((A \times B) \cap S) &= \gamma\left(\left(A \cap (\nabla\psi)^{-1}(B)\right) \times \mathbb{R}^d\right) \\ &= \mu\left(A \cap (\nabla\psi)^{-1}(B)\right) = (\mathrm{Id} \times \nabla\psi)_{\#}\mu(A \times B).\end{aligned}$$

By the monotone class theorem, this identity uniquely determines the restriction of γ to $\mathbb{R}^d_{\not 0} \times \mathbb{R}^d$ as being equal to the restriction of $(\mathrm{Id} \times \nabla\psi)_{\#}\mu$ to $\mathbb{R}^d_{\not 0} \times \mathbb{R}^d$.

Restriction of γ to $\{0\} \times \mathbb{R}^d_{\not 0}$. — We now deal with subsets of $\{0\} \times \mathbb{R}^d_{\not 0}$. Let $B \in \mathcal{B}(\mathbb{R}^d_{\not 0})$. Since ν is the right zero-margin of γ, we can write successively

$$\begin{aligned}\nu(B) &= \gamma(\mathbb{R}^d \times B) \\ &= \gamma(\mathbb{R}^d_{\not 0} \times B) + \gamma(\{0\} \times B) \\ &= (\mathrm{Id} \times \nabla\psi)_{\#}\mu(\mathbb{R}^d_{\not 0} \times B) + \gamma(\{0\} \times B),\end{aligned}$$

where we used the representation of γ on $\mathbb{R}^d_{\not 0} \times \mathbb{R}^d$ in the last equality. This proves the first assertion of the statement.

Assume now that $0 \notin \operatorname{cl} B$. Since B and $\mathbb{R}^d_{\not 0} \times B$ are bounded away from the origin in $\mathbb{R}^d$ and $\mathbb{R}^d \times \mathbb{R}^d$, respectively, both $\nu(B)$ and $\gamma(\mathbb{R}^d_{\not 0} \times B)$ are finite. Hence we have

$$\gamma(\{0\} \times B) = \nu(B) - (\mathrm{Id} \times \nabla\psi)_{\#}\mu(\mathbb{R}^d_{\not 0} \times B). \tag{9}$$

This uniquely determines γ on $\{0\} \times \mathbb{R}^d_{\not 0}$, since for any $A \subset \mathcal{B}(\mathbb{R}^d_{\not 0})$ we can write $\gamma(\{0\} \times A) = \lim_{n\to\infty} \gamma(\{0\} \times (A \setminus \mathbb{B}_{0,1/n}))$.

Definition of γ on $(\mathbb{R}^d \times \mathbb{R}^d) \setminus \{0\}$. — Any Borel set E of $(\mathbb{R}^d \times \mathbb{R}^d) \setminus \{0\}$ can be partitioned into two pieces, $E \cap (\mathbb{R}^d_{\not 0} \times \mathbb{R}^d)$ and $E \cap (\{0\} \times \mathbb{R}^d_{\not 0})$. The mass that γ assigns to each of these pieces is uniquely determined by the representations in the two previous paragraphs.

Proper representation. — Assume $\gamma \in \Gamma^{\mathrm{p}}_{0,\mathrm{cm}}(\mu, \nu)$, so $\gamma(\{0\} \times \mathbb{R}^d_{\not 0}) = 0$, i.e., the restriction of γ to $\{0\} \times \mathbb{R}^d_{\not 0}$ is the zero measure. The same is true for $(\mathrm{Id} \times \nabla\psi)_{\#}\mu$, since μ is only defined on $\mathbb{R}^d_{\not 0}$ and thus, by definition,

$$(\mathrm{Id} \times \nabla\psi)_{\#}\mu\left(\{0\} \times \mathbb{R}^d\right) = \mu\left(\mathbb{R}^d_{\not 0} \cap (\{0\} \times \operatorname{dom}\psi)\right) = \mu(\varnothing) = 0.$$

Moreover, we already know that the measures γ and $(\mathrm{Id} \times \nabla\psi)_{\#}\mu$ coincide on $\mathbb{R}^d_{\not 0} \times \mathbb{R}^d$. Hence, both measures coincide on the union of $\{0\} \times \mathbb{R}^d_{\not 0}$ and $\mathbb{R}^d_{\not 0} \times \mathbb{R}^d$, which is $(\mathbb{R}^d \times \mathbb{R}^d) \setminus \{0\}$, and therefore $\gamma = \operatorname{res}((\mathrm{Id} \times \nabla\psi)_{\#}\mu)$.

We use the latter identity to find a representation of ν, the right zero-margin of γ. For any $B \in \mathcal{B}(\mathbb{R}^d_{\not 0})$, we have

$$\nu(B) = \gamma(\mathbb{R}^d \times B) = (\mathrm{Id} \times \nabla\psi)_{\#}\mu(B) = \mu\left(\mathbb{R}^d_{\not 0} \cap (\nabla\psi)^{-1}(B)\right) = (\nabla\psi)_{\#}\mu(B).$$

It finally follows that $\nu = \operatorname{res}((\nabla\psi)_{\#}\mu)$, as announced. □

3. Proper zero-couplings and representations

The interest in the proper zero-coupling and the proper representation in Eq. (7) is that we can write the "target measure" $\nu \in \mathrm{M}_0(\mathbb{R}^d)$ as a push-forward of the "reference measure" $\mu \in \mathrm{M}_0(\mathbb{R}^d)$, i.e., $\nu = \mathrm{res}((\nabla\psi)_{\#}\mu)$, or, more explicitly,

$$\forall B \in \mathcal{B}(\mathbb{R}^d_{\setminus 0}) : \qquad \nu(B) = (\nabla\psi)_{\#}\mu(B) = \mu\left((\nabla\psi)^{-1}(B) \setminus \{0\}\right),$$

see also the comment after Theorem 2.9. In such a case, $\nabla\psi$ is indeed a transport map pushing forward the reference μ to the target ν. This opens a way towards optimal-transport based statistical methodology without moment conditions as reviewed in [18] but then for Lévy measures and for certain infinite measures arising in multivariate regular variation, as developed in Section 5 below.

For finite measures μ and ν, a sufficient criterion for the existence of such a push-forward representation is that μ is sufficiently smooth and has at least as much mass as ν. This follows from the Brenier–McCann Theorem [9,10,25].

Proposition 3.1. *For $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ with μ vanishing on sets of Hausdorff dimension at most $d-1$, assume that $\nu(\mathbb{R}^d_{\setminus 0}) \leqslant \mu(\mathbb{R}^d_{\setminus 0}) < \infty$. Then there exists a closed convex function $\psi : \mathbb{R}^d \to \mathbb{R} \cup \{0\}$ satisfying $\nu = \mathrm{res}((\nabla\psi)_{\#}\mu)$. Moreover, $\mu(\mathbb{R}^d_{\setminus 0} \setminus \operatorname{dom}\nabla\psi) = 0$ and the measure $\gamma = \mathrm{res}((\mathrm{Id}\times\nabla\psi)_{\#}\mu)$ belongs to $\Gamma^{\mathrm{p}}_{0,\mathrm{cm}}(\mu,\nu)$.*

Remark 3.2. The condition $\nu(\mathbb{R}^d_{\setminus 0}) \leqslant \mu(\mathbb{R}^d_{\setminus 0})$ in Theorem 3.1 is actually necessary for the desired ψ to exist, even when no assumption is made on μ. Indeed, if $\nu = \mathrm{res}((\nabla\psi)_{\#}\mu)$, then $\nu(\mathbb{R}^d_{\setminus 0}) = (\nabla\psi)_{\#}\mu(\mathbb{R}^d_{\setminus 0}) = \mu((\nabla\psi)^{-1}(\mathbb{R}^d_{\setminus 0}) \setminus \{0\}) \leqslant \mu(\mathbb{R}^d_{\setminus 0})$.

In case μ and ν have infinite mass, the situation is more complicated. The problem is that the origin acts as a kind of barrier across which a (cyclically) monotone transport plan cannot move an infinite amount of mass, if the measures are to assign finite mass to sets bounded away from the origin. Example 1.2 illustrates this situation in one dimension, when the measures μ and ν put all their mass on opposite sides of the origin.

Recall the definition of a (cyclically) monotone subset of $\mathbb{R}^d \times \mathbb{R}^d$ in Eqs. (5) and (6) and the collections of proper zero-couplings $\Gamma^{\mathrm{p}}_{0,\mathrm{m}}(\mu,\nu)$ and $\Gamma^{\mathrm{p}}_{0,\mathrm{cm}}(\mu,\nu)$ in Theorem 2.5 with (cyclically) monotone support. By the following lemma, the existence of a proper zero-coupling with monotone support between non-finite measures μ and ν implies that (almost) any point in the support of ν can be "seen" from a point in the support of μ.

Lemma 3.3 (Necessary condition for proper, monotone zero-couplings). *Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ and suppose at least one of them has infinite mass. If there exists $\gamma \in \Gamma^{\mathrm{p}}_{0,\mathrm{m}}(\mu,\nu)$, then the following two properties hold:*

- *for ν-almost all $y \in \operatorname{spt}\nu$, there exists $x \in \operatorname{spt}\mu \setminus \{0\}$ such that $\langle x, y\rangle \geqslant 0$;*
- *for all $y \in \operatorname{spt}\nu$, we have $\sup_{x \in \operatorname{spt}\mu\setminus\{0\}} \langle x/\|x\|, y\rangle \geqslant 0$.*

In the rest of this section we will make the stronger assumption

$$\forall y \in \operatorname{spt}\nu \setminus \{0\} : \exists x \in \operatorname{spt}\mu : \qquad \langle x, y\rangle > 0. \tag{10}$$

We will seek additional conditions under which a zero-coupling with monotone support is proper. A first such condition is that the support of the zero-coupling is in a certain sense homogeneous. This situation

will naturally arise for zero-couplings between limit measures in multivariate regular variation, as in Theorem 5.1 below.

Proposition 3.4. *Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ and assume Eq. (10). Then $\gamma \in \Gamma_{0,\mathrm{m}}(\mu,\nu)$ is proper as soon as its support is homogeneous in the sense that*

$$\exists \alpha, \beta > 0 : \forall \lambda > 0 : \ (x,y) \in \operatorname{spt}\gamma \implies (\lambda^\alpha x, \lambda^\beta y) \in \operatorname{spt}\gamma. \tag{11}$$

The homogeneity assumption (11) made on the support of γ also implies an assumption on the supports of μ and ν. Moreover, this criterion does not allow one to prove directly that any (cyclically) monotone zero-coupling is proper, because of the additional assumption on its support. From now on we will only make assumptions on the zero-marginals μ, ν to derive the desired result.

According to the next result, a zero-coupling with monotone support is proper as soon as μ has sufficient mass in the directions that matter for ν. For $b \in \mathbb{R}^d_{\not\sim 0}$ and $\varepsilon > 0$, consider the cone

$$H_+(b,\varepsilon) = \left\{ x \in \mathbb{R}^d_{\not\sim 0} \mid \langle x/\|x\|, b\rangle > \varepsilon \right\}. \tag{12}$$

Proposition 3.5. *Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ and assume*

$$\forall y \in \operatorname{spt}\nu \setminus \{0\}, \exists \varepsilon > 0 : \qquad \mu\left(H_+(y,\varepsilon)\right) = \infty. \tag{13}$$

Then every $\gamma \in \Gamma_{0,\mathrm{m}}(\mu,\nu)$ is proper, i.e., $\Gamma_{0,\mathrm{m}}(\mu,\nu) = \Gamma^{\mathrm{p}}_{0,\mathrm{m}}(\mu,\nu)$, and therefore also $\Gamma_{0,\mathrm{cm}}(\mu,\nu) = \Gamma^{\mathrm{p}}_{0,\mathrm{cm}}(\mu,\nu)$.

Note that the assumption in Eq. (13) is stronger than the one in Eq. (10). Condition (13) is trivially satisfied if μ has infinite mass and is spherically symmetric.

We define the *positive hull* of a set $C \subset \mathbb{R}^d$ as the smallest cone containing it, i.e.

$$\operatorname{pos}(C) = \{\lambda x \mid x \in C, \lambda \geqslant 0\}.$$

Corollary 3.6. *Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ and assume Eq. (10). Let $\gamma \in \Gamma_{0,\mathrm{m}}(\mu,\nu)$. Then γ is proper as soon as at least one of the two following properties holds:*

(i) $\forall x \in \operatorname{spt}\mu, \forall \delta > 0 : \mu\left(\operatorname{pos}(\mathbb{B}(x,\delta)) \setminus \{0\}\right) = \infty$;
(ii) $\exists \alpha > 0, \forall \lambda > 0 : \mu(\lambda\,\cdot) = \lambda^{-\alpha}\mu(\,\cdot\,)$.

We conclude this section by a simple sufficient criterion in the one-dimensional case.

Proposition 3.7. *Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R})$, and assume*

$$\mu(\mathbb{R}_{>0}) \geqslant \nu(\mathbb{R}_{>0}) \quad \text{and} \quad \mu(\mathbb{R}_{<0}) \geqslant \nu(\mathbb{R}_{<0}).$$

If $\gamma \in \Gamma_0(\mu,\nu)$ has monotone support, then γ is proper. Conversely, if at least one of μ or ν has infinite mass, the existence of $\gamma \in \Gamma^{\mathrm{p}}_{0,\mathrm{m}}(\mu,\nu)$ implies that the above inequalities are both satisfied.

It is unclear if Theorem 3.7 admits an extension to the d-dimensional case in terms of masses assigned to half-spaces of the form $\left\{x \in \mathbb{R}^d \mid \langle x, b\rangle > 0\right\}$.

Proofs

Proof of Theorem 3.1. If $\mu(\mathbb{R}^d_{\not\sim 0}) = 0$, both μ and ν are the null measure and there is nothing to prove, so assume $0 < \mu(\mathbb{R}^d_{\not\sim 0}) < \infty$. Consider the extensions $\tilde{\mu} \in [\mu], \bar{\nu} \in [\nu]$ of μ, ν to $\mathbb{R}^d$ given by $\tilde{\mu}(\{0\}) = 0$, $\bar{\nu}(\{0\}) = \mu(\mathbb{R}^d_{\not\sim 0}) - \nu(\mathbb{R}^d_{\not\sim 0})$, respectively. Since $\nu(\mathbb{R}^d_{\not\sim 0}) \leqslant \mu(\mathbb{R}^d_{\not\sim 0})$ and since both are finite, $\bar{\nu}(\{0\})$ is nonnegative and finite, and $\tilde{\mu}(\mathbb{R}^d) = \bar{\nu}(\mathbb{R}^d)$. By assumption, $\tilde{\mu}$ vanishes on sets of Hausdorff dimension at most $d-1$, allowing one to apply McCann [25, Proof of existence, page 317] to the probability measures $m^{-1}\tilde{\mu}$ and $m^{-1}\bar{\nu}$, where $m = \tilde{\mu}(\mathbb{R}^d) = \bar{\nu}(\mathbb{R}^d)$. This yields the existence of a closed convex function $\psi : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ satisfying $\bar{\nu} = (\nabla\psi)_{\#}\tilde{\mu} = (\nabla\psi)_{\#}\mu$. Apply the operator res (restricting to $\mathbb{R}^d_{\not\sim 0}$) on both sides of the equation to find that $\nu = \operatorname{res}((\nabla\psi)_{\#}\mu)$.

The measure μ puts no mass on the complement of $\operatorname{dom}\nabla\psi$ since

$$\begin{aligned}\mu(\mathbb{R}^d_{\not\sim 0}) = \tilde{\mu}(\mathbb{R}^d) &= \bar{\nu}(\mathbb{R}^d)\\ &= (\nabla\psi)_{\#}\mu(\mathbb{R}^d) = \mu\left(\mathbb{R}^d_{\not\sim 0} \cap (\nabla\psi)^{-1}(\mathbb{R}^d)\right) = \mu\left(\mathbb{R}^d_{\not\sim 0} \cap \operatorname{dom}\nabla\psi\right),\end{aligned}$$

and both sides of the equation are finite.

Finally, we show that the measure $\gamma = \operatorname{res}\big((\mathrm{Id}\times\nabla\psi)_{\#}\,\mu\big)$ is a proper cyclically monotone zero-coupling of μ and ν. First, γ is proper since, by definition,

$$\begin{aligned}\gamma\left(\{0\}\times\mathbb{R}^d_{\not\sim 0}\right) &= (\mathrm{Id}\times\nabla\psi)_{\#}\,\mu\left(\{0\}\times\mathbb{R}^d_{\not\sim 0}\right)\\ &= \mu\left(\mathbb{R}^d_{\not\sim 0} \cap \left(\{0\} \cap (\nabla\psi)^{-1}(\mathbb{R}^d_{\not\sim 0})\right)\right) = \mu(\varnothing) = 0.\end{aligned}$$

Second, for $A \in \mathcal{B}(\mathbb{R}^d_{\not\sim 0})$ and $B \in \mathcal{B}(\mathbb{R}^d)$, we have

$$\gamma(A\times B) = (\mathrm{Id}\times\nabla\psi)_{\#}\mu(A\times B) = \mu\left(A \cap (\nabla\psi)^{-1}(B)\right).$$

On the one hand, putting $B = \mathbb{R}^d$, we find

$$\gamma(A\times\mathbb{R}^d) = \mu\left(A \cap \operatorname{dom}\nabla\psi\right) = \mu(A),$$

so $\varphi_1\gamma = \mu$. On the other hand, putting $A = \mathbb{R}^d_{\not\sim 0}$, we find, if $B \subset \mathbb{R}^d_{\not\sim 0}$,

$$\gamma(\mathbb{R}^d\times B) = \gamma(\mathbb{R}^d_{\not\sim 0}\times B) = \mu\left(\mathbb{R}^d_{\not\sim 0} \cap (\nabla\psi)^{-1}(B)\right) = \nu(B).$$

and thus $\varphi_2\gamma = \nu$. □

Proof of Theorem 3.3. *Proof of (i).* — Define B by removing the "vertical axis" $\{0\}\times\mathbb{R}^d$ from $\operatorname{spt}\gamma$ and projecting the difference set onto the second coordinate and removing the origin, i.e.,

$$B = \operatorname{proj}_2\left(\operatorname{spt}\gamma \setminus (\{0\}\times\mathbb{R}^d)\right) \setminus \{0\},$$

where $\operatorname{proj}_2(x,y) = y$ for $(x,y) \in \mathbb{R}^d\times\mathbb{R}^d$. Clearly, $B \subset \mathbb{R}^d_{\not\sim 0}$. Further, B is a Borel set because proj_2 is continuous and $\operatorname{spt}\gamma \setminus (\{0\}\times\mathbb{R}^d)$ is a countable union of compact sets, specifically $\{(x,y) \in \operatorname{spt}\gamma \mid 1/n \leqslant \|x\| \leqslant n, \|y\| \leqslant n\}$ for $n \in \mathbb{N}$, since $\operatorname{spt}\gamma$ is closed.

