Decomposition of optimal transport plans and entropic regularization for the L^1 Monge-Kantorovich problem on the real line

Armand Ley

April 2025

1 Introduction

1.1 Monge-Kantorovitch problem and decomposition of optimal transport plans

Summary on the optimal transport problem In recent decades, optimal transport theory has attracted considerable attention, due to its many connections to probability, analysis, geometry and other areas of mathematics (see, e.g., [30, 29, 25, 2, 24]). In this paper, we investigate one of the most elementary optimal transport problems, namely the L^1 Monge-Kantorovich problem on the real line. Given two probability measures μ and ν, this problem consists of minimizing the global transport cost function

$$J : \pi \in \mathrm{Marg}(\mu, \nu) \mapsto \int_{\mathbb{R}^2} |y - x| \, \mathrm{d}\pi(x, y),$$

over the set $\mathrm{Marg}(\mu, \nu)$ of measures that satisfy the marginal constraints $\pi(A\times\mathbb{R}) = \mu(A)$ and $\pi(\mathbb{R}\times B) = \nu(B)$ for every pair (A, B) of Borel sets. If $|y-x|$ is replaced by $|y-x|^p$ for some $p > 1$ in the definition of the cost function J, then the set $\mathcal{O}(\mu, \nu)$ of minimizers of J, known as optimal transport plans, reduces to a singleton which consists of the well-known (co)monotone transport. This transport plan, also known as quantile transport plan, is defined as $(G_\mu, G_\nu)_\# \left(\mathcal{L}^1_{\llcorner]0,1[}\right)$, where $\mathcal{L}^1_{\llcorner]0,1[}$ stands for the restriction of the Lebesgue measure to $]0, 1[$, while G_μ and G_ν denote the quantile functions of μ and ν, respectively. For $p = 1$ uniqueness of an optimal transport plan does not hold, and describing the set $\mathcal{O}(\mu, \nu)$ of optimal transport plans becomes more delicate. A key result of this paper is a decomposition theorem for $\mathcal{O}(\mu, \nu)$. More precisely, we construct a set $\mathcal{D} = \{(\mu_i, \nu_i) \; ; \; i \in \mathcal{I}\}$ of pairs of measures such that, for all $i \in \mathcal{I}$, the set $\mathcal{O}(\mu_i, \nu_i)$ admits a tractable description and, for every $\pi \in \mathcal{O}(\mu, \nu)$, there exists a unique family $(\pi_i)_{i\in\mathcal{I}} \in \prod_{i\in\mathcal{I}} \mathcal{O}(\mu_i, \nu_i)$ such that $\pi = \sum_{i\in\mathcal{I}} \pi_i$. This motivates the following notation.

Notation 1.1 (Minkowski sum and direct sum of set of measures)**.** Given a countable family $(E_i)_{i\in\mathcal{I}}$ of subsets of $\mathcal{M}_+(\mathbb{R}^2)$, we denote by $\sum_{i\in\mathcal{I}} E_i$ the image of the map $\phi : (\pi_i)_{i\in\mathcal{I}} \in \prod_{i\in\mathcal{I}} E_i \mapsto \sum_{i\in\mathcal{I}} \pi_i \in \mathcal{M}_+(\mathbb{R}^2)$. If ϕ is injective, we write $\bigoplus_{i\in\mathcal{I}} E_i$ instead of $\sum_{i\in\mathcal{I}} E_i$.

With this notation, our decomposition theorem becomes $\mathcal{O}(\mu, \nu) = \bigoplus_{i\in\mathcal{I}} \mathcal{O}(\mu_i, \nu_i)$. Strictly speaking, the decomposition is formulated in the more general setting of cyclical monotone transport plans (see Theorem 2.36). In the case of measures without atoms and with compact support, it coincides with a previous result by Di Marino and Louet [10, Proposition 3.1]. The second part of the article is devoted to the selection problem associated to the entropic regularization of our transport problem: we defer its detailed presentation and the statement of our results to the second part of this introduction. To motivate our decomposition theorem and to clarify the exposition of the decomposition procedure, we briefly present a related decomposition result for the set of transport plans satisfying the martingale constraint.

Decomposition theorem for the set of martingale transport plans A classical theorem (see Strassen [27]) states that the set $\mathrm{Mart}(\mu,\nu)$ of martingale transport plans, i.e. the transport plans π that admit a disintegration of the form $\pi(\mathrm{d}x,\mathrm{d}y)=\mu(\mathrm{d}x)k_x(\mathrm{d}y)$ with $\int_{\mathbb{R}} y k_x(\mathrm{d}y)=x$, is non-empty if and only if the measures μ and ν are in the convex order, meaning that the potential functions $u_\mu : t\mapsto \int_{\mathbb{R}}|x-t|\ \mathrm{d}\mu(x)$ and $u_\nu : t\mapsto \int_{\mathbb{R}}|x-t|\ \mathrm{d}\nu(x)$ satisfy $u_\mu\le u_\nu$.[1] Assuming $u_\mu\le u_\nu$, let $(I_k)_{k\in\mathcal{K}}$ be the family of connected components of $\{u_\mu<u_\nu\}:=\{x\in\mathbb{R}\ ;\ u_\mu(x)<u_\nu(x)\}$, and define the components of μ by setting $\mu^= = \mu\llcorner_{\{u_\mu=u_\nu\}}$ and $\mu_k=\mu\llcorner_{I_k}$ for each $k\in\mathcal{K}$. Beiglböck and Juillet [4, Theorem A.4] proved that there exists a unique family $((\nu_k)_{k\in\mathcal{K}},\nu^=)$ such that $\nu=\sum_{k\in\mathcal{K}}\nu_k+\nu^=$, $u_{\mu^=}\le u_{\nu^=}$ and $u_{\mu_k}\le u_{\nu_k}$ for each $k\in\mathcal{K}$. Moreover, the inequality $u_{\mu_k}<u_{\nu_k}$ is satisfied in the interior of the support of μ and ν, and they established that $\mathrm{Mart}(\mu,\nu)$ admits the following decomposition:

$$\mathrm{Mart}(\mu,\nu)=\left(\bigoplus_{k\in\mathcal{K}}\mathrm{Mart}(\mu_k,\nu_k)\right)\oplus\mathrm{Mart}(\mu^=,\nu^=).$$

The key ingredient of this result relies on the fact that the set

$$\{x\in\mathbb{R}\ ;\ \forall\pi\in\mathrm{Mart}(\mu,\nu),\pi((]-\infty,x[\times[x,+\infty[)\uplus(]x,+\infty[\times]-\infty,x]))=0\}$$

coincides precisely with the set where the potential functions are equal. In higher dimensions, although the situation is notably more intricate, similar results have recently been established [13, 19, 21].

A decomposition result for cyclically monotone transport plan Our decomposition method for $\mathcal{O}(\mu,\nu)$ closely mirrors that of $\mathrm{Mart}(\mu,\nu)$: the optimality constraint replaces the martingale constraint, and the cumulative distribution functions take the role of the potential functions. In this introduction, we present the decomposition result of $\mathcal{O}(\mu,\nu)$ for atomless measures: in this case, the analogy with the martingale decomposition becomes particularly clear, and our decomposition theorem can be presented more intuitively. In this context, the cumulative distribution functions F_μ^+ and F_ν^+ of μ and ν are continuous, so that the sets $F_\mu^+>F_\nu^+$ and $F_\mu^+<F_\nu^+$ are open, and their connected components, denoted by $(I_k^+)_{k\in\mathcal{K}^+}$ and $(I_k^-)_{k\in\mathcal{K}^-}$, form countable families of open intervals. The components of μ and ν are then defined by restricting them to these connected components. More precisely, we set

$$\begin{cases}\mu^= = \mu\llcorner_{\{F_\mu^+=F_\nu^+\}}\\ \nu^= = \nu\llcorner_{\{F_\mu^+=F_\nu^+\}}\end{cases},\begin{cases}\mu_k^+ = \mu\llcorner_{I_k^+}\\ \nu_k^+ = \nu\llcorner_{I_k^+}\end{cases}\text{ and }\begin{cases}\mu_k^- = \mu\llcorner_{I_k^-}\\ \nu_k^- = \nu\llcorner_{I_k^-}\end{cases}. \tag{1}$$

We write $\mu\le_{\mathrm{F}}\nu$ when the inequality $F_\mu^+\ge F_\nu^+$ holds, with strict inequality $F_\mu^+>F_\nu^+$ on the interior of the union of the supports of μ and ν.[2] We can now state our decomposition result in the case of atomless measures.

[1] The measure μ is smaller than ν in the convex order if $\int f\ \mathrm{d}\mu\le\int f\ \mathrm{d}\nu$ for all convex function f. This characterization trough the ordering of potential functions holds only for measures on $\mathbb{R}$. In the following, we refer to a statement as a *Strassen-type* result if it characterizes a stochastic order via the existence of a transport plan satisfying structural constraints.

[2] For measures that may have atoms, the definition $\le_{\mathrm{F}}$ requires both right continuous *and* left-continuous distribution functions, and strict inequality may also be required at extremal points of the support: we refer the reader to Definition 2.33.

Theorem A. Let μ and ν be two atomless measures on $\mathbb{R}$ such that $\min J < +\infty$, and define the marginal components $((\mu_k^+)_{k\in\mathcal{K}^+}, (\mu_k^-)_{k\in\mathcal{K}^-}, \mu^=)$ and $((\nu_k^+)_{k\in\mathcal{K}^+}, (\nu_k^-)_{k\in\mathcal{K}^-}, \nu^=)$ as in Equation (1).

1. Then $\mu^= = \nu^=$, and for all $k \in \mathcal{K}^+$ (resp. $k \in \mathcal{K}^-$), we have $\mu_k^+ \leq_{\mathrm{F}} \nu_k^+$ (resp. $\nu_k^- \leq_{\mathrm{F}} \mu_k^-$).
2. The set $\mathcal{O}(\mu,\nu)$ admits the following decomposition:

$$\mathcal{O}(\mu,\nu) = \left(\bigoplus_{k\in\mathcal{K}^+} \mathcal{O}(\mu_k^+,\nu_k^+)\right) \oplus \left(\bigoplus_{k\in\mathcal{K}^-} \mathcal{O}(\mu_k^-,\nu_k^-)\right) \oplus \mathcal{O}(\mu^=,\nu^=).$$

As for the decomposition of $\mathrm{Mart}(\mu,\nu)$, the proof of this result relies on the characterization of a specific set, whose elements will be called barrier points. This set, denoted by $\mathcal{B}(\mu,\nu)$ is defined by

$$\mathcal{B}(\mu,\nu) = \{x \in \mathbb{R} \; ; \; \forall \pi \in \mathcal{O}(\mu,\nu), \pi(]-\infty,x[\times]x,+\infty[\cup]x,+\infty[\times]-\infty,x) = 0\},$$

and coincides with the set of points where F_μ^+ and F_ν^+ are equal.[3] In case μ and ν have compact support this result coincides with that of Di Marino and Louet: we mention that their proof relies on the construction of an explicit solution to the dual problem and on the use of complementary slackness to constrain mass displacement for elements of $\mathcal{O}(\mu,\nu)$.[4]. In contrast, our approach is based on cyclical monotonicity, a geometric property of the support of optimal transport plans (see Definition 2.3), which we use to establish our earlier characterization of barrier points. In Section 2, after proving Proposition A, we extend this decomposition result to the case where the measures may have atoms, which, as we will see in detail, requires a more refined analysis. Without delving into technical details at this stage, we note that the general case requires working with both right-continuous *and* left-continuous cumulative distribution functions (see Definition 2.18 and the preceding example), that components may not be mutually singular (see Definition 2.22 and Point 3 of Remark 2.24), and that a refinement of the notion of barrier points is required to control mass repartition at the boundaries of the components (see Proposition 2.52). In Theorem 2.36, after adapting the definition of marginal components, we prove Theorem A for measures (μ,ν) that may have atoms. This result, stated in the broader context of cyclically monotone transport plans, coincides with Theorem A when μ and ν are atomless measures such that $\min J < +\infty$. In the appendix, we prove a refinement of this decomposition theorem, and we address the question of the optimality of both the original decomposition and its refined version.