The measure ν is concentrated on B because (recall γ is a measure on $\mathbb{R}^{2d}_{\not<0}$)

$$\begin{aligned}\nu(\mathbb{R}^d_{\not<0} \setminus B) &= \nu\left(\mathbb{R}^d_{\not<0} \setminus \operatorname{proj}_2\left(\operatorname{spt}\gamma \setminus (\{0\}\times\mathbb{R}^d)\right)\right)\\ &= \gamma\left(\mathbb{R}^d \times \left[\mathbb{R}^d_{\not<0} \setminus \operatorname{proj}_2\left(\operatorname{spt}\gamma \setminus (\{0\}\times\mathbb{R}^d)\right)\right]\right)\\ &\leqslant \gamma\left(\mathbb{R}^{2d}_{\not<0} \setminus \left(\operatorname{spt}\gamma \setminus (\{0\}\times\mathbb{R}^d)\right)\right)\\ &\leqslant \gamma\left(\mathbb{R}^{2d}_{\not<0} \setminus \operatorname{spt}\gamma\right) + \gamma\left(\{0\}\times\mathbb{R}^d_{\not<0}\right) = 0,\end{aligned}$$

where, in the last step, we used the property that γ is proper.

Let $y \in B$. By definition of B, there exists $x \in \mathbb{R}^d$ such that $(x,y) \in \operatorname{spt}\gamma \setminus (\{0\}\times\mathbb{R}^d)$. This implies both $x \neq 0$ and $x \in \operatorname{spt}\mu$: for any open set $U \subset \mathbb{R}^d_{\not<0}$ that contains x, we have $\mu(U) = \gamma(U\times\mathbb{R}^d) > 0$, as $U\times\mathbb{R}^d$ is an open set that contains $(x,y) \in \operatorname{spt}\gamma$.

Since at least one of the two measures μ and ν has infinite mass, so has γ. As $\gamma \in \mathrm{M}_0(\mathbb{R}^d\times\mathbb{R}^d)$, necessarily $\gamma(\mathbb{B}_{2d}(0,\varepsilon) \setminus (0,0)) = \infty > 0$ for all $\varepsilon > 0$. Therefore, $(0,0)$ belongs to $\operatorname{spt}\gamma$ too.

Both (x,y) and $(0,0)$ thus belong to $\operatorname{spt}\gamma$. The monotonicity of the latter set implies $\langle x,y\rangle = \langle x-0, y-0\rangle \geqslant 0$, as required.

Proof of (ii). — For $y = 0$, there is nothing to show. Assume there exists $y \in \operatorname{spt}\nu \setminus \{0\}$ such that

$$\sup_{x\in\operatorname{spt}\mu\setminus\{0\}} \langle x/\|x\|, y\rangle < 0.$$

We will prove that this leads to a contradiction.

Let $N_y(\epsilon) = \left\{z \in \mathbb{R}^d_{\not<0} \mid \|z/\|z\| - y/\|y\|\| < \epsilon\right\}$ for $\epsilon > 0$. For $x \in \operatorname{spt}\mu \setminus \{0\}$ and $z \in N_y(\epsilon)$, the Cauchy–Schwarz inequality yields

$$\begin{aligned}\langle x/\|x\|, z/\|z\|\rangle &= \langle x/\|x\|, z/\|z\| - y/\|y\|\rangle + \langle x/\|x\|, y/\|y\|\rangle\\ &\leqslant \epsilon + \sup_{x\in\operatorname{spt}\mu\setminus\{0\}} \langle x/\|x\|, y/\|y\|\rangle.\end{aligned}$$

Taking $\epsilon < |\sup\{\langle x/\|x\|, y/\|y\|\rangle \mid x \in \operatorname{spt}\mu \setminus \{0\}\}|$ small enough, we get $\langle x,z\rangle < 0$ for all $x \in \operatorname{spt}\mu \setminus \{0\}$ and $z \in N_y(\epsilon)$, where $N_y(\epsilon)$ is an open neighbourhood of y.

Since $y \in \operatorname{spt}\nu \setminus \{0\}$ and since γ is a proper zero-coupling we have

$$0 < \nu(N_y(\epsilon)) = \gamma\left(\mathbb{R}^d \times N_y(\epsilon)\right) = \gamma\left(\operatorname{spt}\gamma \cap (\mathbb{R}^d_{\not<0} \times N_y(\epsilon))\right).$$

The set $\operatorname{spt}\gamma \cap (\mathbb{R}^d_{\not<0} \times N_y(\epsilon))$ has positive measure and is therefore not empty. Let (a,b) be a point in this set; then $a \neq 0$ and $b \in N_y(\epsilon)$. Necessarily $a \in \operatorname{spt}\mu$: for any open neighbourhood $U \subset \mathbb{R}^d_{\not<0}$ of a, we have $\mu(U) = \gamma(U\times\mathbb{R}^d) > 0$, since $U\times\mathbb{R}^d$ is an open neighbourhood of $(a,b) \in \operatorname{spt}\gamma$. By the previous paragraph, we find $\langle a,b\rangle < 0$.

However, since at least one of the two measures μ and ν has infinite mass, we have $(0,0) \in \operatorname{spt}\gamma$ by the argument in the beginning of the proof of Theorem B.7. But $\operatorname{spt}\gamma$ is monotone, and thus $\langle a,b\rangle = \langle a-0, b-0\rangle \geqslant 0$, in contradiction to the conclusion of the preceding paragraph.

We conclude that the assumption at the beginning of the proof of (ii) is wrong, which leads us to the desired conclusion. □

Lemma 3.8. *Let $A \subset \mathbb{R}^d \times \mathbb{R}^d$ satisfy the following two properties:*

(a) *(monotonicity)* $\forall (x,y),(a,b) \in A : \langle x-a, y-b\rangle \geqslant 0$*;*
(b) *(homogeneity)* $\exists \alpha, \beta > 0, \forall \lambda > 0, \forall (x,y) \in A : (\lambda^\alpha x, \lambda^\beta y) \in A$.

Then $x \in \mathrm{cl}(\mathrm{proj}_1 A)$ and $(0,b) \in A$ imply $\langle x, b\rangle \leqslant 0$.

Proof. Let $x \in \mathrm{cl}(\mathrm{proj}_1 A)$ and $(0,b) \in A$. We need to show $\langle x, b\rangle \leqslant 0$. Let $(x_n, y_n) \in A$ for $n \geqslant 1$ be such that $x_n \to x$ as $n \to \infty$. If $y_n = 0$ for all n in an infinite subset N of $\mathbb{N}$, then $\langle x_n, -b\rangle = \langle x_n - 0, 0 - b\rangle \geqslant 0$ for $n \in N$ by monotonicity and thus $\langle x, b\rangle = -\lim_{n\to\infty, n\in N} \langle x_n, -b\rangle \leqslant 0$. Otherwise, let $\lambda > 0$ and put $\lambda_n = \lambda^{1/\beta} \|y_n\|^{-1/\beta}$ and $z_n = y_n/\|y_n\|$. By homogeneity, $(\lambda_n^\alpha x_n, \lambda_n^\beta y_n) \in A$ and thus, by monotonicity,

$$\langle \lambda_n^\alpha x_n - 0, \lambda_n^\beta y_n - b\rangle \geqslant 0,$$

$$\text{from which} \qquad \lambda \langle x_n, z_n\rangle = \lambda_n^\beta \langle x_n, y_n\rangle \geqslant \langle x_n, b\rangle.$$

If z is an accumulation point of $(z_n)_n$, we get $\lambda\langle x, z\rangle \geqslant \langle x, b\rangle$. Since this is true for any $\lambda > 0$, we can take the limit as $\lambda \to 0$ to find $0 \geqslant \langle x, b\rangle$, as required. □

Proof of Theorem 3.4. We apply Theorem 3.8 to $A = \mathrm{spt}\,\gamma$, a set which is both cyclically monotone and thus monotone as well as homogeneous in the sense of Eq. (11). Writing $H_+(x) = \{y \in \mathbb{R}^d \mid \langle x, y\rangle > 0\}$, we get

$$\forall x \in \mathrm{cl}(\mathrm{proj}_1 \mathrm{spt}\,\gamma) : \qquad \mathrm{spt}\,\gamma \cap (\{0\} \times H_+(x)) = \varnothing.$$

Since $\mathrm{spt}\,\mu \subset \mathrm{cl}(\mathrm{proj}_1 \mathrm{spt}\,\gamma)$ by Theorem B.4 and thanks to the assumption on the supports of μ and ν, it follows that

$$\mathrm{spt}\,\gamma \cap (\{0\} \times (\mathrm{spt}\,\nu \setminus \{0\})) \subset \mathrm{spt}\,\gamma \cap \left(\{0\} \times \bigcup_{x \in \mathrm{spt}\,\mu} H_+(x)\right) = \varnothing.$$

As a consequence, $\gamma(\{0\} \times (\mathrm{spt}\,\nu \setminus \{0\})) = 0$. Moreover, since γ is a zero-coupling of μ and ν,

$$\begin{aligned}\gamma\left(\{0\} \times \left(\mathbb{R}^d_{\diagdown 0} \setminus \mathrm{spt}\,\nu\right)\right) &\leqslant \gamma\left(\mathbb{R}^d \times \left(\mathbb{R}^d_{\diagdown 0} \setminus \mathrm{spt}\,\nu\right)\right) \\ &= \nu\left(\mathbb{R}^d_{\diagdown 0} \setminus \mathrm{spt}\,\nu\right) = 0.\end{aligned}$$

We conclude that $\gamma(\{0\} \times \mathbb{R}^d_{\diagdown 0}) = 0$, i.e., γ is a proper zero-coupling. □

Proof of Theorem 3.5. The goal is to show that under the conditions of the statement, $\gamma(\{0\} \times \mathbb{R}^d_{\diagdown 0}) = 0$. A sufficient condition is that $\mathrm{spt}\,\gamma \cap \left(\{0\} \times \mathbb{R}^d_{\diagdown 0}\right) = \varnothing$. We proceed by contradiction and assume the existence of $b \in \mathbb{R}^d_{\diagdown 0}$ such that $(0,b) \in \mathrm{spt}\,\gamma$. Then also $b \in \mathrm{spt}\,\nu$, since for any $0 < r < \|b\|$, we have, by definition of $\mathrm{spt}\,\gamma$,

$$\nu(\mathbb{B}(b,r)) = \gamma(\mathbb{R}^d \times \mathbb{B}(b,r)) > 0,$$

since $\mathbb{R}^d \times \mathbb{B}(b,r)$ is an open set containing $(0,b)$.

From the assumption in the statement, there exists $\varepsilon > 0$ such that $\mu(H_+(b,\varepsilon)) = \infty$, with $H_+(b,\varepsilon)$ defined in Eq. (12). For the latter ε, consider the set

$$A = \{y \in \mathbb{R}^d \mid \exists x_y \in H_+(b,\epsilon) : (x_y, y) \in \mathrm{spt}\,\gamma\}$$

$$= \mathrm{proj}_2\left(\mathrm{spt}\,\gamma \cap \left(H_+(b,\epsilon)\times\mathbb{R}^d\right)\right).$$

(Note that if $\mathrm{spt}\,\gamma$ corresponds to the graph of a mapping $\mathbb{R}^d \to \mathbb{R}^d$, then A is the image of $H_+(b,\varepsilon)$ by the latter mapping.) The set A is a Borel set because the projection proj_2 is continuous and $\mathrm{spt}\,\gamma \cap \left(H_+(b,\epsilon)\times\mathbb{R}^d\right)$ is equal to the countable union of compact sets

$$\{(x,y)\in\mathrm{spt}\,\gamma \mid 1/n \leqslant \|x\| \leqslant n,\ \|y\|\leqslant n, \langle x/\|x\|, b\rangle \geqslant \varepsilon + 1/n\}, \qquad n\in\mathbb{N}.$$

Step 1. — We first show that $A \subset \mathbb{B}(0,\varepsilon)^c$. Let $y\in A$ and let $x_y \in H_+(b,\epsilon)$ be such that $(x_y,y)\in\mathrm{spt}\,\gamma$. Then by the Cauchy–Schwarz inequality, we have

$$\|y\| \geqslant \langle x_y/\|x_y\|, y\rangle.$$

Now because $(0,b)$ and (x_y,y) both belong to the monotone set $\mathrm{spt}\,\gamma$, we have $\langle x_y, y-b\rangle = \langle x_y - 0, y-b\rangle \geqslant 0$, whence $\langle x_y,y\rangle \geqslant \langle x_y,b\rangle$ and thus, from the definition of $H_+(b,\varepsilon)$,

$$\langle x_y/\|x_y\|, y\rangle \geqslant \langle x_y/\|x_y\|, b\rangle \geqslant \varepsilon.$$

The latter two displays prove that $\|y\|\geqslant\varepsilon$, as claimed.

Step 2. — Next we show that $\nu(A)=+\infty$, from which the contradiction will follow, since by definition of $\mathrm{M}_0(\mathbb{R}^d)$ and by Step 1, we should have $\nu(A)\leqslant\nu(\mathbb{B}(0,\varepsilon)^c)<\infty$. From the definition of A, and since $0\notin A$ by Step 1, we have

$$\begin{aligned}\nu(A) &= \gamma(\mathbb{R}^d\times A)\\ &\geqslant \gamma\left(\{(x,y)\in\mathrm{spt}\,\gamma \mid x\in H_+(b,\epsilon), y\in A\}\right)\\ &= \gamma\left(\{(x,y)\in\mathrm{spt}\,\gamma \mid x\in H_+(b,\epsilon)\}\right)\\ &= \gamma\left(H_+(b,\epsilon)\times\mathbb{R}^d\right) = \mu\left(H_+(b,\epsilon)\right) = +\infty,\end{aligned}$$

leading to the stated contradiction. □

Proof of Theorem 3.6. It suffices to check that the condition in Theorem 3.5 is satisfied. Let $y\in\mathrm{spt}\,\nu\setminus\{0\}$. By assumption there exists $x\in\mathrm{spt}\,\mu$ such that $\langle x,y\rangle>0$. We can find $\epsilon>0$ such that $x\in H_+(y,\epsilon)$ and then, by continuity, also $\mathbb{B}(x,\delta)\subset H_+(y,\epsilon)$ for some $\delta>0$. As $x\in\mathrm{spt}\,\mu$, we have $\mu(\mathbb{B}(x,\delta))>0$.

(i) By assumption, $\mu\left(\mathrm{pos}\,\mathbb{B}(x,\delta)\right)=\infty$; note that $0\notin\mathbb{B}(x,\delta)$ by the choice of δ. Since $H_+(y,\epsilon)$ is a cone and $\mathbb{B}(x,\delta)\subset H_+(y,\epsilon)$, the inclusion $\mathrm{pos}\,\mathbb{B}(x,\delta)\subset H_+(y,\epsilon)$ holds too. Therefore, $\mu(H_+(y,\epsilon))=\infty$ and Theorem 3.5 can be applied.

(ii) This is a particular case of the last point. Let $x\in\mathrm{spt}\,\mu$ and $\delta>0$. For any $\lambda>0$, we have $\lambda\mathbb{B}(x,\delta)\subset\mathrm{pos}\,\mathbb{B}(x,\delta)$ and thus $\mu\left(\mathrm{pos}\,\mathbb{B}(x,\delta)\right)\geqslant\mu(\lambda\mathbb{B}(x,\delta))=\lambda^{-\alpha}\mu(\mathbb{B}(x,\delta))$. Since $\mu(\mathbb{B}(x,\delta))>0$, we find assumption (i) by taking the limit as $\lambda\to 0$. □

Proof of Theorem 3.7. If $\mathrm{spt}\,\gamma\cap(\{0\}\times\mathbb{R}_{\diagdown 0})=\varnothing$, then $\gamma\left(\{0\}\times\mathbb{R}_{\diagdown 0}\right)=0$ and thus γ is proper. So assume $\mathrm{spt}\,\gamma$ intersects $\{0\}\times\mathbb{R}_{\diagdown 0}$; we will show that still $\gamma\left(\{0\}\times\mathbb{R}_{>0}\right)=\gamma\left(\{0\}\times\mathbb{R}_{<0}\right)=0$.

Assume there exists $b>0$ such that $(0,b)\in\mathrm{spt}\,\gamma$; if such b does not exist, then $\mathrm{spt}\,\gamma$ does not interset $\{0\}\times\mathbb{R}_{>0}$ and thus $\gamma(\{0\}\times\mathbb{R}_{>0})=0$, and we can move to the next paragraph. Consider $(x,y)\in\mathrm{spt}\,\gamma$ with $x>0$; by monotonicity of $\mathrm{spt}\,\gamma$, we have $(x-0)(y-b)\geqslant 0$, i.e., $y\geqslant b$. It follows that

$$\mathrm{spt}\,\gamma\cap(\mathbb{R}_{>0}\times\mathbb{R})\subset\mathbb{R}_{>0}\times[b,\infty).$$

But then $\gamma(\mathbb{R}_{>0}\times\mathbb{R}) = \gamma(\mathbb{R}_{>0}\times[b,\infty))$ and thus

$$\begin{aligned}\nu(\mathbb{R}_{>0}) \geqslant \nu\left([b,\infty)\right) &= \gamma\left(\mathbb{R}\times[b,\infty)\right)\\ &\geqslant \gamma\left(\mathbb{R}_{>0}\times[b,\infty)\right)\\ &= \gamma\left(\mathbb{R}_{>0}\times\mathbb{R}\right) = \mu(\mathbb{R}_{>0}) \geqslant \nu(\mathbb{R}_{>0}).\end{aligned}$$

As a consequence, the inequalities above become equalities, the common quantity being finite, and thus

$$\gamma\left(\{0\}\times[b,\infty)\right) \leqslant \gamma\left(\mathbb{R}_{\leqslant 0}\times[b,\infty)\right) \leqslant \gamma\left(\mathbb{R}\times[b,\infty)\right) - \gamma\left(\mathbb{R}_{>0}\times[b,\infty)\right) = 0 \qquad \text{and}$$
$$\gamma\left(\{0\}\times(0,b)\right) \leqslant \gamma\left(\mathbb{R}\times(0,b)\right) = \nu((0,b)) = \nu(\mathbb{R}_{>0}) - \nu([b,\infty)) = 0,$$

yielding $\gamma\left(\{0\}\times\mathbb{R}_{>0}\right) = 0$.

Similarly, if there does not exist $b<0$ such that $(0,b)\in\operatorname{spt}\gamma$, then trivially $\gamma(\{0\}\times\mathbb{R}_{<0}) = 0$, while if such a negative number b exists, we still have $\gamma(\{0\}\times\mathbb{R}_{<0}) = 0$ by the same argument as in the previous paragraph.

Reciprocally, assume there exists $\gamma\in\Gamma^{\mathrm{p}}_{0,\mathrm{m}}(\mu,\nu)$ and assume $\mu\left(\mathbb{R}\setminus\{0\}\right) = \infty$ or $\nu\left(\mathbb{R}\setminus\{0\}\right) = \infty$. Then also $\gamma(\mathbb{R}^2\setminus\{(0,0)\}) = \infty$ and thus $\gamma([-r,r]^2) = \infty > 0$ for every $r>0$, which implies $(0,0)\in\operatorname{spt}\gamma$. Moreover, as $\operatorname{spt}\gamma$ is monotone, we find that any $(\tilde{x},\tilde{y})\in\operatorname{spt}\gamma$ satisfies $\tilde{x}\tilde{y} = (\tilde{x}-0)(\tilde{y}-0)\geqslant 0$, i.e., $\operatorname{spt}\gamma\subset[0,\infty)^2\cup(-\infty,0]^2$; intuitively, μ-mass from $\mathbb{R}_{>0}$ (resp. $\mathbb{R}_{<0}$) cannot be sent to $\mathbb{R}_{<0}$ (resp. $\mathbb{R}_{>0}$) by the zero-coupling γ. As a consequence,

$$\begin{aligned}\mu(\mathbb{R}_{>0}) &= \gamma(\mathbb{R}_{>0}\times\mathbb{R})\\ &= \gamma\left(\mathbb{R}_{>0}\times\mathbb{R}_{\geqslant 0}\right)\\ &= \gamma\left(\mathbb{R}_{\geqslant 0}\times\mathbb{R}_{\geqslant 0}\right)\\ &\geqslant \gamma\left(\mathbb{R}_{\geqslant 0}\times\mathbb{R}_{>0}\right) = \nu(\mathbb{R}_{>0}),\end{aligned}$$

where we used that γ is proper, i.e., $\gamma(\{0\}\times\mathbb{R}) = 0$. Bu the same reasoning, we can also prove $\mu(\mathbb{R}_{<0})\geqslant\nu(\mathbb{R}_{<0})$. □

4. Brenier–McCann theorem for zero-couplings

In McCann [25] it is proven that, for a probability measure P on $\mathbb{R}^d$ vanishing on sets of Hausdorff dimension at most $d-1$ and for closed convex functions ϕ and ψ, the identity $(\mathrm{Id}\times\nabla\psi)_{\#}P = (\mathrm{Id}\times\nabla\phi)_{\#}P$ implies $\nabla\phi = \nabla\psi$ P-a.e. In our setting, we prove that the same conclusion holds under appropriate assumptions when an improper term is added as in the representation from Theorem 2.9. The proof is inspired by the one of the uniqueness assertion of Main Theorem in McCann [25], which relies in turn on a refinement of arguments from Aleksandrov [3]. However, the fact that we are only dealing with zero-couplings and that some quantities may be infinite makes the proof substantially more complicated. Recall that we write $\operatorname{res} m$ for the restriction of a Borel measure m on $\mathbb{R}^k$ to $\mathbb{R}^k\setminus\{0\}$.

Theorem 4.1 (Uniqueness of zero-couplings). *Let $\mu,\nu\in\mathrm{M}_0(\mathbb{R}^d)$ and assume μ vanishes on sets of Hausdorff dimension at most $d-1$. Let $\gamma_1,\gamma_2\in\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ and let $\psi_1,\psi_2:\mathbb{R}^d\to\mathbb{R}\cup\{\infty\}$ be closed convex functions such that $\operatorname{spt}\gamma_i\subset\partial\psi_i$ for $i\in\{1,2\}$. If $(0,0)\in\partial\psi_1\cap\partial\psi_2$, then $\nabla\psi_1 = \nabla\psi_2$ μ-almost everywhere and thus $\gamma_1=\gamma_2$. In particular, if at least one of μ or ν has infinite mass, then $\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ is a singleton.*

The consequence $\gamma_1 = \gamma_2$ stated in Theorem 4.1 follows from the representation in Theorem 2.9. In the proof of Theorem 4.1, we actually show the slightly stronger statement that $\nabla\psi_1(x) = \nabla\psi_2(x)$ for any $x \in \operatorname{spt}\mu$ in the domains of both $\nabla\psi_1$ and $\nabla\psi_2$.

In the proof of Theorem 4.1, we cannot proceed directly as in the proof of the Main Theorem in [25] and take any $p \in \operatorname{spt}\mu$ such that $\nabla\psi_1(p) \neq \nabla\psi_2(p)$ to obtain a contradiction, since the pathological set M obtained in that proof via Aleksandrov's lemma must have finite measure, as otherwise an expression of the form infinity minus infinity would arise. This comes at the price of a long build-up in Steps 1–5 in the proof of Theorem 4.1 before arriving at the said contradiction in Step 6.

Remark 4.2. In Theorem 4.1, the assumption $(0,0) \in \partial\psi_1 \cap \partial\psi_2$ implies that the convex functions ψ_1 and ψ_2 attain their global minima at the origin and that 0 is left invariant by their gradients when they are defined at the origin. Because of the nature of zero-couplings, the assumption is essential for the uniqueness of the gradient to hold. Indeed, as a counter-example in dimension $d = 1$, let μ and ν be the Lebesgue measures on $[-1,0]$ and $[0,1]$, respectively. Then $\gamma_1 = \mu \otimes \delta_0 + \delta_0 \otimes \nu$ is a zero-coupling of μ and ν, and belongs to $\Gamma_{0,\mathrm{cm}}(\mu,\nu)$, but so does the measure $\gamma_2 = (\mathrm{Id}\times\nabla\psi)_{\#}\mu$ where $\nabla\psi(x) = x + 1$. In particular, the support of γ_1 is contained in the cyclically monotone set $(\mathbb{R}_{\leqslant 0} \times \{0\}) \cup (\{0\} \times \mathbb{R}_{\geqslant 0})$, and the one of γ_2 in the cyclically monotone set $\{(x, x+1) : x \in \mathbb{R}\}$; incidentally, note that γ_2 is a true coupling rather than just a zero-coupling.

Remark 4.3. When μ and ν have equal finite mass, then their extensions $\tilde{\mu} \in [\mu]$ and $\tilde{\nu} \in [\nu]$ to $\mathbb{R}^d$ defined by $\tilde{\mu}(\{0\}) = \tilde{\nu}(\{0\}) = 0$ have equal finite mass too, and Corollary 14 in McCann [25] yields a unique cyclically monotone coupling π between $\tilde{\mu}$ and $\tilde{\nu}$. It is clear that its restriction $\operatorname{res}\pi$ to $\mathbb{R}^{2d}_{\diagdown 0}$ lies in $\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ while there might be infinitely many other zero-couplings whose supports are cyclically monotone. We refer to the example in Remark 4.2.

Theorem 4.1 concerns the uniqueness of zero-couplings with cyclically monotone support and requires a condition on the origin $(0,0)$ in $\mathbb{R}^d \times \mathbb{R}^d$. This condition is not needed in the next result, which concerns proper representations.

Theorem 4.4 (Uniqueness of proper representations). *Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ with μ vanishing on sets of Hausdorff dimension at most $d-1$, and let ψ_1, ψ_2 be two closed convex functions such that $\nu = \operatorname{res}((\nabla\psi_j)_{\#}\mu)$ and $\mu(\mathbb{R}^d_{\diagdown 0} \setminus \operatorname{dom}\nabla\psi_j) = 0$ for $j = 1, 2$. Then $\nabla\psi_1 = \nabla\psi_2$ μ-almost everywhere.*

A combination of Theorems 2.8, 2.9 and 4.1 leads to a result similar to the Main Theorem in McCann [25], where no moment assumption is required. The major differences are the necessity of at least one of the zero-marginals to have infinite mass, and the addition of an improper term in the representation when there is a need for mass to be sent from the origin outside of the punctured space.

Corollary 4.5 (Brenier–McCann theorem for zero-couplings). *Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ be such that at least one of them has infinite mass and suppose μ vanishes on sets of Hausdorff dimension at most $d-1$. Then $\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ is a singleton with single element γ, say, and there exists a closed convex function ψ on $\mathbb{R}^d$ with $\operatorname{spt}\gamma \subset \partial\psi$ such that*

$$\gamma(A \times B) = (\mathrm{Id}\times\nabla\psi)_{\#}\mu(A \times B) = \mu\left(A \cap (\nabla\psi)^{-1}(B)\right), \qquad A \in \mathcal{B}(\mathbb{R}^d_{\diagdown 0}),\ B \in \mathcal{B}(\mathbb{R}^d),$$

$$\gamma(\{0\} \times B) = \nu(B) - (\nabla\psi)_{\#}\mu(B), \qquad B \in \mathcal{B}(\mathbb{R}^d),\ 0 \notin \mathrm{cl}(B).$$

Although ψ may not be unique, the map $\nabla\psi$ is defined and uniquely determined μ-almost everywhere. If, in addition, γ is proper, then

$$\nu(B) = (\nabla\psi)_{\#}\mu(B) = \mu\left(\mathbb{R}^d_{\setminus 0} \cap (\nabla\psi)^{-1}(B)\right), \qquad B \in \mathcal{B}(\mathbb{R}^d_{\setminus 0}). \tag{14}$$

As we are interested in the existence of a transport map from μ to ν as in Eq. (14), knowing whether the unique cyclically monotone coupling in Theorem 4.5 is proper or not is crucial. Sufficient criteria have been studied in Section 3 and will be applied in Section 5 to zero-couplings between limit measures in the definition of multivariate regular variation.

Remark 4.6. A complication in the proof of Theorem 4.1 is that the convex functions ψ_i are not necessarily continuous at $0 \in \mathbb{R}^d$ relative to their effective domains, $\operatorname{dom}\psi_i = \{\psi_i < \infty\}$. In dimension $d = 2$, define

$$\psi(x,y) = \begin{cases} y^4/x^2 & \text{if } x > 0, \\ 0 & \text{if } (x,y) = (0,0), \\ \infty & \text{otherwise,} \end{cases}$$

a variation on the counterexample in Rockafellar [31, pp. 83–84]. This function is convex (its Hessian matrix has positive trace and determinant), but is not continuous in $(0,0)$ relative to its effective domain, since $\psi(y^2, y) = 1$ for all $y > 0$. Still, it may serve as the potential function of a cyclically monotone transport between two infinite measures in $\mathrm{M}_0(\mathbb{R}^d)$. Let μ be the measure on the half-plane $\{(x,y) \mid x > 0\}$ with Lebesgue density

$$\mu(\mathrm{d}(x,y)) = \frac{x^3}{(x^2+y^2)^3}\mathrm{d}x\,\mathrm{d}y, \qquad x > 0, y \in \mathbb{R}.$$

In polar coordinates $(x,y) = (r\cos\theta, r\sin\theta)$, its density is

$$(T_{\text{polar}})_{\#}\mu(\mathrm{d}(r,\theta)) = (\cos\theta)^3 \mathrm{d}\theta\, r^{-2}\mathrm{d}r, \qquad r > 0, -\pi/2 < \theta < \pi/2,$$

where $T_{\text{polar}}(x,y) = (r,\theta)$ are the polar coordinates of $(x,y) \in \mathbb{R}^2_{\setminus 0}$. We find that $\mu \in \mathrm{M}_0(\mathbb{R}^2)$ and that μ is a possible limit measure in the definition of multivariate regular variation in Section 5 below. The push-forward measure $\nu = (\nabla\psi)_{\#}\mu$ is supported on the opposite half-plane $\{(x,y) \mid x < 0\}$. Its densities in Cartesian and polar coordinates can be computed to be

$$\nu(\mathrm{d}(x,y)) = \frac{64\,|x|^3}{(4x^2+y^2)^3}\mathrm{d}x\,\mathrm{d}y, \qquad x < 0, y \in \mathbb{R};$$

$$(T_{\text{polar}})_{\#}\nu(\mathrm{d}(r,\theta)) = \frac{64\,|\cos\theta|^3}{(1+2\cos^2\theta)^3}\mathrm{d}\theta\, r^{-2}\mathrm{d}r, \qquad r > 0, \pi/2 < \theta < 3\pi/2.$$

Since its angular density is bounded, we have $\nu \in \mathrm{M}_0(\mathbb{R}^2)$ too.

Remark 4.7. While Brenier's Theorem characterizes the unique possible solution to the Kantorovich problem with quadratic distance as cost, Theorem 4.5 provides, when at least one of the zero-marginals has infinite mass, an analogous representation for the unique candidate solution to the variant of this problem used in Guillen et al. [17] to define a transportation metric for Lévy measures.

Proofs

We first prove a convenient extension of Aleksandrov's lemma as stated in [25, Lemma 13]. Here and below, if $B \subset \mathbb{R}^d$, and if ψ is a convex function $\mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$, we define

$$(\partial\psi)^{-1}(B) = \left\{x \in \mathbb{R}^d \mid \exists y \in B : y \in \partial\psi(x)\right\} = \left\{x \in \mathbb{R}^d \mid \partial\psi(x) \cap B \neq \varnothing\right\}.$$

Lemma 4.8 (Aleksandrov's Lemma, extended). *Let $\phi, \psi : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ be convex functions, differentiable at some $p \in \mathbb{R}^d$, with $\nabla\phi(p) \neq \nabla\psi(p)$. Define $M = \{\phi - \phi(p) > \psi - \psi(p)\} \subset \operatorname{dom}\psi$, $X = (\nabla\psi)^{-1}(\partial\phi(M))$ and $\tilde{X} = (\partial\psi)^{-1}(\partial\phi(M))$. Then $X \subset \tilde{X} \subset M$, while p lies at a positive distance from X, i.e., there is $r > 0$ such that $X \cap \mathbb{B}(p, r) = \varnothing$.*

Proof. The inclusion $X \subset \tilde{X}$ follows immediately from the definitions, since $\nabla\psi(x) = y$ is equivalent to $\partial\psi(x) = \{y\}$.

a. In case $\phi(p) = \psi(p) = 0$ and $\nabla\psi(p) = 0$, the result follows from the proof of Lemma 13 in [25] that $X \subset M$. That proof also shows that $\tilde{X} \subset M$ upon noting that by definition of $(\partial\psi)^{-1}$, for any $x \in \tilde{X}$, there indeed exists $y \in \partial\phi(M)$ such that $y \in \partial\psi(x)$.

b. In the general case, define

$$\bar{\phi}(x) = \phi(x) - \phi(p) - \langle\nabla\psi(p), x - p\rangle,$$
$$\bar{\psi}(x) = \psi(x) - \psi(p) - \langle\nabla\psi(p), x - p\rangle.$$

Clearly, $\bar{\phi}(p) = \bar{\psi}(p) = 0$. For all $x \in \operatorname{dom}\partial\phi = \operatorname{dom}\partial\bar{\phi}$, we have $\partial\bar{\phi}(x) = \partial\phi(x) - \nabla\psi(p) = \{y - \nabla\psi(p) \mid y \in \partial\phi(x)\}$, and for all $x \in \operatorname{dom}\partial\psi = \operatorname{dom}\partial\bar{\psi}$, we have $\partial\bar{\psi}(x) = \partial\psi(x) - \nabla\psi(p)$; in particular, $p \in \operatorname{dom}\nabla\bar{\psi}$ and $\nabla\bar{\psi}(p) = 0$. Applying part **a.** to the functions $\bar{\phi}, \bar{\psi}$ at point p, with $M' = \{\bar{\phi} > \bar{\psi}\}$, $\tilde{X}' = (\partial\bar{\psi})^{-1}(\partial\bar{\phi}(M'))$ and $X' = (\nabla\bar{\psi})^{-1}(\partial\bar{\phi}(M'))$, we obtain the inclusions $X' \subset \tilde{X}' \subset M'$ and the fact that p is at positive distance from X'. From the definitions, we have

$$M' = \left\{x \in \mathbb{R}^d \mid \phi(x) - \phi(p) > \psi(x) - \psi(p)\right\} = M,$$

and then

$$\begin{aligned}\tilde{X}' &= \left\{x \in \operatorname{dom}\partial\bar{\psi} \mid \partial\bar{\psi}(x) \cap \partial\bar{\phi}(M') \neq \varnothing\right\}\\ &= \{x \in \operatorname{dom}\partial\psi \mid (\partial\psi(x) - \nabla\psi(p)) \cap (\partial\phi(M) - \nabla\psi(p)) \neq \varnothing\}\\ &= \{x \in \operatorname{dom}\partial\psi \mid \partial\psi(x) \cap \partial\phi(M) \neq \varnothing\} = \tilde{X}.\end{aligned}$$

A similar reasoning yields $X' = X$. The desired result then follows. □

Proof of Theorem 4.1. For convenience, we will change notation and let the two convex functions ψ_1 and ψ_2 be denoted by ψ and ϕ instead, and the two zero-couplings γ_1 and γ_2 with supports contained in their subdifferentials by γ_ψ and γ_ϕ, respectively.

Proof of $\gamma_\psi = \gamma_\phi$

Step 1. — By Theorem 2.9, ψ and ϕ are differentiable μ-almost everywhere, while $\operatorname{spt}\mu$ is contained in $\operatorname{cl}(\operatorname{dom}\psi) \cap \operatorname{cl}(\operatorname{dom}\phi)$, both closures being convex and therefore having boundaries of Hausdorff dimension at most $d - 1$. As we assumed $(0, 0) \in \partial\psi \cap \partial\phi$, the functions ψ and ϕ attain their minima at

0, so that $\psi(0)$ and $\phi(0)$ are certainly finite. Therefore, without losing generality, take $\psi(0) = \phi(0) = 0$. Furthermore, we can restrict attention to the interiors of the domains of ψ and ϕ via

$$\mathcal{D} := \operatorname{int}(\operatorname{dom}\psi) \cap \operatorname{int}(\operatorname{dom}\phi) = \operatorname{int}(\operatorname{dom}\psi \cap \operatorname{dom}\phi), \tag{15}$$

since the above argument shows that $\tilde{\mu}(\mathbb{R}^d \setminus \mathcal{D}) = 0$, where, as before, $\tilde{\mu}$ is the extension of μ to $\mathbb{R}^d$ determined by $\tilde{\mu}(\{0\}) = 0$. Consider the following subsets of $\mathcal{D}$ [see Eq. (47)]:

$$V := \operatorname{dom}\nabla\psi \cap \operatorname{dom}\nabla\phi, \tag{16}$$

$$W := \{y \in V \mid \nabla\psi(y) \neq \nabla\phi(y)\},$$

$$A := \{y \in \mathcal{D} \mid \psi(y) - \phi(y) \neq 0\}. \tag{17}$$

By Alberti and Ambrosio [2, Theorem 2.2] (see also [5]), $\mathcal{D} \setminus V$ is contained in the union of two sets with Hausdorff dimension at most $d-1$, so

$$\tilde{\mu}(\mathcal{D} \setminus V) = 0 \quad \text{and thus} \quad \tilde{\mu}(\mathbb{R}^d \setminus V) = 0. \tag{18}$$

Our claim that $\nabla\psi = \nabla\phi$ μ-almost everywhere is thus equivalent to $\tilde{\mu}(W) = 0$.

We first prove the claim made below the statement of the theorem, namely we show that $\tilde{\mu}(W) = 0$ if and only if for all $x \in \operatorname{spt}\mu \cap V$ we have $\nabla\psi(x) = \nabla\phi(x)$, or in other words, if and only if

$$W \cap \operatorname{spt}\mu = \varnothing. \tag{19}$$

Of course if Eq. (19) holds true then $\tilde{\mu}(W) = 0$. Conversely, if the intersection in Eq. (19) is non-empty, say there exists $x \in W \cap \operatorname{spt}\mu$, then by continuity of the gradients on their domain and the definition of $\operatorname{spt}\mu$, there exists an open ball $\mathbb{B}(x,r)$ such that $\mathbb{B}(x,r) \cap V \subset W$ and $\tilde{\mu}(\mathbb{B}(x,r)) > 0$. But then $\tilde{\mu}(\mathbb{B}(x,r) \cap V) = \tilde{\mu}(\mathbb{B}(x,r)) > 0$ and thus $\tilde{\mu}(W) > 0$.

To prove that $\tilde{\mu}(W) = 0$, we assume the opposite,

$$\tilde{\mu}(W) > 0, \tag{20}$$

and we will construct a subset M of $\mathbb{R}^d_{<0}$ such that $\mathbb{R}^d \times M$ receives different masses from γ_ψ and γ_φ, in contradiction to the assumption of the theorem that both measures are a zero-coupling of μ and ν.