We now highlight a interesting consequence of our approach. For L^p transport with $p > 1$, it is well known that a transport plan is optimal if and only if there exists a set Γ such that, for every $(x,y),(x',y') \in \Gamma$, $|y-x|^p + |y'-x'|^p \leq |y-x'|^p + |y'-x|^p$.[5] We establish that this characterization of optimality also holds for $p = 1$, and provide an interpretation in terms of "crossings" (see Proposition 2.40).

[3] This is not true anymore if μ and ν may have atoms: we refer the reader to Equation (4) for the general characterization of $\mathcal{B}(\mu,\nu)$ in terms of cumulative distribution functions.

[4] Di Marino and Louet proved a slightly different version of this result (see Proposition 2.1 for the exact statement).

[5] This condition corresponds to cyclical monotonicity (see Definition 2.3) for cycles of length two, and is equivalent to $(y'-x')(y-x) \geq 0$. This is the key argument used to show that $\mathcal{O}(\mu,\nu)$ reduces to the monotone transport plan.

1.2 Entropic optimal transport for the L^1 problem on the real line

Summary on the entropic optimal transport problem In the second part of the article, following the same objective as Di Marino and Louet, we investigate the selection problem in entropic optimal transport. Given two probability measures μ and ν, and a regularization parameter $\varepsilon > 0$, the entropic optimal transport problem consists in minimizing the regularized cost function

$$J_\varepsilon : \pi \in \mathrm{Marg}(\mu,\nu) \mapsto \int_{\mathbb{R}^2} |y-x| \, \mathrm{d}\pi(x,y) + \varepsilon \mathrm{Ent}(\pi|\mu\otimes\nu),$$

where the relative entropy $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ is defined as follows:

$$\mathrm{Ent}(\pi|\mu\otimes\nu) = \begin{cases} \int_{\mathbb{R}^2} \log\left(\frac{\mathrm{d}\pi}{\mathrm{d}\mu\otimes\nu}\right)\frac{\mathrm{d}\pi}{\mathrm{d}\mu\otimes\nu} \, \mathrm{d}\mu\otimes\nu & \text{if } \pi \ll \mu\otimes\nu \\ +\infty & \text{otherwise} \end{cases}.$$

When $\varepsilon = 0$, we recover the usual optimal transport problem. Heuristically, since $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ can be interpreted as a measure of divergence from $\mu\otimes\nu$, the entropic term in J_ε encourages the minimizer $\pi_\varepsilon := \mathrm{argmin} J_\varepsilon$ to balance two competing effects: minimizing transport cost *and* being as close as possible to the product measure. This regularization, which applies to more general cost functions, was popularized by Cuturi [9] and has since become a highly active area of research. Its main appeal is likely due to the fact that it transforms the optimal transport problem —which solutions are hard to compute in high dimensions— into a computationally tractable problem. Indeed, adding this regularization term allows the use of robust and efficient algorithms, such as the celebrated Sinkhorn algorithm [26]. We refer to [23] and the numerous references therein for a comprehensive overview of the computational aspects related to optimal transport and to [20] for a general introduction to entropic optimal transport theory.

Since entropic optimal transport is viewed as an approximation of classical optimal transport, understanding its behaviour as the regularization parameter ε tends to 0^+ is of critical importance. Beyond its computational applications, several aspects of this convergence have been studied, including the convergence of minimizers, convergence of solutions to the associated dual problem, stability with respect to the data (marginals and cost function), and convergence rate estimates. In this paper, we are interested in the question of the convergence of the minimizers of J_ε. We refer the works [8, 17] in the context of the squared Euclidean distance, where Γ-convergence methods are used to prove the convergence of the entropic minimizers to the unique optimal transport plan. This result has been extended in [5], where the authors show that for continuous cost functions, every cluster point of entropic minimizers is a cyclically monotone transport plan: in particular, if there exists a unique optimal transport plan, we obtain the convergence toward this transport plan. In the case of measures with finite support, it is known [23, Proposition 4.1] that entropic minimizers converge to $\mathrm{argmin}_{\mathcal{O}(\mu,\nu)}\mathrm{Ent}(\cdot|\mu\otimes\nu)$. Outside these cases, proving the convergence and finding a characterization of the potential limit remains an open question and an active area of research. This selection problem can be reformulated as follows: is an optimal transport plan selected by entropic regularization, and if so, how can it be characterized?

Convergence of minimizers for the L^1 entropic regularization on the real line The most elementary setting in which this selection problem arises is the distance cost on the real line: $\mathcal{O}(\mu,\nu)$ is generally not a singleton, and optimal transport plans are not even absolutely continuous with respect to the product measure.[6] Our first contribution to this selection problem is to use our decomposition result for $\mathcal{O}(\mu,\nu)$ and a Strassen-type theorem due to Kellerer [15] to build a transport plan $\mathcal{K}(\mu,\nu)$ with the following special property: each positive (respectively negative) component of $\mathcal{K}(\mu,\nu)$ is strongly multiplicative, meaning that it coincides with the restriction of a product measure on $\mathrm{F} := \{(x,y) \in \mathbb{R}^2 \ ; \ x \leq y\}$ (respectively on $\tilde{\mathrm{F}} := \{(x,y) \in \mathbb{R}^2 \ ; \ x \geq y\}$). We conjecture that $(\pi_\varepsilon)_{\varepsilon>0}$ converges to $\mathcal{K}(\mu,\nu)$ when $\varepsilon \to 0^+$. In [10, Theorem 4.1], the authors proved $\mathcal{K}(\mu,\nu) = \lim_{\varepsilon\to0^+} \pi_\varepsilon$ when μ and ν are compactly supported, have finite entropy with respect to the Lebesgue measure on $\mathbb{R}$, and there exists $\pi \in \mathcal{O}(\mu - \mu^=, \nu - \nu^=)$ with finite entropy with respect to $\mu \otimes \nu$ (see Sub-subsection 3.2.2).[7] We then adapt the proof from the discrete setting to prove that this conjecture holds whenever there exists an optimal transport plan with finite entropy. Our main convergence result is stated as follows.

Theorem B. Consider a pair (μ,ν) of probability measures. If for all $k \in \mathcal{K}^+$ (respectively $k \in \mathcal{K}^-$), (μ_k^+, ν_k^+) (resp. (μ_k^-, ν_k^-)) forms a pair of mutually singular measures, then $\pi_\varepsilon \xrightarrow[\varepsilon\to0^+]{} \mathcal{K}(\mu,\nu)$.

This covers the case where one measure is atomic and the other has no atoms. In the general case the conjecture is still open, but we prove the following result.

Theorem C. Let (μ,ν) be a pair of probability measures. If π^* is a cluster point of $(\pi_\varepsilon)_{\varepsilon>0}$, then every restriction of π^* to a product set contained in $\{(x,y) \in \mathbb{R}^2 \ ; \ x \leq y\}$ or $\{(x,y) \in \mathbb{R}^2 \ ; \ x \geq y\}$ coincides with a product measure.

This property, known as weak multiplicativity was introduced by Kellerer [15], and its "strict version" has proven useful in studying the Markov property of real-valued point processes [16].

Decomposition and regularized problem in higher dimension In the Monge–Kantorovich problem on $\mathbb{R}^n$ with cost given by the Euclidean norm, a classical decomposition strategy due to Sudakov [28] has proven instrumental to establish the existence of an optimal transport map. The core idea of this decomposition is to disintegrate the measures μ and ν into families $(\mu_\sigma)_\sigma$ and $(\nu_\sigma)_\sigma$ of measures concentrated on one-dimensional "transport rays". This allows one to consider, on each transport ray σ, a solution π_σ to the one-dimensional transport from μ_σ to ν_σ. The delicate step consists in gluing the π_σ together into an optimal transport plan from μ to ν. Assuming $\mu \ll \mathcal{L}^n$, and if each π_σ is chosen as the monotone transport between μ_σ and ν_σ, this gluing procedure is well-defined and yields an optimal transport *map*. For further details on this decomposition, we refer the reader to the lecture notes by Ambrosio [1] as well as to the textbooks [25, Sections 3.1.3 and 3.1.4] and [18, Chapter 18] (which contains an extensive bibliographical note on the historical development of this decomposition). It is worth noting that this type of decomposition differs in several respects from both our decomposition and the martingale transport decomposition. The main difference lies in the fact that, rather than decomposing μ and ν into a family of sub-measures inducing a decomposition of the global space, the decomposition proceeds via disintegration. The assumption $\mu \ll \mathcal{L}^1$ plays a crucial role in this framework. Nevertheless, this method has proven useful in addressing the selection problem arising in the study of L^1 entropic optimal transport for the Euclidean distance cost $c(x,y) = \|y-x\|$. Recently, Aryan and Ghosal proved the following selection result [3]. Let X and Y be compact subsets of $\mathbb{R}^d$ such that $d(X,Y) = \inf(\{\|y-x\| \ ; \ x \in X, y \in Y\}) > 0$,

[6]E.g., if $\mu = \delta_0 + \delta_1 + \mathcal{L}^1_{\llcorner]2,3[}$ and $\nu = \delta_1 + \delta_2 + \mathcal{L}^1_{\llcorner]2,3[}$, then $\mathcal{O}(\mu,\nu) = \mathrm{Marg}(\delta_0+\delta_1, \delta_1+\delta_2) \oplus \{(\mathrm{id},\mathrm{id})_\#\mathcal{L}^1_{\llcorner]2,3[}\}$.

[7]Along with some additional technical assumption, see Theorem 3.9 for the precise statement.

and let μ and ν be probability measures on X and Y with smooth densities with respect to the Lebesgue measure. Then, the family $(\pi_\varepsilon)_{\varepsilon>0}$ converges to an optimal transport plan π^{opt}, and π^{opt} is characterized as the optimal transport plan that minimizes a relative entropy along each transport ray.

1.3 Organization of the paper

In Subsection 2.1, we begin by stating the result of Di Marino and Louet (Proposition 2.1) in the case of atomless marginals. We believe this provides the reader with intuition, while also introducing tools that will be used throughout the remainder of the paper. After introducing cyclically monotone transport plans (Definition 2.3) and barrier points (Definition 2.8), we provide a concise alternative proof of their result. This proof relies on the fact that elements of $\{F_\mu^+ = F_\nu^+\}$ are barrier points (Proposition 2.9) and on a characterization of optimal transport plans for pairs of stochastically ordered measures (Proposition 2.11). Building on this result, we establish our decomposition theorem for atomless marginals (Proposition 2.14).

In Subsection 2.2, we prove a similar decomposition theorem for general measures, which are not necessarily atomless. We begin with the proof of a technical lemma stating that the relative positions of the cumulative distribution functions impose constraints on the regions where optimal transport plans can concentrate (Proposition 2.15). Next, we define the components of the real line (Definition 2.18), the components of the marginals (Definition 2.22), and the components of an optimal transport plan (Definition 2.26). Subsequently, we prove that the components of optimal transport plans posses the correct marginals (Proposition 2.29). We then introduce the reinforced stochastic order (Definition 2.33) and prove that the (non-equal) pairs of marginal components are ordered in this reinforced stochastic order (Proposition 2.35). Finally, we prove our decomposition theorem for $\mathcal{O}(\mu,\nu)$ (Theorem 2.36). At the end of this section, we establish that a transport plan is optimal if and only if it is concentrated on a set of transport paths that do not "freely" cross (Proposition 2.40).