Step 2. — By Lemma A.5, the set

$$W \setminus A = \{y \in V \mid \psi(y) - \phi(y) = 0 \text{ and } \nabla\psi(y) \neq \nabla\phi(y)\}$$

has Hausdorff dimension at most $d-1$. Therefore, $\tilde{\mu}(W \setminus A) = 0$. By (20), we obtain

$$\tilde{\mu}(W \cap A) > 0. \tag{21}$$

Step 3. — Recall $\mathcal{D}$ in (15). For $t \in [-\infty, +\infty]$, define

$$A_t^- := \{y \in \mathcal{D} \mid \psi(y) - \phi(y) < t\},$$

$$A_t^+ := \{y \in \mathcal{D} \mid \psi(y) - \phi(y) > t\},$$

$$A_t^0 := \{y \in \mathcal{D} \mid \psi(y) - \phi(y) = t\}.$$

Then $A_0^- \cup A_0^+$ is a partition of A in (17) and thus, by (21),

$$\tilde{\mu}(W \cap A_0^-) + \tilde{\mu}(W \cap A_0^+) = \tilde{\mu}(W \cap A) > 0.$$

Therefore, $\tilde{\mu}(W \cap A_0^-) > 0$ or $\tilde{\mu}(W \cap A_0^+) > 0$. Without loss of generality, assume

$$\tilde{\mu}(W \cap A_0^-) > 0. \tag{22}$$

In the other case, just switch the roles of ψ and ϕ in the next steps.

Step 4. — The goal of this step is to show that for all $t < 0$, we have $\mu(A_t^-) < \infty$.

Fix $t < 0$. Since $(0,0) \in \partial\psi \cap \partial\phi$, we were able to choose $\psi(0) = \phi(0) = 0$ in Step 1. Consider the set

$$N = \{x \in \operatorname{dom} \nabla\phi \setminus \{0\} \mid \|\nabla\phi(x)\| \geqslant 1\}.$$

From Theorem 2.9 we have for every Borel set C bounded away from 0,

$$\infty > \nu(C) = (\nabla\phi)_{\#}\mu(C) + \gamma(\{0\} \times C) \geqslant (\nabla\phi)_{\#}\mu(C),$$

since $\nu \in \mathrm{M}_0(\mathbb{R}^d)$. With $C = \mathbb{B}(0,1)^c$, we obtain

$$\infty > (\nabla\phi)_{\#}\mu\left(\mathbb{B}(0,1)^c\right) = \mu\left(\{x \in \mathbb{R}^d_{\diagdown 0} \cap \operatorname{dom} \nabla\phi \mid \|\nabla\phi(x)\| \geqslant 1\}\right) = \mu(N).$$

Hence $\mu(N) < \infty$. Letting

$$S = \operatorname{dom} \nabla\phi \setminus (N \cup \{0\}),$$

we have, since $\mathbb{R}^d_{\diagdown 0} \setminus S = \left(\mathbb{R}^d_{\diagdown 0} \setminus \operatorname{dom} \nabla\phi\right) \cup N$ and in view of Eq. (18),

$$S \subset \operatorname{dom} \nabla\phi \subset \operatorname{dom} \phi, \quad \mu(\mathbb{R}^d_{\diagdown 0} \setminus S) < \infty, \text{ and } \|\nabla\phi(x)\| < 1 \text{ for all } x \in S. \tag{23}$$

For any $x \in S$, it thus holds that

$$0 = \phi(0) \geqslant \phi(x) + \langle 0 - x, \nabla\phi(x)\rangle,$$

and therefore also

$$\phi(x) \leqslant \langle x, \nabla\phi(x)\rangle \leqslant \|x\|.$$

We may now show that $\mu(A_t^-) < \infty$. Since $\mu(\mathbb{R}^d_{\diagdown 0} \setminus S) < \infty$ by Eq. (23), it suffices to check that $\mu(S \cap A_t^-) < \infty$. But from the above display and the definition of A_t^- [if $x \in A_t^-$, then $\phi(x) > \psi(x) + |t| \geqslant |t| > 0$], we have the inclusions

$$\begin{aligned} S \cap A_t^- &\subset \{x \in \operatorname{dom} \phi \mid \phi(x) > |t| \text{ and } \phi(x) \leqslant \|x\|\} \\ &\subset \{x \in \mathbb{R}^d \mid \|x\| \geqslant |t|\} = \mathbb{B}(0,|t|)^c. \end{aligned}$$

Hence, $0 \notin \mathrm{cl}(S \cap A_t^-)$ and thus $\mu(S \cap A_t^-) < \infty$, as required.

Step 5. — Recall V in (16). We claim that there exists $z \in V$ with the following three properties:

- $t := \psi(z) - \phi(z) < 0$;
- $\nabla\psi(z) \neq \nabla\phi(z)$;

- for all $\delta > 0$,

$$\mu\left(A_t^- \cap \mathbb{B}(z,\delta)\right) > 0. \tag{24}$$

To prove the claim, define, for every $t < 0$, the sets

$$\begin{aligned}\Omega_t &:= A_t^0 \cap W \cap \operatorname{spt}\mu \\ &= \{x \in V \mid \psi(x) - \phi(x) = t \text{ and } \nabla\psi(x) \neq \nabla\phi(x)\} \cap \operatorname{spt}\mu, \\ \Xi_t &:= \{x \in \Omega_t \mid \exists \delta > 0, \mu(A_t^- \cap \mathbb{B}(x,\delta)) = 0\}.\end{aligned}$$

We need to show that there exists $t < 0$ such that $\Omega_t \setminus \Xi_t \neq \varnothing$. To this end, define

$$\begin{aligned}\mathcal{T}_\Omega &:= \{t \in (-\infty, 0) \mid \Omega_t \neq \varnothing\}, \\ \mathcal{T}_\Xi &:= \{t \in (-\infty, 0) \mid \Xi_t \neq \varnothing\}.\end{aligned}$$

We will show that $\mathcal{T}_\Omega$ is uncountable while $\mathcal{T}_\Xi$ is at most countable, so that there must exist (uncountable many) $t < 0$ such that Ω_t is nonempty and Ξ_t is empty.

By Lemma A.5, for every $t < 0$, the set Ω_t has Hausdorff measure at most $d - 1$ and is therefore a μ-null set. Still, their union over all t less than 0 is

$$\textstyle\bigcup_{t<0} \Omega_t = \bigcup_{t<0} A_t^0 \cap W \cap \operatorname{spt}\mu = A_0^- \cap W \cap \operatorname{spt}\mu,$$

a set receiving positive mass from μ by (22). Therefore, $\mathcal{T}_\Omega$ must be uncountable.

We have $\Xi_t = \bigcup_{n\in\mathbb{N}} \Xi_{t,n}$ and thus $\mathcal{T}_\Xi = \bigcup_{n\in\mathbb{N}} \mathcal{T}_n$, where, for $n \in \mathbb{N}$,

$$\begin{aligned}\Xi_{t,n} &:= \{x \in \Omega_t \mid \mu(A_t^- \cap \mathbb{B}(x, 2/n)) = 0\}, \\ \mathcal{T}_n &:= \{t \in (-\infty, 0) \mid \Xi_{t,n} \neq \varnothing\}.\end{aligned}$$

To show that $\mathcal{T}_\Xi$ is countable, it is sufficient to show that each $\mathcal{T}_n$ is countable.

Define

$$\forall t \in \mathcal{T}_n, \qquad G_t := \textstyle\bigcup_{x\in\Xi_{t,n}} \mathbb{B}(x, 1/n).$$

Each G_t is non-empty and open, and by the argument just given, $G_s \cap G_t = \varnothing$ whenever $s, t \in \mathcal{T}_n$ with $s < t$. Since $\mathbb{R}^d$ is separable, the cardinality of the collection of sets $\{G_t\}_{t\in\mathcal{T}_n}$ is at most countable, and therefore $\mathcal{T}_n$ itself is at most countable, as required. The claim made at the beginning of this step is proved.

Step 6. — Let $t < 0$, and $z \in V$ from Step 5, such that $t = \psi(z) - \phi(z) < 0$. Define

$$M = \{y \in \mathbb{R}^d \mid \psi(y) - \phi(y) < t\} = \{y \in \mathbb{R}^d \mid \psi(y) - \psi(z) < \phi(y) - \psi(z)\}, \tag{25}$$

so that $A_t^- = M \cap \mathcal{D}$. By Step 4, and since $\tilde{\mu}(\mathcal{D}^c) = 0$, we have $\mu(M) = \mu(A_t^-) < \infty$.

If $x \in A_t^-$, it holds that

$$\phi(x) > \psi(x) + |t| \geqslant |t| > 0,$$

where the second inequality above follows from $0 \in \partial\psi(0)$ and $\psi(0) = 0$; hence x is not a minimum of ϕ and thus $0 \notin \partial\phi(x)$. Hence, with $B = \partial\phi(A_t^-)$, we obtain

$$0 \notin B. \tag{26}$$

The set $B = \mathrm{proj}_2((A_t^- \times \mathbb{R}^d) \cap \partial\phi)$ is σ-compact and thus Borel measurable. Indeed, since proj_2 is continuous, the set B is σ-compact as soon as $(A_t^- \times \mathbb{R}^d) \cap \partial\phi$ is σ-compact. But the latter property holds because A_t^- is open (since ϕ and ψ are continuous on the interiors of their domains) and thus σ-compact, while $\partial\phi$ is closed (since ϕ is lower semicontinuous) and thus also σ-compact;[1] the intersection of two σ-compact sets is again σ-compact.

From the assumption that γ_ψ and γ_ϕ are zero-couplings for (μ, ν), Theorem 2.9 and $0 \notin B$, we get

$$\begin{aligned} \nu(B) &= \nabla\psi_\#\mu(B) + \gamma_\psi(\{0\} \times B) \\ &= \nabla\phi_\#\mu(B) + \gamma_\phi(\{0\} \times B). \end{aligned} \tag{27}$$

We now show that

$$\gamma_\psi(\{0\} \times B) = 0. \tag{28}$$

It is enough to show that $(\{0\} \times B) \cap \partial\psi = \varnothing$ or in other words, that if $(x, y) \in \partial\psi$ is such that $y \in B$, then necessarily $x \neq 0$. Thus, let $(x, y) \in \partial\psi$ be such that $y \in B$. Then we have

$$x \in (\partial\psi)^{-1}(y) \subset (\partial\psi)^{-1}(B) = (\partial\psi)^{-1}(\partial\phi(A_t^-)) \subset (\partial\psi)^{-1}(\partial\phi(M)).$$

Now Theorem 4.8 applied at $p = z$ with M defined in Eq. (25) implies $(\partial\psi)^{-1}(\partial\phi(M)) \subset M$, so that $x \in M$. From the definition of M in Eq. (25) and $t \neq 0$, we have $0 \notin M$, and we obtain $x \neq 0$, proving Eq. (28).

Recall from Step 1 that we have $\mu(M \setminus A_t^-) \leqslant \mu(\mathbb{R}^d_{\setminus 0} \setminus \mathcal{D}) = 0$. From Eq. (28) we therefore get

$$\begin{aligned} \nu(B) &= (\nabla\psi)_\#\mu(B) \\ &= \mu[(\nabla\psi)^{-1}(\partial\phi(A_t^-))] && \text{(as } B = \partial\phi(A_t^-) \text{ by definition)} \\ &\leqslant \mu[M \setminus \mathbb{B}(z, \rho)] && \text{(by Lemma 4.8)} \\ &= \mu[A_t^- \setminus \mathbb{B}(z, \rho)] && \text{(since } \mu(M \setminus A_t^-) = 0) \\ &< \mu(A_t^-) < \infty, && \text{(by Steps 4 and 5).} \end{aligned} \tag{29}$$

However, it also holds that

$$\begin{aligned} \nu(B) &\geqslant (\nabla\phi)_\#\mu(B) \\ &= \mu[(\nabla\phi)^{-1}(\partial\phi(A_t^-))] \\ &= \mu(A_t^- \cap \mathrm{dom}\,\nabla\phi) \\ &\geqslant \mu(A_t^-). \end{aligned} \tag{30}$$

Together, Eqs. (29) and (30) provide the desired contradiction. As a consequence, the inequality in Eq. (20) must be false. The proof of the identities $\nabla\psi = \nabla\phi$ μ-almost everywhere and $\gamma_\psi = \gamma_\phi$ is complete.

Proof of the particular case

[1] If $U \subset \mathbb{R}^d$ is open, then $U = \bigcup_{n\geqslant 1} K_n$ with compact $K_n = \{x \in \mathbb{R}^d \mid \|x\| \leqslant n, \mathbb{B}(x, 1/n) \subset U\}$. If $F \subset \mathbb{R}^d$ is closed, then $F = \bigcup_{n\geqslant 1} K_n$ with compact $K_n = \{x \in F \mid \|x\| \leqslant n\}$.

By Theorem 2.8, $\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ is nonempty. Let $\gamma_1,\gamma_2 \in \Gamma_{0,\mathrm{cm}}(\mu,\nu)$. Applying Rockafellar's Theorem A.4 we get two closed convex functions $\psi_1,\psi_2 : \mathbb{R}^d \to \mathbb{R}\cup\{\infty\}$ such that $\operatorname{spt}\gamma_i \subset \partial\psi_i$ for i in $\{1,2\}$. When at least one of μ or ν has infinite mass, Lemma B.7 implies that the assumption $(0,0)\in \operatorname{spt}\gamma_1 \cap \operatorname{spt}\gamma_2$ is satisfied. The conclusion of Step 7 above yields $\gamma_1=\gamma_2$, whence $\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ is a singleton. □

Proof of Theorem 4.4. Since $\nu = \operatorname{res}((\nabla\psi_j)_{\#}\mu)$, $j=1,2$, we must have $\mu(\mathbb{R}^d_{\not 0}) \geqslant \nu(\mathbb{R}^d_{\not 0})$. If μ and thus ν are the zero measure, the result is trivial.

Suppose μ has finite, nonzero mass. We consider their extensions $\tilde\mu \in [\mu]$ and $\bar\nu \in [\nu]$ to $\mathbb{R}^d$ with $\tilde\mu(0)=0$ and $\bar\nu(\{0\}) = \mu(\mathbb{R}^d_{\not 0}) - \nu(\mathbb{R}^d_{\not 0})$. The measures $\bar\nu$ and $\tilde\mu$ verify

$$\bar\nu(\mathbb{R}^d) = \bar\nu(\{0\}) + \nu(\mathbb{R}^d_{\not 0}) = \mu(\mathbb{R}^d_{\not 0}) = \tilde\mu(\mathbb{R}^d) := m < \infty.$$

Thus $\nu_1 = \bar\nu/m$ and $\mu_1 = \tilde\mu/m$ are probability measures. We check that for $\psi\in\{\psi_1,\psi_2\}$, the gradient $\nabla\psi$ pushes μ_1 forward to ν_1 in the traditional sense. This follows from the following identities: for any Borel set $B\subset\mathbb{R}^d_{\not 0}$, we have

$$\nu_1(B) = m^{-1}(\nabla\psi)_{\#}\mu(B) = \mu_1\left((\nabla\psi)^{-1}(B)\right) = (\nabla\psi)_{\#}\mu_1(B)$$

while the equality $\mu(\mathbb{R}^d_{\not 0}\setminus \operatorname{dom}\nabla\psi)=0$ implies $(\nabla\psi)_{\#}\mu(\mathbb{R}^d) = \mu(\mathbb{R}^d_{\not 0}\cap\operatorname{dom}\nabla\psi) = \mu(\mathbb{R}^d_{\not 0})$ and thus

$$\begin{aligned}\nu_1(\{0\}) &= m^{-1}\left(\mu(\mathbb{R}^d_{\not 0}) - \nu(\mathbb{R}^d_{\not 0})\right)\\ &= m^{-1}\left((\nabla\psi)_{\#}\mu(\mathbb{R}^d) - (\nabla\psi)_{\#}\mu(\mathbb{R}^d_{\not 0})\right)\\ &= m^{-1}(\nabla\psi)_{\#}\mu(\{0\}) = (\nabla\psi)_{\#}\mu_1(\{0\}).\end{aligned}$$

Applying the uniqueness part of the Main Theorem in McCann [25] yields $\nabla\psi_1 = \nabla\psi_2$, μ_1-almost everywhere, and thus also μ-almost everywhere.

Assume now that μ has infinite mass. By Theorem 4.1, it is sufficient to show that for $\psi\in\{\psi_1,\psi_2\}$, the measure $\gamma := \operatorname{res}((\mathrm{Id}\times\nabla\psi)_{\#}\mu)$ is a zero-coupling of μ and ν and that $\operatorname{spt}\gamma\subset\partial\psi$, so that $\gamma\in\Gamma_{0,\mathrm{cm}}(\mu,\nu)$. In view of Theorem B.7, we then necessarily have $(0,0)\in\operatorname{spt}\gamma$ and thus $(0,0)\in\partial\psi$, a condition in Theorem 4.1.

First, γ is a zero-coupling of μ and ν, since for $A\in\mathcal{B}(\mathbb{R}^d_{\not 0})$, we have

$$\begin{aligned}\gamma(A\times\mathbb{R}^d) &= \mu\left(A\cap(\nabla\psi)^{-1}(\mathbb{R}^d)\right) = \mu\left(A\cap\operatorname{dom}\nabla\psi\right) = \mu(A),\\ \gamma(\mathbb{R}^d\times A) &= \mu\left(\mathbb{R}^d_{\not 0}\cap(\nabla\psi)^{-1}(A)\right) = \nu(A).\end{aligned}$$

Second, to prove that $\operatorname{spt}\gamma\subset\partial\psi$, we show the equivalent statement that $(x,y)\notin\partial\psi$ implies $(x,y)\notin\operatorname{spt}\gamma$. Let $(x,y)\notin\partial\psi$. Since $\partial\psi$ is closed, there exist open sets V_x,V_y in $\mathbb{R}^d$ such that $x\in V_x$, $y\in V_y$ and $(V_x\times V_y)\cap\partial\psi=\varnothing$. By definition, $\gamma(V_x\times V_y) = \tilde\mu\left(\left\{\tilde x\in V_x\cap\operatorname{dom}\nabla\psi \mid \nabla\psi(\tilde x)\in V_y\right\}\right)$. However, from $\partial\psi\cap(V_x\times V_y)=\varnothing$ we get

$$\left\{\tilde x\in V_x\cap\operatorname{dom}\nabla\psi \mid \nabla\psi(\tilde x)\in V_y\right\} \subset \left\{\tilde x\in V_x \mid \partial\psi(\tilde x)\cap V_y\neq\varnothing\right\} = \varnothing.$$

This shows that $\gamma(V_x\times V_y)=0$, whence $(x,y)\notin\operatorname{spt}\gamma$ as desired. □

Proof of Theorem 4.5. By Theorem 2.8, the set $\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ is not empty, and since at least one of the measures μ and ν has infinite mass, it is a singleton by Theorem 4.1. Theorem 2.9 then yields the desired representations of the unique zero-coupling γ and, in the proper case, of ν. □

5. Tails of transport plans between regularly varying distributions

In probability theory, regular variation underpins asymptotic theory for sample maxima and sums [8,13,28–30] as it provides a coherent framework to describe the behaviour of functions at infinity and thus for the tails of random variables and vectors. A Borel measurable function f defined in a neighbourhood of infinity and taking positive values is regularly varying with index $\tau \in \mathbb{R}$ if $f(\lambda r)/f(r) \to \lambda^\tau$ as $r \to \infty$ for all $\lambda \in (0,\infty)$. A random variable R is said to have a regularly varying upper tail with index $\alpha \in (0,\infty)$ if the function $r \mapsto \mathsf{P}(R > r)$ is regularly varying with index $-\alpha$. The function $t \mapsto b(t) = Q(1-1/t)$, with Q the quantile function of R, is then regularly varying at infinity with index $1/\alpha$, and we have $t\,\mathsf{P}[R/b(t) > \lambda] \to \lambda^{-\alpha}$ as $t \to \infty$ for all $\lambda \in (0,\infty)$. The latter statement means that, in the space $\mathrm{M}_0(\mathbb{R})$, we have $t\,\mathsf{P}[R/b(t) \in \cdot\,] \xrightarrow{0} \nu_\alpha$ as $t \to \infty$ where ν_α is concentrated on $(0,\infty)$ and is determined by $\nu_\alpha((\lambda,\infty)) = \lambda^{-\alpha}$ for all $\lambda \in (0,\infty)$.