In Subsection 2.3, we investigate the optimality of our decomposition. After defining the set $\mathcal{A}$ of admissible decompositions of (μ,ν) (Definition 2.43), we introduce a relation $\preceq_\mathcal{A}$ on $\mathcal{A}$ (Definition 2.45). We then state that $\preceq_\mathcal{A}$ defines a partial order on $\mathcal{A}$ and that the decomposition constructed in Section 2.2 is a minimal element of $(\mathcal{A}, \preceq_\mathcal{A})$ (Theorem 3.6). Given the technical nature of the proof, we defer it to the appendix (Theorem 2.47). However, we prove a key ingredient of its proof: the characterization of barrier points in terms of relative positions of cumulative distribution functions (Proposition 2.52), which relies on the Strassen-type theorem for the reinforced stochastic order (Theorem 2.50).

In Section 3, we study the behaviour of $(\pi_\varepsilon)_{\varepsilon>0}$ as ε tends to 0^+. In Subsection 3.1, we use a straightforward adaptation of a result by Di Marino and Louet (Lemma 3.5) and the fact that the relative entropy of an optimal transport plan equals the sum of the relative entropy of its components (Proposition 3.23) to establish that $\mathcal{K}(\mu,\nu)$ minimizes the relative entropy among optimal transport plans (Theorem 3.6).

In Subsection 3.2, we explain why this implies that, if there exists an optimal transport plan with finite entropy, then $(\pi_\varepsilon)_{\varepsilon>0}$ converges to $\mathcal{K}(\mu,\nu)$.

In Subsection 3.3, after introducing the notion of (large) weak multiplicativity (Definition 3.10), we prove that weak multiplicativity is stable with respect to weak convergence of transport plans (Proposition 3.14). We then apply this stability result to prove that the cluster points of $(\pi_\varepsilon)_{\varepsilon>0}$ are weakly multiplicative (Theorem 3.16).

In Subsection 3.4, after introducing "strict" versions of the reinforced stochastic order, weak multiplicativity, and strong multiplicativity, we prove that, if the (non-equal) pairs of components of (μ,ν) are singular, then π_ε converges to $\mathcal{K}(\mu,\nu)$ as $\varepsilon \to 0^+$. Notably, this encompasses the case of semi-discrete pairs of measures.

In Appendix A, we present a refinement of the decomposition constructed in Section 2.1 (Theorem A.4). This decomposition, obtained by adding mass to the diagonal part of our previous decomposition, is no longer related to the entropy minimization problem, but is preferable in the sense that it has a larger diagonal part.

In Appendix B, by applying another Strassen-type theorem of Kellerer (Proposition B.3), we prove that the fixed part of our refined decomposition is maximal (Lemma B.4). We then use our characterization of barrier points to prove an optimality property of the family of "concentration squares" (Proposition B.9).

In Appendix C, we prove optimality results for our both decompositions. After defining a class $\mathcal{A}^*$ of admissible decompositions that is broader than the class introduced in Section 1.3 (Definition C.2), we introduce a relation $\preceq_{\mathcal{A}^*}$ on the set of admissible decompositions (Definition C.7). Next, we prove that $\preceq_{\mathcal{A}^*}$ defines a partial order on $\mathcal{A}^*$ and that our refined decomposition is a minimal element for $\mathcal{A}^*$ (Theorem C.15). Finally we establish that the non-refined decomposition is the minimal element of $\mathcal{A}$ (Theorem C.17).

1.4 General notations

1. For all $(a,b) \in \mathbb{Z}^2$, we define $[\![a,b]\!] = [a,b] \cap \mathbb{Z}$.

2. Any Polish space E is equipped with its Borel σ-algebra $\mathcal{B}(E)$. We denote by $\mathcal{P}(E)$, $\mathcal{M}_+(E)$, and $\mathcal{M}_+^\sigma(E)$ the set of probability measures, finite positive measures, and σ-finite positive measures on $(E,\mathcal{B}(E))$, respectively. For all $\gamma \in \mathcal{M}_+^\sigma(E)$ and $A \in \mathcal{B}(E)$, we define the restriction of γ to A as $\gamma_{\llcorner A} : B \in \mathcal{B}(E) \mapsto \gamma(A \cap B)$. We will say that γ is concentrated on $A \in \mathcal{B}(E)$ if $\gamma(A^c) = 0$. We will often need to consider the support of the measure, denoted as $\mathrm{spt}(\gamma)$ of a measures γ. Recall that $\mathrm{spt}(\gamma)$ is a closed set, γ is concentrated on $\mathrm{spt}(\gamma)$, and every neighbourhood of an element belonging to the support has positive measure. The pushforward notation will often be needed: if E' is another Polish space, f is measurable, and $f : E \to E'$ and $\gamma \in \mathcal{M}_+^\sigma(E)$, we denote by $f_\# \gamma$ the pushforward measure of γ by f, defined by $f_\#\gamma : B \in \mathcal{B}(E') \mapsto \gamma(f^{-1}(B))$.

3. In the case where $E = \mathbb{R}$, we introduce specific notation. We begin with extremal point of the support: define $s_\gamma = \inf(\mathrm{spt}(\gamma))$ and $S_\gamma = \sup(\mathrm{spt}(\gamma)) \in [-\infty,+\infty]$ (the letter "s" stands for support). Next, let $F_\gamma^+ : t \in \mathbb{R} \mapsto \gamma(]-\infty,t]) \in \mathbb{R}^+$ and $F_\gamma^- : t \in \mathbb{R} \mapsto \gamma(]-\infty,t[) \in \mathbb{R}^+$ denote the cumulative distribution functions of γ, respectively. Recall that F_γ^- is left-continuous, F_γ^+ is right-continuous and $F_\gamma^- \leq F_\gamma^+$. Finally, we denote by $\mathrm{Atom}(\gamma) = \{x \in \mathbb{R} \;;\; \gamma(\{x\}) > 0\}$ the set of atoms of γ: we shall say that γ is atomless when $\mathrm{Atom}(\gamma) = \emptyset$.

4. The following notation concern specific spaces of measures. We define $\mathcal{M}_+^2$ as $\mathcal{M}_+^2 = \{(\mu,\nu) \in \mathcal{M}_+(\mathbb{R})^2 \;;\; \mu(\mathbb{R}) = \nu(\mathbb{R}) > 0\}$: observe that $\mathcal{M}_+^2$ is a proper subset of $\mathcal{M}_+(\mathbb{R})^2$. We also define $\mathcal{M}_1(\mathbb{R}) = \left\{\mu \in \mathcal{M}_+(\mathbb{R}) \;;\; \int_\mathbb{R} |x| \,\mathrm{d}\mu(x) < +\infty\right\}$. Given two measures $\mu,\nu \in \mathcal{M}_+(\mathbb{R})$, we define the set of transport plans from μ to ν by $\mathrm{Marg}(\mu,\nu) = \left\{\pi \in \mathcal{M}_+(\mathbb{R}^2) \;;\; p_{1\#}\pi = \mu \text{ and } p_{2\#}\pi = \nu\right\}$, where $p_1 : (x,y) \in \mathbb{R}^2 \mapsto x \in \mathbb{R}$, $p_2 : (x,y) \in \mathbb{R}^2 \mapsto y \in \mathbb{R}$ stand for the projections from $\mathbb{R}^2$ to $\mathbb{R}$. Recall that the cost function for transport plans $J : \mathrm{Marg}(\mu,\nu) \to \mathbb{R}_+$ associated to the distance is defined

by $J(\pi) = \int_{\mathbb{R}^2} |x-y| \, \mathrm{d}\pi(x,y)$: its minimum, also known as the 1-Wasserstein distance between μ and ν, is denoted by $W_1(\mu,\nu) = \min(J)$. The set $\mathcal{O}(\mu,\nu)$ of optimal transport plans from μ to ν is then defined as the set of minimizers of J, that is, $\mathcal{O}(\mu,\nu) = \{\pi \in \mathrm{Marg}(\mu,\nu) \; ; \; J(\pi) = W_1(\mu,\nu)\}$. Note that $(\mu,\nu) \in \mathcal{M}_1(\mathbb{R}) \times \mathcal{M}_1(\mathbb{R})$ implies that $W_1(\mu,\nu) < +\infty$ and $W_1(\mu,\nu) < +\infty$ implies $(\mu,\nu) \in \mathcal{M}_+^2$. We denote be $\mathfrak{C}(\mu,\nu)$ the set of cyclically monotone transport plan for μ to ν (see Definition 2.3).

5. The following notation for half-planes, adopted from Kellerer, will be useful: we define $\mathrm{F} = \{(x,y) \in \mathbb{R}^2 \; ; \; x \leq y\}$, $\tilde{\mathrm{F}} = \{(x,y) \in \mathbb{R}^2 \; ; \; y \leq x\}$, $\mathrm{G} = \{(x,y) \in \mathbb{R}^2 \; ; \; x < y\}$, $\tilde{\mathrm{G}} = \{(x,y) \in \mathbb{R}^2 \; ; \; x > y\}$ and $\mathrm{D} = \{(x,x) \; ; \; x \in \mathbb{R}\}$. Given $\mathrm{H} \in \mathcal{B}(\mathbb{R}^2)$, we define $\mathrm{Marg}_{\mathrm{H}}(\mu,\nu) = \{\pi \in \mathrm{Marg}(\mu,\nu) \; ; \; \pi(\mathrm{H}^c) = 0\}$.

2 The decomposition result.

2.1 The structure result of Di Marino–Louet

Before presenting our decomposition, we examine the result by Simone Di Marino and Jean Louet regarding the structure of optimal transport plans, when $(\mu,\nu) \in \mathcal{P}(\mathbb{R})^2$ is a pair of atomless marginals [10, Proposition 3.1]. Although their result is not originally presented as a decomposition of the space $\mathcal{O}(\mu,\nu)$ in direct sums of subspaces, we shall see that the proof of this result leads to our decomposition result (for atomless measures). When μ and ν are atomless, the functions F_μ^+ and F_ν^+ are continuous; therefore, the sets $\{F_\mu^+ > F_\nu^+\} := \{x \in \mathbb{R} \; ; \; F_\mu^+(x) > F_\nu^+(x)\}$ and $\{F_\mu^+ < F_\nu^+\} := \{x \in \mathbb{R} \; ; \; F_\mu^+(x) < F_\nu^+(x)\}$ are open and consists of a countable family of open, connected components. Let $(]a_k^+, b_k^+[)_{k\in\mathcal{K}^+}$ and $(]a_k^-, b_k^-[)_{k\in\mathcal{K}^-}$ denote the connected components of $\{F_\mu^+ > F_\nu^+\}$ and $\{F_\mu^+ < F_\nu^+\}$, respectively. Using our notation, the result by Di Marino and Louet can be stated as follows.