More generally, a probability measure $P \in \mathcal{P}(\mathbb{R}^d)$ is regularly varying (see Hult and Lindskog [22, Definition 3.2, Theorem 3.1], Lindskog et al. [23, Definition 3.2, Theorem 3.1], and Resnick [30, Definition 2.1]) if there exists an increasing Borel measurable function b defined in a neighbourhood of infinity and taking positive values such that

$$t\,P(b(t)\,\cdot\,) \xrightarrow{0} \mu, \qquad t \to \infty \tag{31}$$

for some non-zero $\mu \in \mathrm{M}_0(\mathbb{R}^d)$, where $\lambda A = \{\lambda x \mid x \in A\}$ for a subset A of Euclidean speace. The function b must be regularly varying with index $1/\alpha \in (0,\infty)$, say, and the limit measure ν must be homogeneous:

$$\mu(\lambda^{-1/\alpha}\,\cdot\,) = \lambda\,\mu(\,\cdot\,), \qquad \lambda > 0. \tag{32}$$

As a consequence, its support is a (closed) multiplicative cone, that is, $x \in \operatorname{spt}\mu$ and $\lambda \in [0,\infty)$ imply $\lambda x \in \operatorname{spt}\mu$. We call $\alpha \in (0,\infty)$ the index of regular variation of P.

Let $P, Q \in \mathcal{P}(\mathbb{R}^d)$ be regularly varying probability measures with auxiliary functions b_1, b_2 and limit measures $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$, respectively. By Theorem 6 of McCann [25] and by Rockafellar's Theorem A.4, there exists a coupling measure $\pi \in \Pi(P,Q)$ with cyclically monotone support contained in the graph of the subdifferential $\partial\psi$ of some closed convex function ψ on $\mathbb{R}^d$. By Theorem 2.8 and Rockafellar's Theorem A.4, there exists a zero-coupling measure $\gamma \in \Gamma_{0,\mathrm{cm}}(\mu,\nu)$ with cyclically monotone support contained in the subdifferential $\partial\bar\psi$ of some closed convex function $\bar\psi$ on $\mathbb{R}^d$. We will connect the two subdifferentials $\partial\psi$ and $\partial\bar\psi$ via the asymptotic behaviour of $\partial\psi(x)$ when $|x|$ tends to infinity. In view of our aim of providing notions for measures in $\mathrm{M}_0(\mathbb{R}^d)$ that are similar to the center-outward distribution and quantile functions defined in Hallin et al. [19], we believe this connection will be of interest in future work on tail quantile contours.

Write $1_d = (1,\dots,1) \in \mathbb{R}^d$ and let $\operatorname{diag}(a1_d, b1_d)$ be the $(2d)\times(2d)$ diagonal matrix with diagonal $(a,\dots,a,b,\dots,b)$, with d repetitions of $a, b \in \mathbb{R}$. For a diagonal matrix $A = \operatorname{diag}(a_1,\dots,a_d)$ and a scalar $\lambda > 0$, we write $\lambda^A = \exp(A\log(\lambda)) = \operatorname{diag}(\lambda^{a_1},\dots,\lambda^{a_d})$ with $\exp(\,\cdot\,)$ the matrix exponential. The following theorem was summarized in the commutative diagram in Fig. 1 and establishes the link between cyclically monotone couplings and transport plans between regularly varying probability measures on the one hand and those between their exponent measures on the other hand. Recall the Fell topology described in the beginning of Section 2.2 and recall the notation $T \llcorner V = T \cap (V \times \mathbb{R}^d)$ for $T \subset \mathbb{R}^d \times \mathbb{R}^d$ and $V \subset \mathbb{R}^d$. For the subdifferential $\partial\psi$ of a convex function ψ and a scalar $b > 0$, we write $\partial\psi(b\,\cdot\,) = \{(x,y) \mid y \in \partial\psi(bx)\}$.

Theorem 5.1 (Regular variation and monotone transport). *Let P and Q in $\mathcal{P}(\mathbb{R}^d)$ be regularly varying probability measures with auxiliary functions b_1 and b_2, indices $\alpha_1 > 0$ and $\alpha_2 > 0$, and limit*

measures μ and ν in $\mathrm{M}_0(\mathbb{R}^d)$, respectively. Let $\pi \in \Pi_{\mathrm{cm}}(P,Q)$ with $\operatorname{spt}\pi \subset \partial\psi$ for some closed convex function $\psi : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$. If we further assume that

(i) *μ vanishes on sets of Hausdorff dimension not larger than $d-1$,*
(ii) *$\forall y \in \operatorname{spt}\nu \setminus \{0\} : \exists x \in \operatorname{spt}\mu : \langle x, y\rangle > 0$,*

then we have the following properties:

(a) *$\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ is a singleton with single element $\gamma = (\mathrm{Id}\times\nabla\bar\psi)_\#\mu$ for some closed convex function $\bar\psi$ satisfying $\operatorname{spt}\gamma \subset \partial\bar\psi$ and $\bar\psi(0) = \inf_{x\in\mathbb{R}^d}\bar\psi(x) < \infty$. Moreover, γ is proper, so that $\gamma = \operatorname{res}\big((\mathrm{Id}\times\nabla\bar\psi)_\#\mu\big)$ and $\nu = \operatorname{res}\big((\nabla\bar\psi)_\#\mu\big)$.*
(b) *Writing $V = \operatorname{int}(\operatorname{spt}\mu)$ and $B(t) = \operatorname{diag}(b_1(t)1_d, b_2(t)1_d)$ for $t > 0$, we have, as $t \to \infty$,*

$$t\,\operatorname{res}\pi(B(t)\cdot) \xrightarrow{0} \gamma \qquad \text{in } \mathrm{M}_0(\mathbb{R}^d\times\mathbb{R}^d), \tag{33}$$

$$(B(t)^{-1}\partial\psi)\llcorner V \xrightarrow{\mathcal{F}} \big(\partial\bar\psi\big)\llcorner V \qquad \text{in } \mathcal{F}(V\times\mathbb{R}^d). \tag{34}$$

(c) *The measure γ is homogeneous in the sense that, writing $E = \operatorname{diag}(\alpha_1^{-1}1_d, \alpha_2^{-1}1_d)$, for all $\lambda > 0$, we have*

$$\gamma(\lambda^{-E}\,\cdot) = \lambda\gamma, \tag{35}$$

$$\lambda^E \operatorname{spt}\gamma = \operatorname{spt}\gamma. \tag{36}$$

(d) *The gradient $\nabla\bar\psi$ is defined and determined uniquely μ-almost everywhere. We have*

$$\nabla\bar\psi(\lambda x) = \lambda^{\alpha_1/\alpha_2}\,\nabla\bar\psi(x), \qquad \lambda > 0, \tag{37}$$

for μ-almost every $x \in \mathbb{R}^d$. Moreover, the function $\bar\psi$ may be modified such that $\bar\psi(0) = 0$; then it is determined uniquely μ–almost everywhere and satisfies

$$\bar\psi(\lambda x) = \lambda^{\alpha_1/\alpha_2+1}\,\bar\psi(x), \qquad \lambda \geqslant 0, \tag{38}$$

for μ-almost every $x \in \mathbb{R}^d$.

In Eq. (34), the set $B(t)^{-1}\partial\psi$ is the subdifferential of the convex function

$$\psi_{b_1(t),b_2(t)}(x) = (b_1(t)b_2(t))^{-1}\psi(b_1(t)x), \qquad x \in \mathbb{R}^d, \tag{39}$$

the subdifferential of which satisfies

$$\partial\psi_{b_1(t),b_2(t)}(x) = b_2(t)^{-1}\partial\psi(b_1(t)x), \qquad x \in \mathbb{R}^d,$$

by Theorem 5.3 below. We can therefore view Eq. (34) as the convergence of a rescaled version of the subdifferential of ψ, seen as a multivalued map.

If $b_1 = b_2$ and thus $\alpha_1 = \alpha_2$, then the coupling distribution π is regularly varying with the same auxiliary function and index, and with limit measure γ. In that case, the limit measure of the cyclically monotone coupling π between P and Q is equal to the cyclically monotone coupling γ between their limit measures μ and ν. This is the gist of the commutative diagram in Fig. 1. If the two auxiliary functions b_1 and b_2 of P and Q cannot be taken to be the same, for instance because the indices α_1 and α_2 are different, then the coupling distribution π is not regularly varying in the sense of Eq. (31), but we can still think of the convergence relation in Eq. (33) as a kind of extended form of regular variation.

Remark 5.2. The assumptions guaranteeing the cyclically monotone coupling γ to be proper are satisfied as soon as the reference limit measure μ is smooth enough (absolute continuity suffices) and has large enough support. As a consequence, they come at no cost for practical use as one can easily define a reference distribution satisfying them. For example, the reference distribution may be chosen to have a polar decomposition made up of the uniform distribution on the unit sphere for the angular component and a Fréchet distribution for the radial part. If the interest is in maxima only, then we may restrain the angular component to the positive orthant.

In Theorem 5.1, the uniqueness in (a) follows from the smoothness of μ in (i) and Theorem 4.1. The fact that γ is proper and thus γ and ν can be written as push-forwards of μ follows from (ii) and Proposition 3.4 together with homogeneity properties which are the hallmark of regular variation. The same homogeneity properties also explain (c) and (d). Finally, the stability in (b) can be linked back to Theorem 2.6.

Since the functions b_1 and b_2 diverge at infinity, the Fell convergence in Eq. (34) implies that the tail of $\partial\psi$ is asymptotically homogeneous and converges to $\partial\bar{\psi}$. Under additional assumptions, this Fell convergence implies locally uniform convergence if we think of the subdifferentials as multi-valued mappings [32, Theorem 5.44].

The limit measure ν captures the multivariate tail of Q via $Q(A) \approx t^{-1}\nu(b_2(t)^{-1}(A)) = t^{-1}b_2(t)^{\alpha}\nu(A)$ for Borel sets A far away from the origin. If μ is chosen to be spherically symmetric, then quantile or depth contours, however defined, are just spheres, and transporting them via $\nabla\bar{\psi}$ gives the corresponding contours for the target measure ν and thus for the tail quantile contours of Q; the contours are all homothetic to a fixed one by the homogeneity of $\nabla\bar{\psi}$. This provides a basis for the study of tail versions of the center-outward quantiles in [7,18,19].

Proofs

Lemma 5.3 (Scaling). *Let $P, Q \in \mathcal{P}(\mathbb{R}^d)$ and let $\pi \in \Pi(P,Q)$ have support contained in the subdifferential $\partial\psi$ of some closed convex function ψ. For scalar $b_1 > 0$ and $b_2 > 0$, consider the $2d \times 2d$ matrix $B = \operatorname{diag}(b_1 1_d, b_2 1_d)$ and the measures $P_{b_1} = P(b_1\,\cdot)$, $Q_{b_2} = Q(b_2\,\cdot)$ and $\pi_{b_1,b_2} = \pi(B\,\cdot)$. Then $\pi_{b_1,b_2} \in \Pi(P_{b_1}, Q_{b_2})$ and its support $\operatorname{spt}\pi_{b_1,b_2} = B^{-1}\operatorname{spt}\pi$ is contained in the subdifferential of the closed convex function $\psi_{b_1,b_2} = (b_1 b_2)^{-1}\psi(b_1\,\cdot)$, the subdifferential of which is $\partial\psi_{b_1,b_2}(x) = b_2^{-1}\partial\psi(b_1 x)$ for all $x \in \mathbb{R}^d$, or $\partial\psi_{b_1,b_2} = B^{-1}\partial\psi$ as subsets of $\mathbb{R}^d \times \mathbb{R}^d$.*

Proof. The function ψ_{b_1,b_2} is closed and convex since this is true for ψ. We have $(x,y) \in \partial\psi_{b_1,b_2}$ if and only if

$$\forall z \in \mathbb{R}^d, \qquad (b_1 b_2)^{-1}\psi(b_1 z) \geqslant (b_1 b_2)^{-1}\psi(b_1 x) + \langle z - x, y\rangle,$$

which is equivalent to

$$\forall s \in \mathbb{R}^d, \qquad \psi(s) \geqslant \psi(b_1 x) + \langle s - b_1 x, b_2 y\rangle$$

and thus to $(b_1 x, b_2 y) \in \partial\psi$. We find $\partial\psi_{b_1,b_2} = B^{-1}\partial\psi$ and $\partial\psi_{b_1,b_2}(x) = b_2^{-1}\partial\psi(b_1 x)$ for $x \in \mathbb{R}^d$. Clearly, $\operatorname{spt}\pi_{b_1,b_2} = B^{-1}\operatorname{spt}\pi \subset B^{-1}\partial\psi = \partial\psi_{b_1,b_2}$. □

Lemma 5.4. *Let P and Q in $\mathcal{P}(\mathbb{R}^d)$ be regularly varying with auxiliary functions b_1 and b_2, indices $\alpha_1 > 0$ and $\alpha_2 > 0$, and limit measures μ and ν in $\mathrm{M}_0(\mathbb{R}^d)$, respectively. Let ψ be a closed convex function such $\partial\psi$ contains the cyclically monotone support of some $\pi \in \Pi_{\mathrm{cm}}(P,Q)$. Every sequence of*

positive numbers $t_n \to \infty$ then contains a subsequence such that, as $n \to \infty$ along the subsequence, we have writing $B(t) = \operatorname{diag}(b_1(t)1_d, b_2(t)1_d)$,

$$\begin{aligned} t_n \operatorname{res} \pi(B(t_n)\,\cdot\,) &\xrightarrow{0} \gamma && \text{in } \mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d), \\ B(t_n)^{-1}\partial\psi &\xrightarrow{\mathcal{F}} \partial\bar{\psi} && \text{in } \mathcal{F}(\mathbb{R}^d \times \mathbb{R}^d), \end{aligned} \tag{40}$$

for some $\gamma \in \Gamma_{0,\mathrm{cm}}(\mu,\nu)$ and some closed convex function $\bar{\psi}$ satisfying $\operatorname{spt}\gamma \subset \partial\bar{\psi}$ and $\bar{\psi}(0) = \inf_{x\in\mathbb{R}^d}\bar{\psi}(x) < \infty$. If $\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ is a singleton, say $\{\gamma\}$, then we actually have $t \operatorname{res}\pi(B(t)\,\cdot\,) \xrightarrow{0} \gamma$ as $t \to \infty$, and, with $E = \operatorname{diag}(\alpha_1^{-1}1_d, \alpha_2^{-1}1_d)$, the measure γ satisfies

$$\forall \lambda > 0, \qquad \gamma(\lambda^{-E}\,\cdot\,) = \lambda\gamma(\,\cdot\,); \tag{41}$$

in particular, its support satisfies $\lambda^E \operatorname{spt}\gamma = \operatorname{spt}\gamma$ for all $\lambda > 0$.

Proof of Lemma 5.4. Choose a sequence $1 \leqslant t_n \to \infty$ and consider the Borel measures $\bar{\mu}_n = t_n P(b_1(t_n)\,\cdot\,)$ and $\bar{\nu}_n = t_n Q(b_2(t_n)\,\cdot\,)$, both with mass t_n. In view of Lemma 5.3, the function $\psi_n = (b_2(t_n)b_1(t_n))^{-1}\psi(b_1(t_n)\,\cdot\,)$ is closed and convex, and its subdifferential $\partial\psi_n = B(t_n)^{-1}\partial\psi$ contains the support of the measure $\pi_n = t_n\,\pi(B(t_n)\,\cdot\,)$ that couples $\bar{\mu}_n$ and $\bar{\nu}_n$.

By regular variation, $\operatorname{res}\bar{\mu}_n \xrightarrow{0} \mu$ and $\operatorname{res}\bar{\nu}_n \xrightarrow{0} \nu$ in $\mathrm{M}_0(\mathbb{R}^d)$ as $n \to \infty$, with μ and ν not depending on the sequence t_n. Moreover, for every n, the measure $\gamma_n = \operatorname{res}\pi_n$ is a zero-coupling between $\mu_n = \operatorname{res}\bar{\mu}_n$ and $\nu_n = \operatorname{res}\bar{\nu}_n$. By Theorem 2.6, there exists a subsequence of t_n, relabelled as t_n again, that satisfies $\gamma_n \xrightarrow{0} \gamma \in \Gamma_{0,\mathrm{cm}}(\mu,\nu)$ in $\mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$ and $\partial\psi_n \xrightarrow{\mathcal{F}} \partial\bar{\psi}$ as $n \to \infty$ along the subsequence, with $\bar{\psi}$ a closed convex function satisfying $\operatorname{spt}\gamma \subset \operatorname{gph}\partial\bar{\psi}$, the existence of which is a consequence of Rockafellar's Theorem A.4.

Since μ and ν have infinite mass and since their zero-coupling γ is supported by $\partial\bar{\psi}$, Theorem B.7 implies that $\bar{\psi}(0)$ is finite.