Proposition 2.1 (Structure result by Di Marino–Louet). *Assume that μ and ν in $\mathcal{P}(\mathbb{R})$ are atomless and have compact support. For all $\pi \in \mathcal{O}(\mu,\nu)$, there exists a set $S_\pi \in \mathcal{B}(\mathbb{R}^2)$ such that $\pi(S_\pi) = 1$ and:*

1. $[S_\pi \cap (\{F_\mu^+ = F_\nu^+\} \times \mathbb{R})] \subset \mathrm{D}$;
2. *For all $k \in \mathcal{K}^+$, $[S_\pi \cap (]a_k^+, b_k^+[\times \mathbb{R})] \subset \mathrm{F} \cap (\mathbb{R} \times]a_k^+, b_k^+[)$;*
3. *For all $k \in \mathcal{K}^-$, $[S_\pi \cap (]a_k^-, b_k^-[\times \mathbb{R})] \subset \tilde{\mathrm{F}} \cap (\mathbb{R} \times]a_k^-, b_k^-[)$.*

Interpreting a transport plan $\pi \in \mathrm{Marg}(\mu,\nu)$ as a mass displacement from the distribution μ to the distribution ν, a point (x,y) represents a transport path and $\pi(x,y)$ the amount of mass displaced from x to y. Observe that F represents forward transports, $\tilde{F}$ represents backward transports, and D corresponds to fixed mass. From this perspective, Point 1 of Proposition 2.1 states that the mass on the set $\{F_\mu^+ = F_\nu^+\}$ remains fixed under any optimal transport for the distance cost. Point 2 (resp. Point 3) of Proposition 2.1 states that for an optimal displacement, the mass of $\mu(dx)$ on a connected component of $\{F_\mu^+ > F_\nu^+\}$ (resp. $\{F_\mu^+ < F_\nu^+\}$) remains within the same component and moves forward (resp. backward). Interpreting a transport as a measure on the plan $\mathbb{R}^2$, Proposition 2.1 can be viewed as an assertion about the concentration region of elements of $\mathcal{O}(\mu,\nu)$. Specifically, every element of $\mathcal{O}(\mu,\nu)$ is concentrated on the following disjoint union of triangles and diagonal parts:

$$\biguplus_{k\in\mathcal{K}^+} (\mathrm{F} \cap]a_k^+, b_k^+[^2) \biguplus \biguplus_{k\in\mathcal{K}^-} (\tilde{\mathrm{F}} \cap]a_k^-, b_k^-[^2) \biguplus \left(\mathrm{D} \cap \{F_\mu^+ = F_\nu^+\}^2\right).$$

Example 2.2. Define $\mu = \mathbb{1}_{]0,1[} \cdot \mathcal{L}^1$ and $\nu = \left(2\mathbb{1}_{]1/8,1/4[} + 2\mathbb{1}_{]3/8,1/2[} + \mathbb{1}_{]1/2,3/4[} + 2\mathbb{1}_{]3/4,7/8[}\right) \cdot \mathcal{L}^1$. We have $\{F_\mu^+ > F_\nu^+\} =]0,1/4[\cup]1/4,1/2[$, $\{F_\mu^+ < F_\nu^+\} =]3/4,1[$ and $\{F_\mu^+ = F_\nu^+\} = [1/2,3/4]$. According to Proposition 2.1, for every optimal displacement from μ to ν the mass moves as illustrated in Figure 1 and is concentrated in the orange region shown in Figure 3.

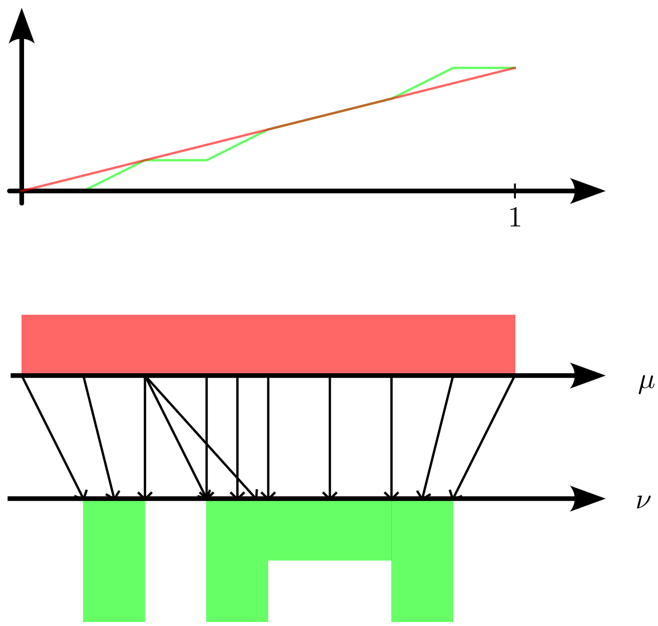

Figure 1: Representation of F_μ^+ (in red) and F_ν^+ (in green) and mass displacement of elements of $\mathcal{O}(\mu,\nu)$

We propose a brief alternative proof of Proposition 2.1, which will provide the opportunity to introduce notions used throughout the article. Our proof relies on cyclical monotonicity, and does not use the construction and manipulation of an explicit Kantorovich potential that appears in the proof of Proposition 3.1 in [10]. Cyclical monotonicity is a standard tool in optimal transport theory: for further details, we refer to the monograph [30, Chapter 5].

Definition 2.3. A set $\Gamma \subset \mathbb{R}^2$ is said to be cyclically monotone if, for every $n \in \mathbb{N}^*$ and $((x_i,y_i))_{i\in[\![1,n]\!]} \in \Gamma^n$, $\sum_{i=1}^n |y_i - x_i| \leq \sum_{i=1}^n |y_{i+1} - x_i|$ (with the convention $y_{n+1} = y_1$). A measure $\pi \in \mathcal{M}_+(\mathbb{R}^2)$ is said to be cyclically monotone if it is concentrated on a cyclically monotone set.

Notation 2.4. For any pair $(\gamma_1,\gamma_2) \in \mathcal{M}_+(\mathbb{R}^2)$, define $\mathfrak{C}(\gamma_1,\gamma_2)$ as the set of cyclically monotone measures in $\mathrm{Marg}(\gamma_1,\gamma_2)$.

Remark 2.5. For any pair $(\gamma_1,\gamma_2) \in \mathcal{M}_+^2$, $\mathfrak{C}(\gamma_1,\gamma_2)$ is non-empty. Indeed, denoting by G_{γ_1} and G_{γ_2} the quantile functions of γ_1 and γ_2, respectively, the reader may verify that the monotone transport plan $(G_{\gamma_1}, G_{\gamma_2})_\# \left(\mathbb{1}_{[0,1]} \cdot \mathcal{L}^1\right)$ belongs to $\mathfrak{C}(\gamma_1,\gamma_2)$.

The following result is classical; see [30, Theorem 5.10] for a proof with more general cost functions.

Theorem 2.6. *Assume that* $(\gamma_1,\gamma_2) \in \mathcal{M}_+(\mathbb{R})^2$ *and* $W_1(\gamma_1,\gamma_2) < +\infty$. *Then* $\mathcal{O}(\gamma_1,\gamma_2) = \mathfrak{C}(\gamma_1,\gamma_2)$.

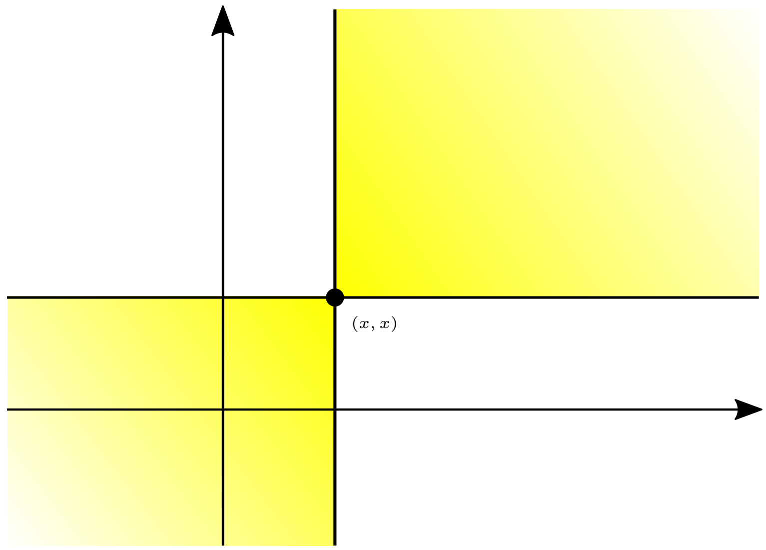

Figure 2: Concentration region of optimal transport plans when $x \in \mathbb{R}$ is a barrier point.

The following remark states that a sub-measure of a cyclically monotone transport plan remains cyclically monotone for its own marginals. We next derive a corresponding result for optimal transport plans. For any measurable space $(E, \mathcal{T})$ and any pair $(\gamma_1, \gamma_2) \in \mathcal{M}_+(E)^2$, we write $\gamma_1 \leq \gamma_2$ if, for all $A \in \mathcal{T}$, $\gamma_1(A) \leq \gamma_2(A)$.

Remark 2.7 (Cyclicity and optimality of sub-measures)**.**

1. Consider $(\gamma_1, \gamma_2) \in \mathcal{M}_+(\mathbb{R})^2$, $\pi \in \mathfrak{C}(\gamma_1, \gamma_2)$ and $\pi^* \in \mathcal{M}(\mathbb{R}^2)$. If $\pi^* \leq \pi$, then $\pi^* \in \mathfrak{C}(p_{1\#}\pi^*, p_{2\#}\pi^*)$. Indeed, if Γ is a cyclically monotone set on which π is concentrated, then π^* is also concentrated on Γ.

2. Consider $(\gamma_1, \gamma_2) \in \mathcal{M}_+(\mathbb{R})^2$ such that $W_1(\gamma_1, \gamma_2) < +\infty$, $\pi \in \mathcal{O}(\gamma_1, \gamma_2)$, and $\pi^* \in \mathcal{M}_+(\mathbb{R}^2)$. If $\pi^* \leq \pi$, then $\pi^* \in \mathcal{O}(p_{1\#}\pi^*, p_{2\#}\pi^*)$. Indeed, since

$$W_1\left(p_{1\#}\pi^*, p_{2\#}\pi^*\right) \leq \int_{\mathbb{R}^2} |y-x| \, \mathrm{d}\pi^*(x,y) \leq \int_{\mathbb{R}^2} |y-x| \, \mathrm{d}\pi(x,y) = W_1(\gamma_1, \gamma_2) < +\infty,$$

by Theorem 2.6, $\mathfrak{C}(\gamma_1, \gamma_2) = \mathcal{O}(\gamma_1, \gamma_2)$ and $\mathfrak{C}(p_{1\#}\pi^*, p_{2\#}\pi^*) = \mathcal{O}(p_{1\#}\pi^*, p_{2\#}\pi^*)$. Thus, $\pi^* \in \mathcal{O}(p_{1\#}\pi^*, p_{2\#}\pi^*)$ directly follows from the previous point.

We now introduce the key concept of barrier point, which plays a central role in proving Proposition 2.1 and is used throughout the remainder of the article.

Definition 2.8 (Barrier points)**.** For all $x \in \mathbb{R}$, define $\mathrm{C}(x) = (]-\infty, x[\times]x, +\infty[) \cup (]x, +\infty[\times]-\infty, x[)$. We say that $x \in \mathbb{R}$ is a barrier point for $(\mu, \nu) \in \mathcal{M}_+^2$ if, for all $\pi \in \mathfrak{C}(\mu, \nu)$, we have $\pi(\mathrm{C}(x)) = 0$. We denote by $\mathcal{B}(\mu, \nu)$ the set of barrier points for (μ, ν).

Intuitively, a barrier point of (μ, ν) is a point through which no optimal transport from μ to ν can move mass. Visually, $x \in \mathcal{B}(\mu, \nu)$ if every $\pi \in \mathfrak{C}(\mu, \nu)$ is concentrated in the yellow region shown in Figure

2. The next result provides a sufficient condition for a point to be a barrier point. Its proof relies on cyclical monotonicity and is postponed to the next subsection, following Proposition 2.15, which presents a more general statement.

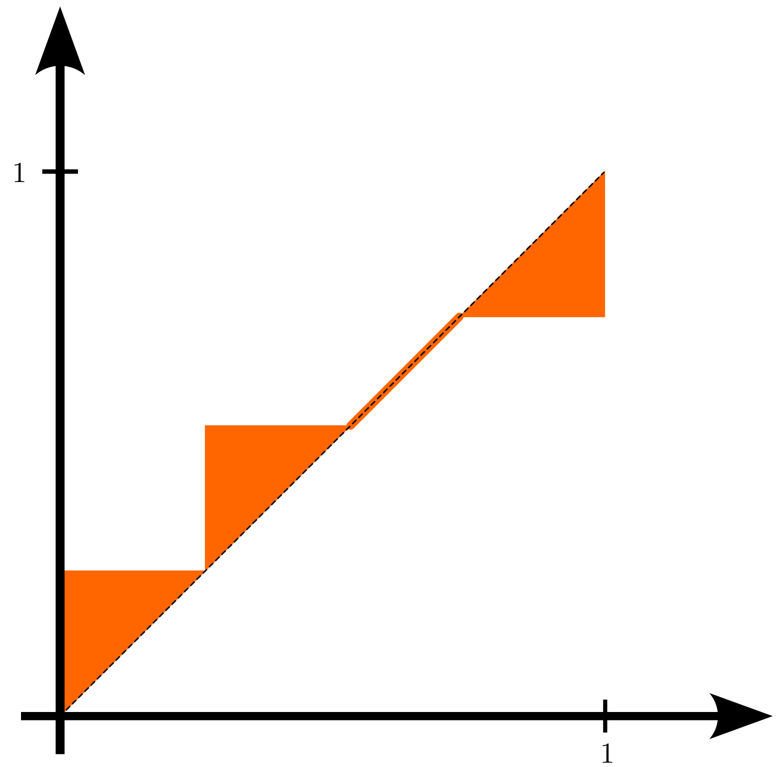

Figure 3: Illustration of the concentration area of optimal transport plans.

Proposition 2.9. *Let $\mu, \nu \in \mathcal{M}_+(\mathbb{R})$ satisfy $W_1(\mu, \nu) < +\infty$. The inclusion $\{F_\mu^+ = F_\nu^+\} \subset \mathcal{B}(\mu, \nu)$ is satisfied.*

This proposition is an important step toward the proof of Proposition 2.1. To see this, observe that in Example 2.2, 1/4 and 1/2 lie in $\{F_\mu^+ = F_\nu^+\}$, and thus both belong to $\mathcal{B}(\mu, \nu)$. Hence, the mass cannot exit the connected component $]1/4, 1/2[$, which is precisely the content of Point 2 of Proposition 2.1. To prove the "moving forward" part of Point 2, we use the fact that, for measures in the stochastic order, a transport is optimal for the distance cost if and only if all the mass moves forward (see Proposition 2.11).

Definition 2.10 (Stochastic Order)**.** We say that $\gamma_1 \in \mathcal{M}_+(\mathbb{R})$ is smaller than $\gamma_2 \in \mathcal{M}_+(\mathbb{R})$ in the stochastic order if, for all bounded increasing function $f : \mathbb{R} \to \mathbb{R}$, $\int_\mathbb{R} f \; \mathrm{d}\gamma_1 \leq \int_\mathbb{R} f \; \mathrm{d}\gamma_2$. In this case, we write $\gamma_1 \leq_{\mathrm{st}} \gamma_2$.

It is well known that $\gamma_1 \leq_{\mathrm{st}} \gamma_2$ if and only if both $\gamma_1(\mathbb{R}) = \gamma_2(\mathbb{R})$ and $F_{\gamma_1}^+ \geq F_{\gamma_2}^+$. We now recall the Strassen-type theorem for the order $\leq_{\mathrm{st}}$, along with the characterization of the set of transport plans for measures in the stochastic order. We refer the reader to [15, Proposition 4.1] for a proof of the first point. The second point will be proved in Subsection 2.2, immediately following Proposition 2.15.

Proposition 2.11. *Consider $(\gamma_1, \gamma_2) \in \mathcal{M}_+(\mathbb{R})^2$.*

1. ***Strassen-type theorem for*** $\leq_{\mathrm{st}}$: $\gamma_1 \leq_{\mathrm{st}} \gamma_2 \iff \mathrm{Marg}_{\mathrm{F}}(\gamma_1, \gamma_2) \neq \emptyset$.
2. ***Characterization of optimality:*** *If $\gamma_1 \leq_{\mathrm{st}} \gamma_2$, then the equality $\mathfrak{C}(\gamma_1, \gamma_2) = \mathrm{Marg}_{\mathrm{F}}(\gamma_1, \gamma_2)$ holds. When $W_1(\gamma_1, \gamma_2) < +\infty$, we obtain $\mathcal{O}(\gamma_1, \gamma_2) = \mathrm{Marg}_{\mathrm{F}}(\gamma_1, \gamma_2)$.*

We prove now the structure result of Di Marino–Louet.

Proof of Proposition 2.1. Let $\mathrm{C}(x)$ be defined as in Definition 2.8, and define $B = \left(\bigcup_{k\in\mathcal{K}^+}\{a_k^+, b_k^+\}\right) \cup \left(\bigcup_{k\in\mathcal{K}^-}\{a_k^-, b_k^-\}\right)$, $S_\pi = \mathrm{spt}(\pi) \cap (B^c \times B^c)$. Since μ, ν are atomless and B is countable, $\pi(S_\pi) = 1$. To prove Point 2, fix $k \in \mathcal{K}_+$. Since a_k^+ and b_k^+ belong to $\{F_\mu^+ = F_\nu^+\}$, by Proposition 2.9, it follows that they also belong to $\mathcal{B}(\mu, \nu)$. Thus $\pi\left(\mathrm{C}(a_k^+) \cup \mathrm{C}(b_k^+)\right) = 0$, which implies $\mathrm{spt}(\pi) \subset [\mathrm{C}(a_k^+) \cup \mathrm{C}(b_k^+)]^c =]-\infty, a_k^+]^2 \cup [a_k^+, b_k^+]^2 \cup [b_k^+, +\infty[^2$, and therefore:

$$\begin{cases} S_\pi \cap (]a_k^+, b_k^+[\times\mathbb{R}) \subset]a_k^+, b_k^+[^2 \subset \mathbb{R}\times]a_k^+, b_k^+[\\ S_\pi \cap (\mathbb{R}\times]a_k^+, b_k^+[) \subset]a_k^+, b_k^+[^2 \subset]a_k^+, b_k^+[\times\mathbb{R} \end{cases}. \tag{2}$$

To complete the proof of Point 2, it remains to show that $S_\pi \cap (]a_k^+, b_k^+[\times\mathbb{R})$ is contained in F. Let us define $\pi_k^+ = \pi_{\llcorner]a_k^+, b_k^+[^2}$, $\mu_k^+ = \mu_{\llcorner]a_k^+, b_k^+[}$ and $\nu_k^+ = \nu_{\llcorner]a_k^+, b_k^+[}$. As π is concentrated on $\mathrm{spt}(\pi)$, the first inclusion of the first line and the second inclusion of the second line of Equation (2) imply that $\pi_k^+ = \pi_{\llcorner]a_k^+, b_k^+[\times\mathbb{R}}$. Similarly, the second inclusion of the first line and the first inclusion of the second line of Equation 2 yield $\pi_k^+ = \pi_{\llcorner \mathbb{R}\times]a_k^+, b_k^+[}$. Thus, $\pi_k^+ \in \mathrm{Marg}(\mu_k^+, \nu_k^+)$. By Remark 2.7, it follows that $\pi_k^+ \in \mathcal{O}(\mu_k^+, \nu_k^+)$. Since $\mu_k^+(\mathbb{R}) = F_{\mu_k^+}(b_k^+) - F_{\mu_k^+}(a_k^+) = F_{\nu_k^+}(b_k^+) - F_{\nu_k^+}(a_k^+) = \nu_k^+(\mathbb{R})$ and $F_{\mu_k^+} - F_{\nu_k^+} = \mathbb{1}_{]a_k^+, b_k^+[}(F_\mu^+ - F_\nu^+) \geq 0$, we conclude that $\mu_k^+ \leq_{\mathrm{st}} \nu_k^+$. By Proposition 2.11, we obtain $\pi_k^+(\mathrm{F}^c) = 0$, i.e., $\pi(\mathrm{F}^c \cap (]a_k^+, b_k^+[\times\mathbb{R})) = 0$. Thus, $\mathrm{spt}(\pi) \subset \mathrm{F} \cup (]a_k^+, b_k^+[\times\mathbb{R})^c$, which implies $S_\pi \cap (]a_k^+, b_k^+[\times\mathbb{R}) \subset \mathrm{F}$, thereby completing the proof of Point 2 in Proposition 2.1. Since the proof of Point 3 is similar, it remains to prove Point 1. Consider $(x, y) \in [S_\pi \cap (\{F_\mu^+ = F_\nu^+\} \times \mathbb{R})]$ and suppose $x < y$. As $x \notin B$, the situations $]x, y[\subset \{F_\mu^+ < F_\nu^+\}$ and $]x, y[\subset \{F_\mu^+ > F_\nu^+\}$ are excluded. Since F_μ^+ and F_ν^+ are continuous, there exists $z \in]x, y[$ such that $F_\mu^+(z) = F_\nu^+(z)$. Proposition 2.9 implies $\pi(\mathrm{C}(z)) = 0$, whereas $(x, y) \in \mathrm{C}(z)$. This contradicts $(x, y) \in \mathrm{spt}(\pi)$, therefore $x \geq y$. Similarly $x \leq y$, which proves $x = y$ and completes the proof of Point 1. □

Remark 2.12. Our alternative proof of the statement of Di Marino and Louet requires only $W_1(\mu, \nu) < +\infty$, rather than compact support of the measures. In Theorem 2.14, we complete this result by a decomposition result. We will generalize Theorem 2.14 for marginals that may have atoms in Theorem 2.36 and Theorem A.4 (and $\mathfrak{C}(\mu, \nu)$ instead of $\mathcal{O}(\mu, \nu)$).

We fix the notation used in the previous proof concerning the components of the marginals.

Notation 2.13 (Components of the marginals)**.** For all $k \in \mathcal{K}^+$ (resp. $k \in \mathcal{K}^-$), define $\mu_k^+ = \mu_{\llcorner]a_k^+, b_k^+[}$, $\nu_k^+ = \nu_{\llcorner]a_k^+, b_k^+[}$ (resp. $\mu_k^- = \mu_{\llcorner]a_k^-, b_k^-[}$, $\nu_k^- = \nu_{\llcorner]a_k^-, b_k^-[}$). Then, define $\mu^= = \mu_{\llcorner \{F_\mu^+ = F_\nu^+\}}$ and $\nu^= = \nu_{\llcorner \{F_\mu^+ = F_\nu^+\}}$.