Suppose that $\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ is a singleton, say $\{\gamma\}$. The convergence in Eq. (40) along some subsequence of an arbitrary sequence $t_n \to \infty$ implies $t \operatorname{res}\pi(B(t)\,\cdot\,) \xrightarrow{0} \gamma$ as $t \to \infty$; otherwise, we could find a sequence $t_n \to \infty$ along which the M_0-convergence to γ would not take place. The matrix $B(t)$ is invertible for every $t > 0$ in an appropriate neighbourhood of infinity and we have $B(\lambda t)^{-1}B(t) \to \lambda^{-E}$ as $t \to \infty$ by regular variation of b_1 and b_2. For $\lambda \in (0,\infty)$, we have, on the one hand,

$$\lambda t \operatorname{res}\pi(B(t)\,\cdot\,) \xrightarrow{0} \lambda\gamma, \qquad t \to \infty,$$

and on the other hand,

$$\lambda t \operatorname{res}\pi(B(t)\,\cdot\,) = \lambda t \operatorname{res}\pi(B(\lambda t)B(\lambda t)^{-1}B(t)\,\cdot\,) \xrightarrow{0} \gamma(\lambda^{-E}\,\cdot\,), \qquad t \to \infty, \tag{42}$$

from which Eq. (41) follows. To see Eq. (42), note that $\pi(B(t)\,\cdot\,) = (B(t)^{-1})_{\#}\pi$ and that, for any bounded and continuous function $f : \mathbb{R}^{2d}_{\setminus 0} \to \mathbb{R}$ that vanishes on a neighbourhood of the origin, we have

$$\begin{aligned} \lambda t\,(\operatorname{res}\pi(B(t)\,\cdot\,))\,(f) &= \lambda t \int_{\mathbb{R}^2_{\setminus 0}} f(B(t)^{-1}x)\,\mathrm{d}\pi(x) \\ &= \lambda t \int_{\mathbb{R}^2_{\setminus 0}} f(B(\lambda t)^{-1}B(\lambda t)B(t)^{-1}x)\mathrm{d}\pi(x) \\ &= \lambda t\,(\operatorname{res}\pi(B(\lambda t)\,\cdot\,))\,(f_{\lambda,t}), \end{aligned}$$

where $f_{\lambda,t}(x) := f(B(\lambda t)B(t)^{-1}x) \to f_{\lambda,\infty}(x) = f(\lambda^E x)$ as $t \to \infty$ for all $x \in \mathbb{R}^{2d}_{\setminus 0}$. Since the latter convergence takes place uniformly on compact subsets of $\mathbb{R}^{2d}_{\setminus 0}$ and since the functions $f_{\lambda,t}$ for large t and f_λ all vanish on a common neighourhood of the origin, we find by an extension of the mapping theorem in Hult and Lindskog [22, Theorem 2.5] that

$$\lambda t\,(\operatorname{res}\pi(B(t)\cdot))\,(f) \to \gamma(f_{\lambda,\infty}) = \gamma(f(\lambda^E\cdot)) = (\gamma(\lambda^{-E}\cdot))(f), \qquad t \to \infty,$$

as required. □

Proof of Theorem 5.1. (a) Since μ and ν are limit measures in the definition of regular variation, they are nonzero, and the homogeneity in Eq. (32) then implies $\mu(\mathbb{R}^d_{\setminus 0}) = \nu(\mathbb{R}^d_{\setminus 0}) = \infty$. By assumption (i) and Theorem 4.5, it follows that $\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ is a singleton, $\{\gamma\}$, say. Theorem 5.4 then implies that $\operatorname{spt}\gamma$ satisfies the homogeneity assumption in Theorem 3.4, which, in combination with assumption (ii), implies that γ is proper. The existence and properties of $\bar\psi$ follow from the same Theorem 5.4 too. Finally, the expressions for γ and ν in terms of $\nabla\bar\psi$ and μ follow from Theorem 4.5.

(b) The M_0-convergence in Eq. (33) is clear from (a) and Theorem 5.4. By Theorem 5.3, $\operatorname{spt}(\operatorname{res}\pi(B(t)\cdot))$ is a subset of $B(t)^{-1}\partial\psi$, which is the subdifferential of the function $\psi_{b_1(t),b_2(t)}$ in Eq. (39). Since $\Gamma_{0,\mathrm{cm}}(\mu,\nu)$ is a singleton from (a), and as $\operatorname{res}\pi(B(t)\cdot) \xrightarrow{0} \gamma$ as $t \to \infty$, we can apply Theorem 2.7 to any sequence $0 < t_n \to \infty$. Again from the proof of Lemma 5.4, we have $\operatorname{spt}\gamma \subset \partial\bar\psi$ for $\bar\psi$ from (a). It follows that in $\mathcal{F}(V \times \mathbb{R}^d)$, we have

$$B(t_n)^{-1}\partial\psi\llcorner V \xrightarrow{\mathcal{F}} \partial\bar\psi\llcorner V, \qquad n \to \infty.$$

Since this is true for every sequence $t_n \to \infty$, the convergence in Eq. (34) follows.

(c) The two properties of γ are already proven in Lemma 5.4.

(d) Define $T = \operatorname{spt}\gamma$ and $T(x) = \{y \in \mathbb{R}^d \mid (x,y) \in T\}$ for $x \in \mathbb{R}^d$. Further, put $\operatorname{dom}T = \{x \in \mathbb{R}^d \mid T(x) \neq \varnothing\} = \operatorname{proj}_1(T)$. By Eq. (36), we have $y \in T(x)$ if and only if $\lambda^{1/\alpha_2}y \in T(\lambda^{1/\alpha_1}x)$, so that

$$\forall\lambda > 0, x \in \mathbb{R}^d: \qquad T(\lambda x) = \lambda^{\alpha_1/\alpha_2}T(x), \tag{43}$$

an identity in which one and hence both sets can be empty.

By Eq. (43), the set $\operatorname{dom}T$ is a multiplicative cone, which furthermore contains the origin. Further, note that $\operatorname{dom}T = \operatorname{proj}_1(\operatorname{spt}\gamma)$, so $\mu(\mathbb{R}^d \setminus \operatorname{dom}T) = 0$ and, by Lemma B.4 and Remark B.5, $\operatorname{spt}\mu = \operatorname{cl}(\operatorname{dom}T)$. As $\operatorname{spt}\gamma \subset \partial\bar\psi$, also $\operatorname{dom}T \subset \operatorname{dom}\partial\bar\psi$. Replace $\bar\psi$ by $\bar\psi - \bar\psi(0)$ to ensure $\bar\psi(0) = 0$ without changing the subdifferential or gradient of $\bar\psi$.

Let $x \in \operatorname{dom}T$ and define $f(t) = \bar\psi(tx)$ for $t \geqslant 0$. The convex function f is finite on $[0,\infty)$, as $\operatorname{dom}T \subset \operatorname{dom}\partial\bar\psi$ and $\operatorname{dom}T$ is a multiplicative cone. Its subdifferential $\partial f(t)$ contains the set $\langle\partial\bar\psi(tx), x\rangle = \{\langle y,x\rangle : y \in \partial\bar\psi(tx)\}$ (just check the inequality in the definition of a subdifferential); therefore, it also contains the set $\langle T(tx), x\rangle$. As a consequence, f has a left derivative f'_- and a right derivative f'_+ satisfying $f'_-(t) \leqslant \inf\langle T(tx),x\rangle \leqslant \sup\langle T(tx),x\rangle \leqslant f'_+(t)$ at every $t > 0$. As furthermore

$$\forall t \in (0,\infty): \qquad f(t) = \int_{(0,t)} f'_-(s)\,\mathrm{d}s = \int_{(0,t)} f'_+(s)\,\mathrm{d}s, \tag{44}$$

(e.g. [31], Corollary 24.2.1), Eq. (43) implies

$$\bar\psi(tx) = \int_{(0,t)} \sup\langle T(sx),x\rangle\,\mathrm{d}s = \sup\langle T(x),x\rangle \int_{(0,t)} s^{\alpha_1/\alpha_2}\,\mathrm{d}s. \tag{45}$$

We find that $\bar\psi(x)$ is uniquely determined by T for $x \in \operatorname{dom} T$ and thus μ-almost everywhere, and also that Eq. (38) holds for such x.

As in Lemma A.6, let the function $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ be the closure of the convex hull of the restriction of $\bar\psi$ to $D = \operatorname{dom} T$. Its epigraph is

$$\operatorname{epi}\phi = \operatorname{cl}\left(\operatorname{conv}(\operatorname{epi}\bar\psi \cap (D \times \mathbb{R}))\right).$$

By Lemma A.6(c), we have $\partial\bar\psi(x) \subset \partial\phi(x)$ for all $x \in D$, and therefore

$$\operatorname{spt}\gamma \subset \partial\bar\psi \cap (D \times \mathbb{R}^d) \subset \partial\phi.$$

Theorem 2.9 yields $\mu(\mathbb{R}^d_{\searrow 0} \setminus \operatorname{dom}\nabla\phi) = 0$, that is, $\nabla\phi$ is defined μ-almost everywhere.

From Lemma A.6(b), we get $\phi(x) = \bar\psi(x)$ for $x \in D$. Since D is a multiplicative cone and since $\bar\psi(\lambda x) = \lambda^{\alpha_1/\alpha_2+1}\bar\psi(x)$ for all $x \in D$ and $\lambda \geqslant 0$, we have

$$\phi(\lambda x) = \lambda^{\alpha_1/\alpha_2+1}\phi(x), \qquad \lambda > 0,\ x \in D. \tag{46}$$

Take $x \in D$ and $\lambda > 0$. By Lemma A.6(b)–(c), we have

$$v \in \partial\phi(x) \iff \forall z \in D : \phi(z) \geqslant \phi(x) + \langle v, z - x\rangle.$$

Therefore, as D is a multiplicative cone,

$$\begin{aligned} w \in \partial\phi(\lambda x) &\iff \forall z \in D : \phi(\lambda z) \geqslant \phi(\lambda x) + \langle w, \lambda z - \lambda x\rangle \\ &\iff \forall z \in D : \phi(z) \geqslant \phi(x) + \langle \lambda^{-\alpha_1/\alpha_2} w, z - x\rangle \\ &\iff \lambda^{-\alpha_1/\alpha_2} w \in \partial\phi(x), \end{aligned}$$

where we used Eq. (46) in the second equivalence and Lemma A.6(b)–(c) in the third one. We conclude

$$\partial\phi(\lambda x) = \lambda^{\alpha_1/\alpha_2}\partial\phi(x), \qquad \lambda > 0,\ x \in D.$$

As a consequence, $D \cap \operatorname{dom}\nabla\phi$ is a multiplicative cone and $\nabla\phi(\lambda x) = \lambda^{\alpha_1/\alpha_2}\nabla\phi(x)$ for all $\lambda > 0$ and all $x \in D \cap \operatorname{dom}\nabla\phi$. In view of Lemma A.6(e), Eq. (37) must hold for all $x \in D \cap \operatorname{dom}\nabla\phi$ and therefore μ-almost everywhere. □

6. Discussion

We have studied measure transportation between infinite Borel measures on Euclidean space by cyclically monotone mappings. The measures under consideration are finite on sets bounded away from the origin, and so are their zero-couplings, which have the prescribed margins apart from a possible contribution at the origin. We have shown that a cyclically monotone zero-coupling always exists, and we have identified conditions under which it is unique and can be represented as the restriction to the punctured space of the push-forward of one measure by a cyclically monotone mapping, that is, the gradient of a closed convex function. In addition, we have established stability properties of cyclically monotone zero-couplings with respect to M_0-convergence of measures and Fell convergence of subgradients.

The cost function was implicit in our work through the notion of cyclical monotonicity. In contrast to classical optimal transport theory, we did not assume finiteness of transport costs or moments, and no minimization of an expected cost was required. Allowing the measures to have potentially infinite mass

leads to new phenomena compared to the case of probability measures. In particular, the possible presence of a sink or infinite reservoir of mass at the origin may preclude representations in terms of push-forward measures. This observation motivated the introduction of the notion of proper zero-couplings and the study of conditions ensuring their existence.

We have further connected these results to the theory of regular variation. Specifically, we have shown how cyclically monotone couplings between regularly varying probability measures converge at high levels to cyclically monotone zero-couplings between their limit measures. In this way, the present work provides a rigorous link between monotone transport for regularly varying distributions and monotone transport for their tail measures, thereby laying a theoretical foundation for the use of transport-based methods in multivariate extreme value analysis.

Several directions for further research remain open. One natural extension is the study of transport problems under cyclical monotonicity with respect to more general cost functions beyond the quadratic cost considered here. Another direction concerns the development of statistical methodology based on cyclically monotone transport for infinite measures, including inference and simulation methods for multivariate extremes and related stochastic models such as Lévy processes.

References

[1] Albergo, M. S. and E. Vanden-Eijnden (2022). Building normalizing flows with stochastic interpolants. *arXiv preprint arXiv:2209.15571*.

[2] Alberti, G. and L. Ambrosio (1999). A geometrical approach to monotone functions in $\mathbb{R}^n$. *Mathematische Zeitschrift 230*, 259–316.

[3] Aleksandrov, A. (1942). Existence and uniqueness of a convex surface with a given integral curvature. *C. R. (Doklady) Acad. Sci. URSS(N.S.) 35*, 131–134.

[4] Ambrosio, L., E. Brué, D. Semola, et al. (2021). *Lectures on optimal transport*, Volume 130. Springer.

[5] Anderson, R. D. and V. L. Klee (1952). Convex functions and upper semi-continuous collections. *Duke Math. J. 19*, 349–357.

[6] Beiglböck, M. (2015). Cyclical monotonicity and the ergodic theorem. *Ergodic Theory Dyn. Syst. 35(3)*.

[7] Beirlant, J., S. Buitendag, E. Del Barrio, M. Hallin, and F. Kamper (2020). Center-outward quantiles and the measurement of multivariate risk. *Insurance: Mathematics and Economics 95*, 79–100.

[8] Bingham, N. H., C. M. Goldie, J. L. Teugels, and J. Teugels (1989). *Regular Variation*. Cambridge University Press.

[9] Brenier, Y. (1987). Décomposition polaire et réarrangement monotone des champs de vecteurs. *CR Acad. Sci. Paris Sér. I Math. 305*, 805–808.

[10] Brenier, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics 44*(4), 375–417.

[11] Catalano, M., H. Lavenant, A. Lijoi, and I. Prünster (2021). A Wasserstein index of dependence for random measures. *arXiv preprint arXiv:2109.06646*.

[12] Chernozhukov, V., A. Galichon, M. Hallin, and M. Henry (2017). Monge–kantorovich depth, quantiles, ranks and signs. *The Annals of Statistics 45*, 223–256.

[13] de Haan, L. and A. Ferreira (2006). *Extreme Value Theory: An Introduction*. Springer New York.

[14] De Pascale, L., A. Kausamo, and K. Wyczesany (2023). 60 years of cyclic monotonicity: a survey. *arXiv preprint arXiv:2308.07682*.

[15] de Valk, C. and J. Segers (2019). Tails of optimal transport plans for regularly varying probability measures. *arXiv preprint arXiv:1811.12061v2*.

[16] Figalli, A. and N. Gigli (2010). A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions. *Journal de mathématiques pures et appliquées 94*(2), 107–130.

[17] Guillen, N., C. Mou, and S. Swiech (2019). Coupling Lévy measures and comparison principles for viscosity solutions. *Transactions of the American Mathematical Society 372*(10), 7327–7370.

[18] Hallin, M. (2022). Measure transportation and statistical decision theory. *Annual Review of Statistics and Its Application 9*(1), 401–424.
[19] Hallin, M., E. Del Barrio, J. Cuesta-Albertos, and C. Matrán (2021). Distribution and quantile functions, ranks and signs in dimension d: a measure transportation approach. *Annals of Statistics 49*, 1139–1165.
[20] Huesmann, M. (2016). Optimal transport between random measures. *Ann. Inst. Henri Poincaré Probab. Stat. 52(1)*, 196–232.
[21] Huesmann, M. and K.-T. Sturm (2013). Optimal transport from Lebesgue to Poisson. *The Annals of Probability 41*(4), 2426–2478.
[22] Hult, H. and F. Lindskog (2006). Regular Variation for Measures on Metric Spaces. *Publ. Inst. Math. 80*(94), 121–140.
[23] Lindskog, F., S. I. Resnick, and J. Roy (2014). Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps. *Probability Surveys 11*, 270–314.
[24] Lipman, Y., R. T. Chen, H. Ben-Hamu, M. Nickel, and M. Le (2022). Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*.
[25] McCann, R. J. (1995). Existence and uniqueness of monotone measure-preserving maps. *Duke Mathematical Journal 80*(2), 309–323.
[26] Molchanov, I. (2017). *Theory of Random Sets*. Springer Cham.
[27] Pratelli, A. (2008). On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Mathematische Zeitschrift 258*(3), 677–690.
[28] Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Science & Business Media.
[29] Resnick, S. I. (2008). *Extreme Values, Regular Variation and Point Processes*. Springer Science & Business Media.
[30] Resnick, S. I. (2024). *The art of finding hidden risks: Hidden regular variation in the 21st century*. Springer Nature.
[31] Rockafellar, R. T. (1970). *Convex Analysis*. Princeton: Princeton University Press.
[32] Rockafellar, R. T. and R. J.-B. Wets (1998). *Variational Analysis*. New York : Springer.
[33] Rüschendorf, L. and S. T. Rachev (1990). A characterization of random variables with minimum L^2-distance. *Journal of Multivariate Analysis 32*(1), 48–54.
[34] Santambrogio, F. (2015). *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*. Cham: Birkhäuser/Springer.
[35] Schachermayer, W. and J. Teichmann (2009). Characterization of optimal transport plans for the Monge–Kantorovich problem. *Proceedings of the American Mathematical Society 137*(2), 519–529.
[36] Segers, J. (2022). Graphical and uniform consistency of estimated optimal transport plans. *arXiv preprint arXiv:2208.02508*.
[37] Villani, C. (2009). *Optimal Transport: Old and New*, Volume 338. Springer.

Appendix A: Background

A.1. Portmanteau and Prohorov theorem for infinite measures

The space $\mathrm{M}_0(\mathbb{R}^d)$ enjoys useful properties similar to the classical Portmanteau theorem or Prohorov theorem for probability measures which will be of constant use. We refer to [22,23] for details and quote here the results we need.

Theorem A.1 (Portmanteau Theorem 2.4 in [22]). *Let $\mu, \mu_n \in \mathrm{M}_0(\mathbb{R}^d), n \geqslant 1$. The following statements are equivalent.*

(i) $\mu_n \xrightarrow{0} \mu$ *as* $n \to \infty$,

(ii) $\mu_n(A) \to \mu(A)$ *as* $n \to \infty$ *for every* $A \in \mathcal{S}_0$ *satisfying* $\mu(\mathrm{bnd}\, A) = 0$ *and* $\mathrm{cl}(A) \cap \{0\} = \varnothing$ *(closure in* $\mathbb{R}^d$ *as a whole),*

(iii) $\mu(\mathrm{int}(A)) \leqslant \liminf \mu_n(\mathrm{int}(A) \leqslant \limsup \mu_n(\mathrm{cl}(A)) \leqslant \mu(\mathrm{cl}(A))$ *for every* $A \in \mathbb{R}^d_{\searrow 0}$ *such that* $\mathrm{cl}(A) \cap \{0\} = \varnothing$ *(closure in* $\mathbb{R}^d$ *as a whole).*

Theorem A.2 (Prohorov Theorem 2.7 in [22]). *The set* $M \subset \mathrm{M}_0(\mathbb{R}^d)$ *is relatively compact if and only if there exists a sequence* $r_i > 0, i \geqslant 1$ *satisfying* $r_i \downarrow 0$ *as* $i \to \infty$ *such that for each* $i \geqslant 1$ *we have*

$$\sup_{\mu \in M} \mu(\mathbb{R}^d \setminus \mathbb{B}_{0,r_i}) < \infty$$

and for each $\eta > 0$ *there exists a compact set* $K_i \subset \mathbb{R}^d \setminus \mathbb{B}_{0,r_i}$ *such that*

$$\sup_{\mu \in M} \mu\big(\mathbb{R}^d \setminus (\mathbb{B}_{0,r_i} \cup K_i)\big) < \eta.$$

Theorem A.2 gives a path to prove the M_0-convergence of μ_n to μ by showing that (i) the set $\{\mu_n\}_{n \geqslant 1}$ is relatively compact and (ii) all its accumulation points are equal to μ.

A.2. Convex functions and cyclical monotonicity

In this subsection we recall convex functions and cyclical monotonicity, which are connected to optimal transport.

A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is said to be convex if the inequality

$$f\left((1-\tau)x + \tau y\right) \leqslant (1-\tau)\, f(x) + \tau f(y)$$

holds for every $x, y \in \mathbb{R}^d$ and $\tau \in (0,1)$; equivalently, f is a convex function if its epigraph $\mathrm{epi}\, f = \big\{(x,t) \in \mathbb{R}^d \times \mathbb{R} \mid f(x) \leqslant t\big\}$ is a convex set. The convex function f is proper if $f(x) < +\infty$ for some x in $\mathbb{R}^d$ and it is closed if its epigraph is closed, or equivalently, if f is lower semi-continuous. The effective domain of f is the set $\mathrm{dom}\, f = \{x \in \mathbb{R}^d \mid f(x) \in \mathbb{R}\}$.

For any proper, convex function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ and any point $\bar{x}$ lying in $\mathrm{dom}\, f$, we call subgradient of f at $\bar{x}$ any point $v \in \mathbb{R}^d$ satisfying

$$f(x) \geqslant f(\bar{x}) + \langle v, x - \bar{x} \rangle$$

for every x in $\mathbb{R}^d$. We let $\partial f(\bar{x})$ denote the set containing all such v and call it the subdifferential of f at $\bar{x}$. Further, $\partial f = \big\{(\bar{x}, v) \in \mathrm{dom}\, f \times \mathbb{R}^d \mid v \in \partial f(\bar{x})\big\}$. The domain of ∂f, notation $\mathrm{dom}\, \partial f$, is the set of all x at which $\partial f(x)$ is not empty. The function f is differentiable at x if and only $\partial f(x)$ is a singleton. Combining Rockafellar [31, Theorem 6.3, Theorem 23.4, and Corollary 25.1.1], we have in fact

$$\left.\begin{aligned} \mathrm{dom}\, \nabla f \subset \mathrm{int}(\mathrm{dom}\, \partial f) &= \mathrm{int}(\mathrm{dom}\, f), \\ \mathrm{dom}\, \partial f \ \subset \quad \mathrm{dom}\, f\ &, \\ \mathrm{cl}(\mathrm{dom}\, \partial f) = \ \mathrm{cl}(\mathrm{dom}\, f)&. \end{aligned}\right\} \tag{47}$$

The concept of (cyclical) monotonicity is linked to the subdifferentials of convex functions.

Definition A.3 (Cyclically monotone maps and sets). *A set $T \subset \mathbb{R}^d \times \mathbb{R}^d$ is cyclically monotone if, for every integer $n \geqslant 1$ and all $(x_1, y_1), \ldots, (x_n, y_n) \in T$ with $y_{n+1} := y_1$, we have*

$$\sum_{i=1}^{n} \langle x_i, y_i \rangle \geqslant \sum_{i=1}^{n} \langle x_i, y_{i+1} \rangle. \tag{48}$$

If Eq. (48) holds for $n = 2$, then T is monotone. The set T is maximal (cyclically) monotone if it is not contained in a larger subset of $\mathbb{R}^d \times \mathbb{R}^d$ that is also (cyclically) monotone.

For a family $\mathcal{G}$ of subsets of $\mathbb{R}^d \times \mathbb{R}^d$, we write $\mathcal{G}_{(c)m}$ for the subfamily consisting of (cyclically) monotone sets in $\mathcal{G}$. Likewise we add m(c)m for maximal (cyclically) monotonicity. The following theorem due to Rockafellar establishes a connection between convex functions and cyclically monotone sets. It is taken from Theorems 24.8 and 24.9 in [31] and Theorem 12.25 in [32].

Theorem A.4 (Rockafellar's Theorem). (a) *A set $T \subset \mathbb{R}^d \times \mathbb{R}^d$ is cyclically monotone if and only if it is contained in the subdifferential of a closed convex function.*

(b) *A set $T \subset \mathbb{R}^d \times \mathbb{R}^d$ has the form $T = \partial\psi$ for some proper, closed convex function $\psi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ if and only if T is maximal cyclically monotone. Then ψ is determined by T uniquely up to an additive constant.*

The following lemma is used in the proof of Theorem 4.1.

Lemma A.5. *Let ψ and ϕ be convex functions on $\mathbb{R}^d$. For every $t \in \mathbb{R}$, the set*

$$\{x \in \operatorname{dom} \nabla\psi \cap \operatorname{dom} \nabla\phi \mid \psi(x) - \phi(x) = t \text{ and } \nabla\psi(x) \neq \nabla\phi(x)\} \tag{49}$$

has Hausdorff dimension at most $d - 1$.

Proof. Let X be the set in Eq. (49). By McCann's Implicit Function Theorem [25, Theorem 17 and Corollary 19], every point x in X has an open neighbourhood N_x such that the $(d-1)$-dimensional Hausdorff measure of the set

$$\{y \in N_x \mid \psi(y) - \phi(y) = t\}$$

is finite. The collection of sets $\{N_x \mid x \in X\}$ forms an open cover of X. By the Lindelöf property of $\mathbb{R}^d$, there exists a countable subcover, i.e., a countable set $Q \subset X$ such that X is a subset of $\bigcup_{x \in Q} N_x$ and then also of

$$\bigcup_{x \in Q} \{y \in N_x \mid \psi(y) - \phi(y) = t\}.$$

The latter set is a countable union of sets of Hausdorff dimension at most $d - 1$ and has therefore Hausdorff dimension at most $d - 1$ as well. □

The convex hull of a function $g : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ is the function $\operatorname{conv} g$ with epigraph $\operatorname{epi}(\operatorname{conv} g) = \operatorname{conv}(\operatorname{epi} g)$, the convex hull of the epigraph of g. It is the greatest convex function majorized by g [31, p. 36]. The closure of a function f is the function $\operatorname{cl} f$ whose epigraph is $\operatorname{epi}(\operatorname{cl} f) = \operatorname{cl}(\operatorname{epi} f)$; it is the same as the lower semicontinuous hull of f. The following lemma is used in the proof of Theorem 5.1.

Lemma A.6 (Restrictions of convex functions). *Let $\psi : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ be a closed convex function and let $D \subset \mathbb{R}^d$ be such that $D \cap \operatorname{dom} \psi \neq \varnothing$. Define the function ϕ by $\phi = \operatorname{cl}(\operatorname{conv} g)$, where $g(x) = \psi(x)$ if $x \in D$ and $g(x) = \infty$ if $x \in \mathbb{R}^d \setminus D$. Then ϕ is a closed convex function with the following properties:*

(a) $\phi \geqslant \psi$;
(b) $\phi(x) = \psi(x)$ *for* $x \in D$;
(c) $\partial\phi(x) = \{v \in \mathbb{R}^d : \forall z \in D,\ \psi(z) \geqslant \psi(x) + \langle z - x, v\rangle\} \supset \partial\psi(x)$ *for* $x \in D$;
(d) $\operatorname{dom}\nabla\phi \subset \operatorname{dom}\partial\psi$;
(e) $D \cap \operatorname{dom}\nabla\phi \subset \operatorname{dom}\nabla\psi$ *and* $\nabla\phi(x) = \nabla\psi(x)$ *for all* $x \in D \cap \operatorname{dom}\nabla\phi$.

Proof. Since $\operatorname{epi}\phi = \operatorname{cl}(\operatorname{epi}(\operatorname{conv} g)) = \operatorname{cl}(\operatorname{conv}(\operatorname{epi} g))$, the epigraph of ϕ is closed and convex. Hence ϕ is a closed and convex function.

(a) As $\operatorname{epi} g = \operatorname{epi}\psi \cap (D \times \mathbb{R})$, we have $\operatorname{epi}\psi \supset \operatorname{epi} g$ and thus, since $\operatorname{epi}\psi$ is convex and closed, $\operatorname{epi}\psi \supset \operatorname{conv}(\operatorname{epi} g)$ and then $\operatorname{epi}\psi \supset \operatorname{cl}(\operatorname{conv}(\operatorname{epi} g)) = \operatorname{epi}\phi$. Hence $\psi \leqslant \phi$.

(b) Since $\operatorname{epi}\phi \supset \operatorname{epi} g$, we must have $\phi(x) \leqslant g(x) = \psi(x)$ for $x \in D$, and thus $\phi(x) = \psi(x)$ for such x.

(c) Let $x \in D$ and write $V = \{v \in \mathbb{R}^d : \forall z \in D,\ \psi(z) \geqslant \psi(x) + \langle z - x, v\rangle\}$. By definition of the subdifferential, $V \supset \partial\psi(x)$. Likewise, in view of (b), also $V \supset \partial\phi(x)$. We need to show that $V \subset \partial\phi(x)$. Let $v \in V$. By definition,

$$\operatorname{epi} g = \operatorname{epi}\psi \cap (D \times \mathbb{R}) \subset \{(z, \lambda) \in \mathbb{R}^d \times \mathbb{R} : \lambda \geqslant \psi(x) + \langle z - x, v\rangle\}.$$

The set on the right-hand side is a closed half-space in $\mathbb{R}^{d+1}$. Therefore, it also contains $\operatorname{cl}(\operatorname{conv}(\operatorname{epi} g)) = \operatorname{epi}\phi$. But that means that $\phi(z) \geqslant \psi(x) + \langle z - x, v\rangle$ for all $z \in \mathbb{R}^d$. As $\psi(x) = \phi(x)$, we find that $v \in \partial\phi(x)$. Therefore, $V \subset \partial\phi(x)$ as required.

(d) From $\psi \leqslant \phi$ we infer $\operatorname{dom}\phi \subset \operatorname{dom}\psi$. Therefore, by (47),

$$\operatorname{dom}\nabla\phi \subset \operatorname{int}(\operatorname{dom}\phi) \subset \operatorname{int}(\operatorname{dom}\psi) = \operatorname{dom}\partial\psi.$$

(e) Let $x \in D \cap \operatorname{dom}\nabla\phi$ and write $\nabla\phi(x) = y$. On the one hand, by (d), $x \in \operatorname{dom}\partial\psi$ and thus $\partial\psi(x) \neq \varnothing$. On the other hand, by (c), $\partial\psi(x) \subset \partial\phi(x) = \{y\}$. We conclude that $\partial\psi(x) = \{y\}$, so that $x \in \operatorname{dom}\nabla\psi$ and $\nabla\psi(x) = y = \nabla\phi(x)$. □

Appendix B: Zero-couplings and their support

This section relies heavily on the work in Segers [36] on the convergence of sequences of couplings $\pi_n \in \Pi(P_n, Q_n)$ where P_n, Q_n are sequences of probability measures on $\mathbb{R}^d$, and the convergence of closed cyclically monotone sets containing their supports, i.e., $T_n \in \mathcal{F}(\mathbb{R}^d \times \mathbb{R}^d)$ satisfying $\operatorname{spt}\pi_n \subset T_n$. Although being formulated for probability measures, most of the results in [36] are in fact given with a demonstration that extends straightforwardly to M_0. Indeed, thanks to Theorem 2.7 (Prohorov) and Theorem 2.4 (Portmanteau) for M_0 (which are quite different from the ones in the set of probability measures, see Section A.1) in [22], most of the arguments can be reused almost directly. In what follows, we adapt the results given in [36] to our context.

Lemma B.1. *The maps* φ_1, φ_2 *in Theorem 2.1 are continuous as functions* $\mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d) \to \mathrm{M}_0(\mathbb{R}^d)$.

Proof of Lemma B.1. We only treat φ_1, as φ_2 can be treated in exactly the same way. It suffices to check that $\gamma_n \xrightarrow{0} \gamma$ in $\mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$ implies $\varphi_1\gamma_n \xrightarrow{0} \varphi_1\gamma$ in $\mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$ as $n \to \infty$. To do so, we consider the Portmanteau Theorem 2.1(iii) in Lindskog et al. [23] (see Theorem A.1) and apply it twice. Let $A \subset \mathbb{R}^d_{\setminus 0}$ be bounded away from the origin. Then $\operatorname{int}(A)$, $\operatorname{int}(A) \times \mathbb{R}^d$ and $\operatorname{cl}(A)$, $\operatorname{cl}(A) \times \mathbb{R}^d$ are open and closeds set, respectively, and bounded away from the origin in $\mathbb{R}^d$ or $\mathbb{R}^d \times \mathbb{R}^d$. Since $\gamma_n \xrightarrow{0} \gamma$ as $n \to \infty$ in $\mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$, we have

$$\gamma\big(\operatorname{int}(A) \times \mathbb{R}^d\big) \leqslant \liminf_{n\to\infty} \gamma_n\big(\operatorname{int}(A) \times \mathbb{R}^d\big) \leqslant \limsup_{n\to\infty} \gamma_n\big(\operatorname{cl}(A) \times \mathbb{R}^d\big) \leqslant \gamma\big(\operatorname{cl}(A) \times \mathbb{R}^d\big),$$

which we can rewrite as

$$(\varphi_1\gamma)(\mathrm{int}(A)) \leqslant \liminf_{n\to\infty}(\varphi_1\gamma_n)(\mathrm{int}(A)) \leqslant \limsup_{n\to\infty}(\varphi_1\gamma_n)(\mathrm{cl}(A)) \leqslant (\varphi_1\gamma)(\mathrm{cl}(A)).$$

It follows that $\varphi_1\gamma_n \xrightarrow{0} \varphi_1\gamma$ as $n \to \infty$, as required. □

For our purpose, the main consequence of Theorem B.2 below is that if a sequence $(\gamma_n, T_n) \in \Gamma_0(\mu_n, \nu_n) \times \mathcal{F}(\mathbb{R}^d \times \mathbb{R}^d)$ satisfying $\mathrm{spt}\,\gamma_n \subset T_n$ converges in the product topology to some (γ, T), then the limit pair satisfies $\mathrm{spt}\,\gamma \subset T$. It is adapted from Lemma D.11 (Lower semi-continuity of the support map) in [36]. The lim inf of a sequence F_n in $\mathcal{F}(\mathbb{E})$ is defined as the set of points $x \in \mathbb{E}$ such that there exist $x_n \in F_n$ such that $x_n \to x$ in $\mathbb{E}$; the set $\liminf_n F_n$ is necessarily closed in $\mathbb{E}$ too. We recall that a map f from a topological space $\mathbb{D}$ into $\mathcal{F}(\mathbb{E})$ is lower semi-continuous if $f^{-1}(\mathcal{F}_G)$ is open for any open set G in $\mathbb{E}$, where $\mathcal{F}_G = \{F \in \mathcal{F}(\mathbb{E}) : F \cap G \neq \varnothing\}$; which, in the case where the space $\mathbb{D}$ is metrizable with metric ρ and $\mathbb{E}$ is a locally compact Hausdorff topological space (LCHS), is equivalent to $f(d) \subset \liminf_n f(d_n)$ whenever $d_n \to d$ in $\mathbb{D}$ [36, lemma D.8].

Lemma B.2 (Lower semi-continuity of the support map). *Let $\mathbb{E}$ be a non-empty open subset of $\mathbb{R}^k$ for some $k \geqslant 1$. If $\mu_n \xrightarrow{0} \mu$ in $\mathrm{M}_0(\mathbb{E})$—seen as a subset of $\mathrm{M}_0(\mathbb{R}^d)$—as $n \to \infty$, then $\mathrm{spt}\,\mu \subset \liminf_{n\to\infty} \mathrm{spt}\,\mu_n$, that is, the map*

$$\mathrm{spt} : \mathrm{M}_0(\mathbb{E}) \to \mathcal{F}(\mathbb{E}) : \mu \mapsto \mathrm{spt}\,\mu$$

is lower semi-continuous. In particular, $\{(\mu, F) \in \mathrm{M}_0(\mathbb{E}) \times \mathcal{F}(\mathbb{E}) : \mathrm{spt}\,\mu \subset F\}$ is closed in the product topology.

Proof of Lemma B.2. The proof is adapted from the one of Lemma D.11 in [36], with minor changes to deal with the M_0 space. Let $G \subset \mathbb{E}$ be open and suppose that $\mathrm{spt}\,\mu \cap G \neq \varnothing$. By Lemma D.8(ii) in [36] we need to show that $\mathrm{spt}(\mu_n) \cap G \neq \varnothing$ for all large n. For $\varepsilon > 0$ small enough, the set $G_\varepsilon = G \setminus \mathrm{cl}\,\mathbb{B}_{0,\varepsilon}$ is a non-empty open set bounded away from 0. Since $\mu \in \mathrm{M}_0(\mathbb{E}) \subset \mathrm{M}_0(\mathbb{R}^d)$ we have $\mathrm{spt}\,\mu \neq \{0\}$ as $\mathrm{spt}\,\mu$ is by definition equal to $\mathrm{spt}\,\tilde{\mu}$ and $\tilde{\mu}(\{0\}) = 0$, so, decreasing ε if needed, we can also suppose $G_\varepsilon \cap \mathrm{spt}\,\mu \neq \varnothing$. It follows from the Portmanteau lemma for M_0-convergence (Theorem A.1, see Theorem 2.1 in [23]) that

$$\liminf_{n\to\infty} \mu_n(G) \geqslant \liminf_{n\to\infty} \mu_n(G_\varepsilon) \geqslant \mu(G_\varepsilon) > 0.$$

Since $\mathrm{spt}(\mu_{(n)}) \cap G \neq \varnothing$ if and only if $\mu_{(n)}(G) > 0$ (since G is open), the first part of the Lemma is proved. The set $\{(\mu, F) : \mathrm{spt}\,\mu \subset F\}$ is closed in view of Lemma D.8(iv) in [36] since $\mathbb{E}$ is a LCHS space. □

The following result is adapted from Lemma 4.1 in [36] and will be useful when dealing with converging sequences. Let $\mathrm{M}_{0,cm}(\mathbb{R}^d \times \mathbb{R}^d)$ denote the subset of $\mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$ of measures whose support is cyclically monotone in the sense of Theorem A.3. Recall that $\mathcal{F}_{\mathrm{cm}}$ denotes the collection of closed subsets of $\mathbb{R}^d \times \mathbb{R}^d$ that are cyclically monotone.

Lemma B.3 (Closure and compactness properties of sets of zero-coupling measures).

(a) *For $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$, the sets $\mathrm{M}_{0,cm}(\mathbb{R}^d \times \mathbb{R}^d)$ and $\Gamma_{0,\mathrm{cm}}(\mu, \nu)$ are closed in $\mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$.*

(b) *The sets*

$$\big\{(\mu,\nu,\gamma)\mid \mu,\nu\in \mathrm{M}_0(\mathbb{R}^d),\gamma\in\Gamma_0(\mu,\nu)\big\} \qquad \textit{and}$$
$$\big\{(\mu,\nu,\gamma)\mid \mu,\nu\in \mathrm{M}_0(\mathbb{R}^d),\gamma\in\Gamma_{0,\mathrm{cm}}(\mu,\nu)\big\}$$

are closed in $\mathrm{M}_0(\mathbb{R}^d)\times\mathrm{M}_0(\mathbb{R}^d)\times\mathrm{M}_0(\mathbb{R}^d\times\mathbb{R}^d)$.

(c) *If* $K,L\subset \mathrm{M}_0(\mathbb{R}^d)$ *are compact, then*

$$\{(\mu,\nu,\gamma)\mid \mu\in K,\nu\in L,\gamma\in\Gamma_0(\mu,\nu)\} \tag{50}$$

is compact too.

Proof of Lemma B.3. We deal with the different assertions separately.