Theorem 2.14. *Consider a pair $(\mu, \nu) \in \mathcal{P}(\mathbb{R})^2$ of atomless measures with compact support. Using the notation introduced in Notation 2.13, we define $\mathfrak{P} = \prod_{k\in\mathcal{K}^+} \mathcal{O}(\mu_k^+, \nu_k^+) \times \prod_{k\in\mathcal{K}^-} \mathcal{O}(\mu_k^-, \nu_k^-) \times \mathcal{O}(\mu^=, \nu^=)$.*

1. *For all $k \in \mathcal{K}^+$ (resp. $\mathcal{K}^-$), we have $\mu_k^+ \leq_{\mathrm{st}} \nu_k^+$ (resp. $\nu_k^- \leq_{\mathrm{st}} \mu_k^-$), and $\mu^= = \nu^=$.*
2. *The map $\varphi : \mathfrak{P} \to \mathcal{O}(\mu, \nu)$ defined by*

$$\varphi((\pi_k^+)_{k\in\mathcal{K}^+}, (\pi_k^+)_{k\in\mathcal{K}^+}, \pi^=) = \sum_{k\in\mathcal{K}^+} \pi_k^+ + \sum_{k\in\mathcal{K}^-} \pi_k^- + \pi^=$$

is a bijection. Furthermore, its inverse $\varphi^{-1} : \mathcal{O}(\mu,\nu) \to \mathfrak{P}$ is given by

$$\varphi^{-1}(\pi) = \left(\left(\pi \llcorner_{]a_k^+, b_k^+[^2} \right)_{k\in\mathcal{K}^+}, \left(\pi \llcorner_{]a_k^-, b_k^-[^2} \right)_{k\in\mathcal{K}^-}, \pi\llcorner_{\{F_\mu^+ = F_\nu^+\}^2} \right). \tag{3}$$

Proof. 1. From the proof of Proposition 2.1, for all $k \in \mathcal{K}^+$, we have $\mu_k^+ \leq_{\mathrm{st}} \nu_k^+$ and similarly, for all $k \in \mathcal{K}^-$, $\nu_k^- \leq_{\mathrm{st}} \mu_k^-$. By Point 1 of Proposition 2.1, $\pi\llcorner_{\{F_\mu^+ = F_\nu^+\}\times\mathbb{R}}$ is concentrated on D and its first marginal is $\mu^=$. Therefore, $\pi\llcorner_{\{F_\mu^+ = F_\nu^+\}\times\mathbb{R}} = (\mathrm{id}, \mathrm{id})_\# \mu^= = \pi\llcorner_{\{F_\mu^+ = F_\nu^+\}^2}$. Similarly $\pi\llcorner_{\mathbb{R}\times\{F_\mu^+ = F_\nu^+\}} = (\mathrm{id}, \mathrm{id})_\# \nu^= = \pi\llcorner_{\{F_\mu^+ = F_\nu^+\}^2}$. Hence $\mu^= = \nu^=$.

2. To prove that φ is surjective, fix $\pi \in \mathcal{O}(\mu,\nu)$. We established in the proof of Proposition 2.1 that, for all $k \in \mathcal{K}^+$, π_k^+, defined by $\pi_k^+ = \pi\llcorner_{]a_k^+, b_k^+[^2}$ belongs to $\mathcal{O}(\mu_k^+, \nu_k^+)$. Similarly, for all $k \in \mathcal{K}^-$, π_k^-, defined by $\pi_k^- = \pi\llcorner_{]a_k^-, b_k^-[^2}$ belongs to $\mathcal{O}(\mu_k^-, \nu_k^-)$. By the first point, $\pi^= := (\mathrm{id}, \mathrm{id})_\# \mu^=$ belongs to $\mathcal{O}(\mu^=, \nu^=)$. Hence $\mathcal{F}$, defined by $\mathcal{F} = ((\pi_k^+)_{k\in\mathcal{K}^+}, (\pi_k^-)_{k\in\mathcal{K}^-}, \pi^=)$ belongs to $\mathfrak{P}$. Moreover, the measure $\tilde{\pi}$, defined as $\tilde{\pi} = \sum_{k\in\mathcal{K}^+} \pi_k^+ + \sum_{k\in\mathcal{K}^-} \pi_k^- + \pi^=$ has first marginal $\sum_{k\in\mathcal{K}^+} \mu_k^+ + \sum_{k\in\mathcal{K}^-} \mu_k^- + \mu^= = \mu$ and corresponds to the restriction of π to $\biguplus_{k\in\mathcal{K}^+}(]a_k^+, b_k^+[^2) \uplus \biguplus_{k\in\mathcal{K}^-}(]a_k^-, b_k^-[^2) \uplus (\{F_\mu^+ = F_\nu^+\}^2)$. Thus, $\pi(\mathbb{R}^2) = \mu(\mathbb{R}) = \tilde{\pi}(\mathbb{R})$ and $\tilde{\pi} \leq \pi$, which implies $\tilde{\pi} = \pi$. Finally, we have shown that $\mathcal{F} \in \mathfrak{P}$ and $\varphi(\mathcal{F}) = \pi$. Therefore, φ is surjective. Note that, as $\pi \in \mathcal{O}(\mu,\nu)$ and $\mathcal{F} \in \mathfrak{P}$, integrating $(x, y) \mapsto |y - x|$ along the equality $\varphi(\mathcal{F}) = \pi$ yields $W_1(\mu,\nu) = \sum_{k\in\mathcal{K}^+} W_1(\mu_k^+, \nu_k^+) + \sum_{k\in\mathcal{K}^-} W_1(\mu_k^-, \nu_k^-)$. This ensures that φ takes values in $\mathcal{O}(\mu,\nu)$. Indeed, for every $\mathcal{F} = ((\pi_k^+)_{k\in\mathcal{K}^+}, (\pi_k^-)_{k\in\mathcal{K}^-}, \pi^=)$, we have $\varphi(\mathcal{F}) \in \mathrm{Marg}(\mu,\nu)$ and

$$\begin{aligned} J(\varphi(\mathcal{F})) &= \sum_{k\in\mathcal{K}^+} \int_{\mathbb{R}^2} |y - x| \, \mathrm{d}\pi_k^+(x,y) + \sum_{k\in\mathcal{K}^-} \int_{\mathbb{R}^2} |y - x| \, \mathrm{d}\pi_k^-(x,y) + \int_{\mathbb{R}^2} |y - x| \, \mathrm{d}\pi^=(x,y) \\ &= \sum_{k\in\mathcal{K}^+} W_1(\mu_k^+, \nu_k^+) + \sum_{k\in\mathcal{K}^-} W_1(\mu_k^-, \nu_k^-) = W_1(\mu,\nu), \end{aligned}$$

which implies $\varphi(\mathcal{F}) \in \mathcal{O}(\mu,\nu)$. The injectivity of φ and the validity of Equation (3) follow directly from the fact that $\{]a_k^+, b_k^+[^2\}_{k\in\mathcal{K}^+} \cup \{]a_k^-, b_k^-[^2\}_{k\in\mathcal{K}^-} \cup \{\{F_\mu^+ = F_\nu^+\}^2\}$ forms a collection of disjoint sets. $\square$

2.2 Generalized decomposition result

In this subsection, we aim to extend Theorem 2.14 to any pair of measures, including those with atoms. More precisely, for any pair $(\mu,\nu) \in \mathcal{M}_+^2$, we will prove a decomposition for $\mathfrak{C}(\mu,\nu)$. Note that we neither require the measures to be atomless nor that $W_1(\mu,\nu) < +\infty$, and that we decompose $\mathfrak{C}(\mu,\nu)$ instead of $\mathcal{O}(\mu,\nu)$. By Theorem 2.6, this result also covers the decomposition of $\mathcal{O}(\mu,\nu)$ when $W_1(\mu,\nu) < +\infty$. Note that $\mu(\mathbb{R}) = \nu(\mathbb{R})$ is the minimal assumption possible: otherwise $\mathrm{Marg}(\mu,\nu)$ is empty and there is nothing to investigate: this is implicitly assumed from now on, as we consider pair of marginals in $\mathcal{M}_+^2$ (see Notation 1.4, Point 4). The following result extends Proposition 2.9 and serves as a fundamental step toward our decomposition.

Proposition 2.15. *Consider $(\mu,\nu)\in\mathcal{M}_+^2$, $x\in\mathbb{R}$ and $\pi\in\mathfrak{C}(\mu,\nu)$.*

1. *If $F_\mu^+(x)\geq F_\nu^+(x)$, then $\pi(]x,+\infty[\times]-\infty,x])=0$.*
2. *If $F_\mu^-(x)\geq F_\nu^-(x)$, then $\pi([x,+\infty[\times]-\infty,x[)=0$.*
3. *If $F_\mu^+(x)\leq F_\nu^+(x)$, then $\pi(]-\infty,x]\times]x,+\infty[)=0$.*
4. *If $F_\mu^-(x)\leq F_\nu^-(x)$, then $\pi(]-\infty,x[\times[x,+\infty[)=0$.*

Proof. Assume that x satisfies $F_\mu^+(x)\geq F_\nu^+(x)$. To reach a contradiction, suppose that $\pi(]x,+\infty[\times]-\infty,x])>0$. Observe that $\pi(]-\infty,x]\times]-\infty,x])+\pi(]-\infty,x]\times]x,+\infty[)=F_\mu^+(x)\geq F_\nu^+(x)=\pi(]-\infty,x]\times]-\infty,x])+\pi(]x,+\infty[\times]-\infty,x])$, which implies $\pi(]-\infty,x]\times]x,+\infty[)\geq\pi(]x,+\infty[\times]-\infty,x])>0$. Let $\Gamma\in\mathcal{B}(\mathbb{R}^2)$ be a cyclically monotone set on which π is concentrated. Then, $\pi(\Gamma\cap(]x,+\infty[\times]-\infty,x]))=\pi(]x,+\infty[\times]-\infty,x])>0$ and $\pi(\Gamma\cap(]-\infty,x]\times]x,+\infty[))=\pi(]-\infty,x]\times]x,+\infty[)>0$. Hence, there exists $(x_1,y_1)\in(]x,+\infty[\times]-\infty,x])\cap\Gamma$ and $(x_2,y_2)\in(]-\infty,x]\times]x,+\infty[)\cap\Gamma$. Define $a=\min(x_2,y_1), b=\max(x_2,y_1), c=\min(x_1,y_2)$ and $d=\max(x_1,y_2)$. We have the inequalities $x_2\leq x<y_2$, $y_1\leq x<x_1$, and $a\leq b\leq x<c\leq d$. Thus, $|y_2-x_1|+|y_1-x_2|=(d-c)+(b-a)<(d-c)+2(c-b)+(b-a)=(d+c)-(a+b)=(y_2+x_1)-(y_1+x_2)=(y_2-x_2)+(x_1-y_1)=|y_2-x_2|+|y_1-x_1|$, which contradicts the cyclical monotonicity of Γ. Therefore, $\pi(]x,+\infty[\times]-\infty,x])=0$. The reader may verify that the arguments for the three other points follow are analogue. ☐

Remark 2.16. Observe that the proof of Proposition 2.15 only relies on the cyclical monotonicity for two pairs of points. Indeed, we used only that the set Γ on which π is concentrated satisfies:

$$\forall(x_1,y_1),(x_2,y_2)\in\Gamma, |y_1-x_1|+|y_2-x_2|\leq|y_1-x_2|+|y_2-x_1|.$$

By Points 2 and 3, we deduce $\{F_\mu^+=F_\nu^+\}=\{F_\mu^+\geq F_\nu^+\}\cap\{F_\mu^+\leq F_\nu^+\}\subset\mathcal{B}(\mu,\nu)$, thus establishing Proposition 2.9. For a more general result, see Proposition 2.52 and Remark 2.53, where the following equality is shown:

$$\mathcal{B}(\mu,\nu)=\{F_\mu^+=F_\nu^+\}\cup\{F_\mu^-=F_\nu^-\}\cup\{F_\nu^-\leq F_\mu^-<F_\mu^+\leq F_\nu^+\}\cup\{F_\mu^-\leq F_\nu^-<F_\nu^+\leq F_\mu^+\}. \quad (4)$$

We now apply Proposition 2.15 to establish Point 2 of Proposition 2.11.