Proof of (a). As $\Gamma_{0,\mathrm{cm}}(\mu,\nu)=\Gamma_0(\mu,\nu)\cap \mathrm{M}_{0,cm}(\mathbb{R}^d\times\mathbb{R}^d)$, it suffices to show that $\Gamma_0(\mu,\nu)$ and $\mathrm{M}_{0,cm}(\mathbb{R}^d\times\mathbb{R}^d)$ are closed.

According to Theorem B.1, the maps $\varphi_1,\varphi_2:\mathrm{M}_0(\mathbb{R}^d\times\mathbb{R}^d)\to\mathrm{M}_0(\mathbb{R}^d)$ sending a measure in $\mathrm{M}_0(\mathbb{R}^d\times\mathbb{R}^d)$ to its zero-marginals are continuous. As a consequence, $\Gamma_0(\mu,\nu)=\varphi_1^{-1}(\{\mu\})\cap\varphi_2^{-1}(\{\nu\})$ is closed.

Next, we have $\mathrm{M}_{0,cm}=\mathrm{spt}^{-1}(\mathcal{F}_{\mathrm{cm}})$ where the map $\mathrm{spt}:\mathrm{M}_0(\mathbb{R}^d\times\mathbb{R}^d)\to\mathcal{F}(\mathbb{R}^d\times\mathbb{R}^d)$ sends a measure to its support. The set $\mathrm{spt}^{-1}(\mathcal{F}_{\mathrm{cm}})$ is closed by lower semi-continuity of spt (Lemma B.2) and the fact that the complement of $\mathcal{F}_{\mathrm{cm}}$ can be written as a union over finite intersections of sets of the form $\mathcal{F}_G=\{F\in\mathcal{F}(\mathbb{R}^d\times\mathbb{R}^d)\mid F\cap G\neq\varnothing\}$ for open $G\subset\mathbb{R}^d\times\mathbb{R}^d$ [36, Lemma 3.2].

Proof of (b). If $\gamma_n\in\Gamma_0(\mu_n,\nu_n)$ for all n and if $\mu_n\xrightarrow{0}\mu$, $\nu_n\xrightarrow{0}\nu$ and $\gamma_n\xrightarrow{0}\gamma$, then also $\gamma\in\Gamma_0(\mu,\nu)$ using arguments similar to those used for (a) and relying on the continuity of φ_i. It follows that $\{(\mu,\nu,\gamma)\mid\mu,\nu\in\mathrm{M}_0,\gamma\in\Gamma_0(\mu,\nu)\}$ is closed in $\mathrm{M}_0(\mathbb{R}^d)\times\mathrm{M}_0(\mathbb{R}^d)\times\mathrm{M}_0(\mathbb{R}^d\times\mathbb{R}^d)$. The intersection of the former set with the closed set $\mathrm{M}_0(\mathbb{R}^d)\times\mathrm{M}_0(\mathbb{R}^d)\times\mathrm{M}_{0,cm}(\mathbb{R}^d\times\mathbb{R}^d)$ is the set

$$\big\{(\mu,\nu,\gamma)\mid\mu,\nu\in\mathrm{M}_0,\gamma\in\Gamma_{0,\mathrm{cm}}(\mu,\nu)\big\},$$

which is thus closed as well.

Proof of (c). Let $M=\bigcup_{\mu\in K,\nu\in L}\Gamma_0(\mu,\nu)$. The set

$$\{(\mu,\nu,\gamma)\mid\mu\in K,\nu\in L,\gamma\in\Gamma_0(\mu,\nu)\}=\{(\mu,\nu,\gamma)\mid\gamma\in\Gamma_0(\mu,\nu)\}\cap\big(K\times L\times\mathrm{M}_0(\mathbb{R}^d\times\mathbb{R}^d)\big)$$

is closed by (b) and contained in $K\times L\times M$. Since K,L are compact subsets of $\mathrm{M}_0(\mathbb{R}^d)$, it suffices to prove that M is relatively compact in $\mathrm{M}_0(\mathbb{R}^d\times\mathbb{R}^d)$ to conclude that $K\times L\times M$ is relatively compact, yielding the desired compactness of the set of interest.

Let $\epsilon>0$. Since K and L are compact in $\mathrm{M}_0(\mathbb{R}^d)$, the Prohorov-like characterization of relative compactness in $\mathrm{M}_0(\mathbb{R}^d)$ stated in Theorem A.2 [22, Theorem 2.7] yields the existence of two positive sequences $(r_i^K)_{,i\geqslant1}$ and $(r_i^L)_{i\geqslant1}$ decreasing to zero, so that for each $i\geqslant1$ there exist compact sets $C_i^K\subset\mathbb{R}^d\setminus\mathbb{B}(0,r_i^K)$ and $C_i^L\subset\mathbb{R}^d\setminus\mathbb{B}(0,r_i^L)$ such that

$$\sup_{\mu\in K}\mu\big(\mathbb{R}^d\setminus\mathbb{B}(0,r_i^K)\big)<\infty, \qquad \sup_{\mu\in K}\mu\big(\mathbb{R}^d\setminus(\mathbb{B}(0,r_i^K)\cup C_i^K)\big)\leqslant\epsilon/2,$$
$$\sup_{\nu\in L}\nu\big(\mathbb{R}^d\setminus\mathbb{B}(0,r_i^L)\big)<\infty, \qquad \sup_{\nu\in L}\nu\big(\mathbb{R}^d\setminus(\mathbb{B}(0,r_i^L)\cup C_i^L)\big)\leqslant\epsilon/2.$$

It is easy to see that we can choose common sequences of numbers and sets by setting

$$r_i = \max(r_i^K, r_i^L), \qquad C_i = (C_i^K \cup C_i^L) \setminus \mathbb{B}(0, r_i),$$

to satisfy the above inequalities for μ and ν simultaneously.

For the Euclidean norm, $\|(x,y)\| \geqslant \sqrt{2}r$ implies $\max(\|x\|, \|y\|) \geqslant r$, where $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ and $r > 0$. As a consequence, we have, indicating the dimensions of the balls by a subscript,

$$\mathbb{R}^{2d} \setminus \mathbb{B}_{2d}(0, \sqrt{2}r_i) \subset \big((\mathbb{R}^d \setminus \mathbb{B}_d(0,r_i)) \times \mathbb{R}^d\big) \cup \big(\mathbb{R}^d \times (\mathbb{R}^d \setminus \mathbb{B}_d(0,r_i))\big),$$

and thus, for every $i \geqslant 1$,

$$\sup_{\gamma \in M} \gamma\big(\mathbb{R}^{2d} \setminus \mathbb{B}_{2d}(0, \sqrt{2}r_i)\big) \leqslant \sup_{\mu \in K} \mu\big(\mathbb{R}^d \setminus \mathbb{B}_d(0,r_i)\big) + \sup_{\nu \in L} \nu\big(\mathbb{R}^d \setminus \mathbb{B}_d(0,r_i)\big) < \infty.$$

Finally, define the compact set

$$K_i = \big(C_i \cup \operatorname{cl} \mathbb{B}_d(0,r_i)\big)^2 \setminus \mathbb{B}_{2d}(0, \sqrt{2}r_i),$$

which is obviously contained in $\mathbb{R}^{2d} \setminus \mathbb{B}_{2d}(0, \sqrt{2}r_i)$. For all $(x,y) \in \mathbb{R}^{2d}$, we have

$$\begin{aligned}
&(x,y) \notin K_i \cup \mathbb{B}_{2d}(0, \sqrt{2}r_i) \\
&\implies (x,y) \notin \big(C_i \cup \operatorname{cl} \mathbb{B}_d(0,r_i)\big)^2 \\
&\implies x \notin C_i \cup \operatorname{cl} \mathbb{B}_d(0,r_i) \text{ or } y \notin C_i \cup \operatorname{cl} \mathbb{B}_d(0,r_i) \\
&\implies x \notin C_i \cup \mathbb{B}_d(0,r_i) \text{ or } y \notin C_i \cup \mathbb{B}_d(0,r_i),
\end{aligned}$$

whence, in $\mathbb{R}^{2d}$,

$$[K_i \cup \mathbb{B}_{2d}(0, \sqrt{2}r_i)]^c \subset [(C_i \cup \mathbb{B}_d(0,r_i))^c \times \mathbb{R}^d] \cup [\mathbb{R}^d \times (C_i \cup \mathbb{B}_d(0,r_i))^c].$$

It follows that

$$\begin{aligned}
\sup_{\gamma \in M} \gamma\big([K_i \cup \mathbb{B}_{2d}(0, \sqrt{2}r_i)]^c\big) &\leqslant \sup_{\mu \in K} \mu\big([C_i \cup \mathbb{B}_d(0,r_i)]^c\big) + \sup_{\nu \in L} \nu\big([C_i \cup \mathbb{B}_d(0,r_i)]^c\big) \\
&\leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we have shown that the set M satisfies the two criteria in Theorem A.2, confirming that M is relatively compact, as required. □

The following lemma is similar to Lemma C.1 in [36] which concerns probability measures. Recall that $\varphi_1\gamma$ and $\varphi_2\gamma$ are the zero-marginals of $\gamma \in \mathrm{M}_0(\mathbb{R}^d)$ while $\phi_1\gamma$ and $\phi_2\gamma$ are its ordinary marginals (Theorem 2.2 and Theorem 2.3). Let $\operatorname{proj}_1 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ be the projection mapping $(x,y) \mapsto x$. Seeing a subset T of $\mathbb{R}^d \times \mathbb{R}^d$ as (the graph of) a multivalued mapping from $\mathbb{R}^d$ into $\mathbb{R}^d$, we also write $\operatorname{dom} T = \operatorname{proj}_1(T) = \big\{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^d : (x,y) \in T\big\}$.

Lemma B.4 (Support of a zero-margin). *Let $\gamma \in \mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$ and let $\mu = \varphi_1\gamma \in \mathrm{M}_0(\mathbb{R}^d)$ denote its left zero-marginal. Then $\operatorname{spt}\mu \subset \operatorname{spt}\phi_1\gamma = \operatorname{cl}(\operatorname{proj}_1(\operatorname{spt}\gamma))$. As a consequence, if $\operatorname{spt}\gamma \subset T$ for some $T \subset \mathbb{R}^d \times \mathbb{R}^d$, then $\operatorname{spt}\mu \subset \operatorname{cl}(\operatorname{dom} T)$. In particular, if T is maximal monotone, then $\operatorname{int}(\operatorname{spt}\mu) \subset \operatorname{dom} T$.*

Proof of Lemma B.4. We consider the extension $\tilde{\gamma}$ of γ to a positive Borel measure on $\mathbb{R}^d \times \mathbb{R}^d$ and write $\bar{\mu} = \phi_1 \tilde{\gamma} = \phi_1 \gamma$ for its left marginal. Since μ is the restriction of $\bar{\mu}$ to $\mathbb{R}^d_{\setminus 0}$, we have $\operatorname{spt}\mu \subset \operatorname{spt}\bar{\mu}$.

We show that $\operatorname{spt}\bar{\mu}$ is equal to $F = \operatorname{cl}(\operatorname{proj}_1(\operatorname{spt}\tilde{\gamma}))$. First, to show that $\operatorname{spt}\bar{\mu} \supset F$, it suffices to show that $\operatorname{spt}\bar{\mu} \supset \operatorname{proj}_1(\operatorname{spt}\tilde{\gamma})$, as $\operatorname{spt}\bar{\mu}$ is closed. But the latter inclusion follows easily: if $x \in \operatorname{proj}_1(\operatorname{spt}\tilde{\gamma})$, then there exists y such that $(x, y) \in \operatorname{spt}\tilde{\gamma}$; for any open neighbourhood U of x, the set $U \times \mathbb{R}^d$ is an open neighbourhood of (x, y), and thus $\bar{\mu}(U) = \tilde{\gamma}(U \times \mathbb{R}^d) > 0$, which, as U was arbitrary, implies $x \in \operatorname{spt}\bar{\mu}$.

Second, to show that $\operatorname{spt}\bar{\mu} \subset F$, we need to show that $\bar{\mu}(F^c) = 0$, as F is closed. We have $\bar{\mu}(F^c) = \tilde{\gamma}(F^c \times \mathbb{R}^d)$ and

$$F^c \times \mathbb{R}^d \subset (\operatorname{proj}_1(\operatorname{spt}\tilde{\gamma}))^c \times \mathbb{R}^d \subset (\operatorname{spt}\tilde{\gamma})^c,$$

which is a $\tilde{\gamma}$-null set.

If $\operatorname{spt}\gamma \subset T$ for some $T \subset \mathbb{R}^d \times \mathbb{R}^d$, then $\operatorname{proj}_1(\operatorname{spt}\gamma) \subset \operatorname{dom} T$ and thus $\operatorname{spt}\bar{\mu} \subset \operatorname{cl}(\operatorname{dom} T)$.

Finally, assume T contains $\operatorname{spt}\gamma$ and is maximal monotone. By Theorem 12.41 (near convexity of domains and ranges) in [32], there exists a convex subset C of $\mathbb{R}^d$ such that $C \subset \operatorname{dom} T \subset \operatorname{cl} C$. As $\operatorname{spt}\mu \subset \operatorname{cl}(\operatorname{dom} T)$, we have $\operatorname{spt}\mu \subset \operatorname{cl} C$, hence $\operatorname{int}(\operatorname{spt}\mu) \subset \operatorname{int}\operatorname{cl} C$. Moreover, the convexity of C yields $\operatorname{ri} C = \operatorname{ri}\operatorname{cl} C$, where ri is the relative interior of C with respect to the smallest affine subspace that contains C. As a consequence, we can write

$$\begin{aligned}\operatorname{int}(\operatorname{spt}\mu) \subset \operatorname{int}\operatorname{cl} C &\subset \operatorname{ri}\operatorname{cl} C \\ &= \operatorname{ri} C \subset C \subset \operatorname{dom} T,\end{aligned}$$

and the statement is proved. □

Remark B.5. Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ and $\gamma \in \Gamma_0(\mu, \nu)$, and write $\bar{\mu} = \phi_1 \gamma$ for its left margin (Theorem 2.3). Since $\mu = \operatorname{res}\bar{\mu}$ we have the inclusions $\operatorname{spt}\mu \subset \operatorname{spt}\bar{\mu} \subset \operatorname{spt}\mu \cup \{0\}$. It follows by Theorem B.4 that $\operatorname{spt}\mu \setminus \{0\} = \operatorname{cl}(\operatorname{proj}_1(\operatorname{spt}\gamma)) \setminus \{0\}$.

Recall that we write $T \llcorner V = T \cap (V \times \mathbb{R}^d)$ for $T \subset \mathbb{R}^d \times \mathbb{R}^d$ and $V \subset \mathbb{R}^d$. The following lemma will be useful to prove the graphical convergence, relative to the interior V of the support of μ, of a sequence of maximal cyclically monotone sets $T_n \in \mathcal{F}(\mathbb{R}^d \times \mathbb{R}^d)$, for $n \in \mathbb{N}$, containing the support of γ_n by allowing one to prove that the limit points of $T_n \llcorner V$ are all the same. It is adapted from Lemma 4.2 in [36]. For $T \subset \mathbb{R}^d \times \mathbb{R}^d$, we write $T^{-1} = \{(y, x) \in \mathbb{R}^d \times \mathbb{R}^d \mid (x, y) \in T\}$, $\operatorname{dom} T = \operatorname{proj}_1(T)$ and $\operatorname{rge} = \operatorname{proj}_2(T)$.

Lemma B.6 (Uniqueness of maximal monotone extension on the interior of the support). *Let $\gamma \in \Gamma_0(\mu, \nu)$ for $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ and let $S, T \in \mathcal{F}_{\mathrm{mm}}(\mathbb{R}^d \times \mathbb{R}^d)$ satisfy $\operatorname{spt}\gamma \subset S \cap T$. Put $V = \operatorname{int}(\operatorname{spt}\mu)$ and $W = \operatorname{int}(\operatorname{spt}\nu)$. Then $V \subset \operatorname{dom} S \cap \operatorname{dom} T$ and $S \llcorner V = T \llcorner V$, as well as $W \subset \operatorname{rge} S \cap \operatorname{rge} T$ and $S^{-1} \llcorner W = T^{-1} \llcorner W$.*

Proof of Lemma B.6. Let $\tilde{\gamma}$ be the extension of γ to $\mathbb{R}^d \times \mathbb{R}^d$ defined by $\tilde{\gamma}(B) = \gamma(B \setminus \{(0,0)\})$ for $B \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d)$. Let $\hat{\mu} = \phi_1 \tilde{\gamma}$ and $\hat{\nu} = \phi_2 \tilde{\gamma}$ be its left- and right-marginals; by definition, $\tilde{\gamma}$ lies in $\Pi(\hat{\mu}, \hat{\nu})$. According to Theorem B.4 and Lemma C.1 in Segers [36], it is immediate that both $\hat{V} = \operatorname{int}\operatorname{spt}\hat{\mu}$ and V are contained in both $\operatorname{dom} S$ and $\operatorname{dom} T$, since $\operatorname{spt}\gamma = \operatorname{spt}\tilde{\gamma}$. The rest of the proof of Lemma 4.2 in [36] can then be reused directly since it exclusively relies on properties of maximal monotone maps and not on the finiteness of the measures considered there. It follows that $S(x) = T(x)$ for all $x \in \hat{V}$. Since $\operatorname{spt}\mu \subset \operatorname{spt}\hat{\mu}$, we have $S(x) = T(x)$ for all $x \in V$ too.

The statements about W follow by switching the roles of μ and ν, upon noting that $\operatorname{dom} S^{-1} = \operatorname{rge} S$. □

Lemma B.7. *Let $\mu, \nu \in \mathrm{M}_0(\mathbb{R}^d)$ with $\mu(\mathbb{R}^d_{\setminus 0}) = \infty$ or $\nu(\mathbb{R}^d_{\setminus 0}) = \infty$ (or both), and let $\gamma \in \Gamma_0(\mu, \nu)$ be supported by $\partial\psi$ for some closed convex function ψ. Then $(0,0) \in \operatorname{spt}\gamma \subset \partial\psi$ and $\psi(0) = \inf_{x \in \mathbb{R}^d} \psi(x)$.*

Proof. The zero-coupling measure γ must have infinite mass: if, for instance, $\mu(\mathbb{R}^d_{\setminus 0}) = \infty$, then also $\gamma(\mathbb{R}^d_{\setminus 0} \times \mathbb{R}^d) = \mu(\mathbb{R}^d_{\setminus 0}) = \infty$. By Theorem 2.4, γ belongs to $\mathrm{M}_0(\mathbb{R}^d \times \mathbb{R}^d)$. Therefore, necessarily $\gamma(\mathbb{B}_d(0,\varepsilon) \times \mathbb{B}_d(0,\varepsilon)) = \infty > 0$ for every $\varepsilon > 0$, which means that $\operatorname{spt}\gamma$ meets $\mathbb{B}_d(0,\varepsilon) \times \mathbb{B}_d(0,\varepsilon)$ for every $\varepsilon > 0$. As a consequence, $(0,0) \in \operatorname{spt}\gamma \subset \partial\psi$, the latter inclusion holding by assumption. But then also $0 \in \operatorname{dom}\partial\psi \subset \operatorname{dom}\psi$ in view of Eq. (47). By definition of the subdifferential we find, as required,

$$\forall x \in \mathbb{R}^d, \qquad \psi(x) \geqslant \psi(0) + \langle x - 0, 0\rangle = \psi(0). \qquad \square$$