Proof of Point 2 of Proposition 2.11. Assume $\pi\in\mathrm{Marg}_{\mathrm{F}}(\mu,\nu)$. For all $n\geq1$ and $((x_i,y_i))_{i\in[\![1,n]\!]}\in\mathrm{F}^n$, we have $\sum_{i=1}^n|y_i-x_i|=\sum_{i=1}^n y_i-x_i=\sum_{i=1}^n y_{i+1}-x_i\leq\sum_{i=1}^n|y_{i+1}-x_i|$. Hence F is cyclically monotone. Since $\pi(\mathrm{F}^c)=0$, $\pi\in\mathfrak{C}(\mu,\nu)$. Assume now $\pi\in\mathfrak{C}(\mu,\nu)$. Since $F_\mu^+\geq F_\nu^+$, by Point 1 of Proposition 2.15, for all $x\in\mathbb{R}$, $\pi(]x,+\infty[\times]-\infty,x])=0$. Since $\mathrm{F}^c=\cup_{r\in\mathbb{Q}}]r,+\infty[\times]-\infty,r]$, we obtain $\pi(\mathrm{F}^c)=0$. ☐

We have shown that F is a cyclically monotone set, and by a similar argument, $\tilde{\mathrm{F}}$ is also cyclically monotone.

Remark 2.17. When $(\mu,\nu) \in \mathcal{M}_1(\mathbb{R})^2$, Point 2 of Proposition 2.11 seems to be well known. In this case $\mathfrak{C}(\mu,\nu) = \mathcal{O}(\mu,\nu)$ and its proof relies on the fact that, for all $\pi \in \mathrm{Marg}(\mu,\nu)$, $\int |y-x| \ \mathrm{d}\pi \geq \left|\int y - x \ \mathrm{d}\pi(x,y)\right| = \int y \ \mathrm{d}\nu(y) - \int x \ \mathrm{d}\mu(x)$.

As in the decomposition of Di Marino Louet, we begin by partitioning the real line according the relative position of the cumulative distribution functions. Next, we define a set $\{(\mu_k^+,\nu_k^+)\}_{k\in\mathcal{K}^+} \cup \{(\mu_k^-,\nu_k^-)\}_{k\in\mathcal{K}^-} \cup \{(\mu^=,\nu^=)\}$ of measure components of (μ,ν), and finally, we prove that

$$\mathfrak{C}(\mu,\nu) = \left(\bigoplus_{k\in\mathcal{K}^+} \mathfrak{C}(\mu_k^+,\nu_k^+)\right) \oplus \left(\bigoplus_{k\in\mathcal{K}^-} \mathfrak{C}(\mu_k^-,\nu_k^-)\right) \oplus \mathfrak{C}(\mu^=,\nu^=).$$

Naturally, when dealing with measures that have atoms, the approach of Di Marino and Louet must be adapted First, for the decomposition of the real line, observe that $\{F_\mu^+ > F_\nu^+\}$ need not be open in general, making it difficult to consider the connected components. More importantly, if we consider the connected components of $\{F_\mu^+ > F_\nu^+\}$ and define the components of the marginal by restriction of μ and ν to the connected components, our decomposition would not be sufficiently fine. For instance, consider $\mu = \mathbb{1}_{]0,1[} \cdot \mathcal{L}^1 + 2\delta_1$ and $\nu = 2\mathbb{1}_{]1/2,2[} \cdot \mathcal{L}^1$. Since $\{F_\mu^+ > F_\nu^+\} =]0,2[$, the set of component would be $\{(\mu,\nu)\}$, which is associated to the equality $\mathfrak{C}(\mu,\nu) = \mathfrak{C}(\mu,\nu)$. However, noting that 1 belongs to $\{F_\mu^- = F_\nu^-\} \subset \mathcal{B}(\mu,\nu)$, the reader may verify that $\mathfrak{C}(\mu,\nu)$ can be decomposed as

$$\mathfrak{C}(\mu,\nu) = \mathfrak{C}(\mu_1,\nu_1) \oplus \mathfrak{C}(\mu_2,\nu_2),$$

where $\mu_1 = \mathbb{1}_{]0,1[} \cdot \mathcal{L}^1$, $\nu_1 = 2\mathbb{1}_{[1/2,1]} \cdot \mathcal{L}^1$, $\mu_2 = 2\delta_1$ and $\nu_2 = 2\mathbb{1}_{]1,2[} \cdot \mathcal{L}^1$. This shows that the decomposition of the real line obtained by taking the connected components of $\{F_\mu^+ > F_\nu^+\}$ is not fine enough. The following definition addresses this issue by considering the signs of both $F_\mu^+ - F_\nu^+$ *and* $F_\mu^- - F_\nu^-$, rather than solely the sign of $F_\mu^+ - F_\nu^+$.

Definition 2.18 (Components of the real line)**.** Consider $(\mu,\nu) \in \mathcal{M}_+^2$.

1. Define $E^+ = \{F_\mu^+ > F_\nu^+\} \cap \{F_\mu^- > F_\nu^-\}$, $E^- = \{F_\mu^+ < F_\nu^+\} \cap \{F_\mu^- < F_\nu^-\}$ and $E^= = \mathbb{R} \setminus (E^+ \cap E^-)$. Since $F_\mu^+ - F_\nu^+$ and $F_\mu^- - F_\nu^-$ are left-continuous and right-continuous, respectively, the reader may verify that E^- and E^+ are open, while $E^=$ is closed. Since E^+ and E^- are disjoint, $(E^+,E^-,E^=)$ forms a disjoint covering of $\mathbb{R}$.

2. Since E^+ and E^- are open sets, each admits a countable family of open connected component. Let $(]a_k^+,b_k^+[)_{k\in\mathcal{K}^+}$ and $(]a_k^-,b_k^-[)_{k\in\mathcal{K}^-}$ denote the family of connected component of E^+ and E^-, respectively.

3. Let B_l^+ denote the set $\left\{a_k^+ \ ; \ k \in \mathcal{K}^+\right\}$ of the left boundary points of connected components of E^+. Similarly, define $B_r^+ = \left\{b_k^+ \ ; \ k \in \mathcal{K}^+\right\}$, $B_l^- = \left\{a_k^- \ ; \ k \in \mathcal{K}^-\right\}$ and $B_r^- = \left\{b_k^- \ ; \ k \in \mathcal{K}^-\right\}$. Then, we denote by $B_l = B_l^+ \cup B_l^-$, $B_r = B_r^+ \cup B_r^-$ and $B = B_l \cup B_r$ the sets of left boundaries, right boundaries and overall boundaries of the connected components of E^+ and E^-, respectively. Since these connected components form a countable family, B is countable. Moreover, note that $B \subset E^=$.

Example 2.19. We introduce the following example, which will be referenced multiple times. Define

$$\begin{cases}\mu = \mathbb{1}_{[0,1]}\cdot\mathcal{L}^1 + \delta_1 + \mathbb{1}_{[1,2]}\cdot\mathcal{L}^1 + \delta_2 + \mathbb{1}_{[2,3]}\cdot\mathcal{L}^1 + \delta_3 + \mathbb{1}_{[3,4]}\cdot\mathcal{L}^1 + \delta_4 + \mathbb{1}_{[4,5]}\cdot\mathcal{L}^1 + \delta_5 + \mathbb{1}_{[5,6]}\cdot\mathcal{L}^1\\ \nu = \mathbb{1}_{[\frac{1}{2},1]}\cdot\mathcal{L}^1 + \delta_1 + \mathbb{1}_{[1,2]}\cdot\mathcal{L}^1 + 2\delta_2 + \frac{1}{2}\mathbb{1}_{[2,3]}\cdot\mathcal{L}^1 + \frac{3}{2}\delta_3 + \frac{1}{2}\mathbb{1}_{[3,4]}\cdot\mathcal{L}^1 + \delta_4 + \mathbb{1}_{[4,5]}\cdot\mathcal{L}^1 + \delta_5 + \mathbb{1}_{[5,6]}\cdot\mathcal{L}^1\end{cases}. \tag{5}$$

Then, we obtain $E^+ =]0,2[$, $E^- =]2,3[\cup]3,4[$, and $E^= =]-\infty,0]\cup\{2,3\}\cup[4,+\infty[$. Hence, the boundary sets are given by $B_l^+ = \{0\}, B_l^- = \{2,3\}, B_r^+ = \{2\}, B_r^- = \{3,4\}, B_l = \{0,2,3\}, B_r = \{2,4\}$, and $B = \{0,2,3,4\}$.

Remark 2.20.

1. Observe that the dependence of E^+, E^-, $E^=$, a_k^+, b_k^+, B_l^+, B_l^-, B_r^+, B_r^-, B_l, and B_r on (μ,ν) has been omitted here for clarity. We also write $\mathcal{B}$ instead of $\mathcal{B}(\mu,\nu)$ (see Definition 2.8). In cases of ambiguity, we make the dependence explicit — for example by writing $E^+(\mu,\nu)$ instead of E^+.

2. Our decomposition is symmetric: reversing the transport from ν to μ conserves the component, up to a sign change. More precisely, $E^\pm(\mu,\nu) = E^\mp(\nu,\mu)$, $E^=(\mu,\nu) = E^=(\nu,\mu)$, $B_l^\pm(\mu,\nu) = B_l^\mp(\nu,\mu)$, $B_r^\pm(\mu,\nu) = B_r^\mp(\nu,\mu)$, $B_l(\mu,\nu) = B_l(\nu,\mu)$, $B_r(\mu,\nu) = B_r(\nu,\mu)$ and $B(\mu,\nu) = B(\nu,\mu)$.

3. Note that (B_l^+, B_l^-, B_l^c) and (B_r^+, B_r^-, B_r^c) are two families of disjoint subsets of $\mathbb{R}$. However, all their nine pairwise intersection – namely $B_l^+\cap B_r^+, B_l^+\cap B_r^-, B_l^-\cap B_r^+, B_l^-\cap B_r^-, B_l^+\setminus B_r, B_l^-\setminus B_r, B_r^+\setminus B_l, B_r^-\setminus B_l$ and $B_l^c\cap B_r^c$ – can be non-empty in general. For example, in Example 2.19, we have $0\in B_l^+\cap B_r^c$, $2\in B_l^-\cap B_r^+$, $3\in B_r^-\cap B_l^-$, and $4\in B_r^-\cap B_l^c$.

4. Note that B represents the set of boundaries of connected components of $E^+\cup E^-$, but does not coincide with the topological boundary of that set. For instance, consider $\mu = \sum_{n\geq 0}\frac{1}{2^{n+1}}\delta_{1-\frac{1}{2^n}}$ and $\nu = \mathbb{1}_{[0,1]}\cdot\mathcal{L}^1$, where the set $E^+(\mu,\nu)$ is given by $E^+(\mu,\nu) = \cup_{n\geq 0}\left]1-\frac{1}{2^n}, 1-\frac{1}{2^{n+1}}\right[$. Therefore, 1 lies in the topological boundary of E^+, but is not an element of B.

5. The relative position of F_μ^- and F_ν^+ is not prescribed in $E^+(\mu,\nu)$. For instance, in Example 2.19, $1/2\in E^+(\mu,\nu)\cap\{F_\mu^- > F_\nu^+\}$ whereas $1\in E^+(\mu,\nu)\cap\{F_\mu^- < F_\nu^+\}$.

Lemma 2.21. *Consider $(\mu,\nu)\in\mathcal{M}_+^2$. Then, $E^=(\mu,\nu)\subset\mathcal{B}(\mu,\nu)$.*

Proof. We have $E^= = (E^-\cup E^+)^c = (E^-)^c\cap(E^+)^c = [\{F_\mu^-\geq F_\nu^-\}\cup\{F_\mu^+\geq F_\nu^+\}]\cap[\{F_\mu^-\leq F_\nu^-\}\cup\{F_\mu^+\leq F_\nu^+\}]$. By applying Proposition 2.15, it follows that $E^=\subset\mathcal{B}$. □

The inclusion $\mathcal{B}(\mu,\nu)\subset E^=$ is also satisfied as we shall prove in Proposition 2.52. In the general case, the components of μ and ν cannot be defined solely by restriction to the connected components of E^+, E^- and $E^=$. For instance, consider $\mu = \delta_{-1} + 6\delta_0 + \delta_1$ and $\nu = \delta_{-3} + 2\delta_{-2} + 3\delta_0 + \delta_2 + \delta_3$. We have $E^- =]-3,0[$ and $E^+ =]0,3[$ The reader may verify that that the mass at point 0 should split as follows: $(F_\nu^-(1) - F_\mu^-(1)) = 2$ units of mass should go from 0 to to $[2,0[$, $(F_\mu^+(1) - F_\nu^+(1)) = 1$ unit go from 0 to $]0,3]$ and 3 units of 0 does not move. The following definition of the marginal components accounts for the possibility that mass at boundary points may be shared among a fixed part and mass moving to the left or right; conversely boundary points can receive mass from the left and right. The amount of mass allocated to each part is expressed using differences of the cumulative distribution functions, consistently with our illustrative example. In the following definition and throughout the paper, when referring to intervals

$[a,b]$, $]a,b]$ or $[a,b[$, where a and b that may be infinite, we use the following convention: $[-\infty,x[=]-\infty,x[$, $]x,+\infty]=]x,+\infty[$ for all $x\in\mathbb{R}$, and $[-\infty,+\infty]=\mathbb{R}$. For every $\gamma\in\mathcal{M}_+(\mathbb{R})$, we also set $F^{\pm}_\gamma(-\infty)=0$ and $F^{\pm}_\gamma(+\infty)=1$.

Definition 2.22 (Components of the marginals)**.** Consider $(\mu,\nu)\in\mathcal{M}_+^2$.

1. For all $k\in\mathcal{K}^+$, define $\mu_k^+=(F_\mu^+(a_k^+)-F_\nu^+(a_k^+))\delta_{a_k^+}+\mu\llcorner_{]a_k^+,b_k^+[}$ and $\nu_k^+=\nu\llcorner_{]a_k^+,b_k^+[}+(F_\mu^-(b_k^+)-F_\nu^-(b_k^+))\delta_{b_k^+}$. Then, set $\mu^+=\sum_{k\in\mathcal{K}^+}\mu_k^+$ and $\nu^+=\sum_{k\in\mathcal{K}^+}\nu_k^+$.
2. For all $k\in\mathcal{K}^-$, define $\mu_k^-=\mu\llcorner_{]a_k^-,b_k^-[}+\left(F_\nu^-(b_k^-)-F_\mu^-(b_k^-)\right)\delta_{b_k^-}$ and $\nu_k^-=(F_\nu^+(a_k^-)-F_\mu^+(a_k^-))\delta_{a_k^-}+\nu\llcorner_{]a_k^-,b_k^-[}$. Then, set $\mu^-=\sum_{k\in\mathcal{K}^-}\mu_k^-$ and $\nu^-=\sum_{k\in\mathcal{K}^-}\nu_k^-$.
3. Define $\mu_1^==\mu\llcorner_{B^c\cap E^=}$, $\nu_1^==\nu\llcorner_{B^c\cap E^=}$, $\mu_2^==\sum_{x\in B}(\mu(x)-\mu^+(x)-\mu^-(x))\delta_x$, $\nu_2^==\sum_{x\in B}(\nu(x)-\nu^+(x)-\nu^-(x))\delta_x$, $\mu^==\mu_1^=+\mu_2^=$ and $\nu^==\nu_1^=+\nu_2^=$.
4. Define $\mathcal{D}_\mathcal{K}=\{(\mu_k^+,\nu_k^+)\}_{k\in\mathcal{K}^+}\cup\{(\mu_k^-,\nu_k^-)\}_{k\in\mathcal{K}^-}\cup\{(\mu^=,\mu^=)\}$.

Example 2.23. In Example 2.19, we have $\mu_1^+=\mathbb{1}_{[0,1]}\cdot\mathcal{L}^1+\delta_1+\mathbb{1}_{[1,2]}\cdot\mathcal{L}^1$, $\nu_1^+=\mathbb{1}_{[1/2,1]}\cdot\mathcal{L}^1+\delta_1+\mathbb{1}_{[1,2]}\cdot\mathcal{L}^1+\frac{1}{2}\delta_2$, $\mu_1^-=\mathbb{1}_{[2,3]}\cdot\mathcal{L}^1$, $\nu_1^-=\frac{1}{2}\delta_2+\frac{1}{2}\mathbb{1}_{[2,3]}\cdot\mathcal{L}^1$, $\mu_2^-=\mathbb{1}_{[3,4]}\cdot\mathcal{L}^1$, $\nu_2^-=\frac{1}{2}\delta_3+\frac{1}{2}\mathbb{1}_{[3,4]}$, $\mu_1^==\mathbb{1}_{[4,5]}\cdot\mathcal{L}^1+\delta_5+\mathbb{1}_{[5,6]}\cdot\mathcal{L}^1=\nu_1^=$, $\mu_2^==\delta_2+\delta_3+\delta_4=\nu_2^=$, and $\mu^==\delta_2+\delta_3+\delta_4+\mathbb{1}_{[4,5]}\cdot\mathcal{L}^1+\delta_5+\mathbb{1}_{[5,6]}\cdot\mathcal{L}^1=\nu^=$. Figure 4 illustrates how the mass of ν at point 2 is allocated between ν_1^+, $\nu^=$ and ν_1^-.

Remark 2.24.
1. As expected, the measures μ and ν coincide with sum of their respective components. Indeed, for all $x\in B$, $\mu^-(x)+\mu^+(x)+\mu_1^=(x)+\mu_2^=(x)=\mu^+(x)+\mu^-(x)+0+(\mu(x)-\mu^-(x)-\mu^+(x))=\mu(x)$, which implies $(\mu^-+\mu^++\mu_1^=+\mu_2^=)\llcorner_B=\mu\llcorner_B$. Moreover,
$$\begin{aligned}\mu\llcorner_{B^c}&=\mu\llcorner_{B^c\cap E^+}+\mu\llcorner_{B^c\cap E^-}+\mu\llcorner_{B^c\cap E^=}\\&=\mu\llcorner_{E^+}+\mu\llcorner_{E^-}+\mu\llcorner_{B^c\cap E^=}\\&=\mu^+_{\llcorner B^c}+\mu^-_{\llcorner B^c}+\mu_1^=+0=\left(\mu^-+\mu^++\mu_1^=+\mu_2^=\right)\llcorner_{B^c}.\end{aligned}$$
Therefore $\mu=\mu^++\mu^-+\mu_1^=+\mu_2^==\sum_{k\in\mathcal{K}^+}\mu_k^++\sum_{k\in\mathcal{K}^-}\mu_k^-+\mu^=$. Similarly, $\nu=\sum_{k\in\mathcal{K}^+}\nu_k^++\sum_{k\in\mathcal{K}^-}\nu_k^-+\nu^=$.
2. The symmetry between transport from μ to ν and transport from ν to μ still holds. According to Point 2 of Remark 2.20, $\left(]a_k^\mp,b_k^\mp[\right)_{k\in\mathcal{K}^\mp}$ is the family of connected components of $E^\pm(\nu,\mu)$. It follows directly that, for each $k\in\mathcal{K}^\mp$, the pair of marginal components associated to the component $]a_k^\mp,b_k^\mp[$ is $\left(\nu_k^\mp,\mu_k^\mp\right)$.
3. Observe that $\{\mu_k^+\}_{k\in\mathcal{K}^+}\cup\{\nu_k^-\}_{k\in\mathcal{K}^-}$ forms a set of singular measures. However, μ_k^+ and μ_j^- can both assign mass to a_k^+. This happens precisely when $b_j^-=a_k^+$, $F_\mu^+(a_k^+)>F_\nu^+(a_k^+)$ and $F_\nu^-(b_j^-)>F_\mu^-(b_j^-)$. Likewise, (μ_k^+,ν_j^+) and $(\mu_k^+,\mu_2^=)$ may not be pairs of singular measures.
4. The fixed part $\mu^=$ of μ splits into two singular measures $\mu_1^=$ and $\mu_2^=$, each requiring different mathematical treatments. The first part, $\mu_1^=$, represents the mass of μ that lies outside the closure of any positive of negative component of the decomposition: We will later show that this mass is fixed

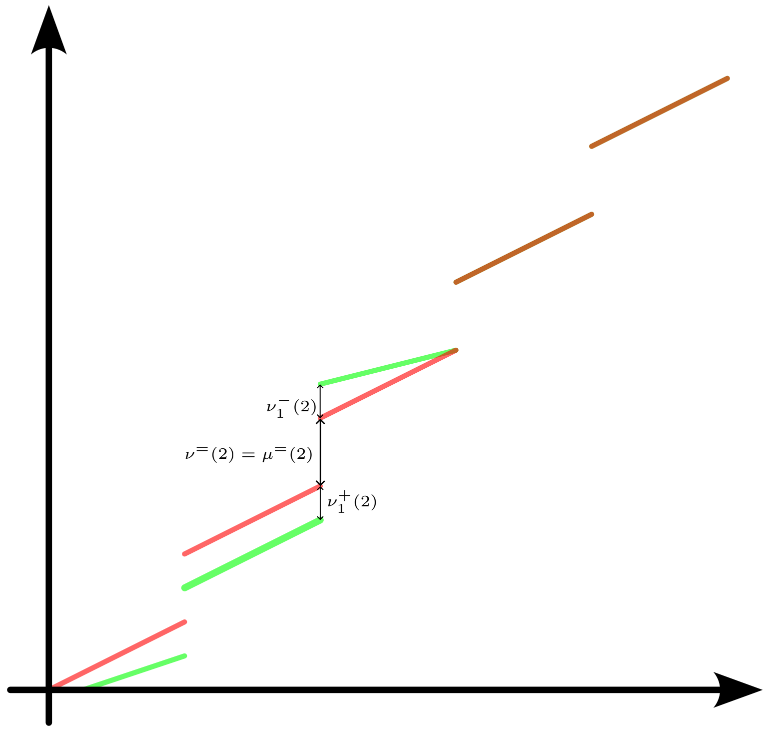

Figure 4: Splitting of point 2 in Example 2.19

by every optimal transport plan. The second part, $\mu_2^=$, is atomic and corresponds to the part mass of μ located at boundary points that is not allocated to any positive or negative component of our decomposition. We explicitly compute $\mu_2^=(x)$ in the next lemma and later prove that $\mu_2^=$ measures the amount of mass at boundary points that is fixed by every optimal transport. In the atomless case, $\mu_2^= = 0$ and $E^=(\mu,\nu) = \{F_\mu^+ = F_\nu^+\}$. Thus $\mu^= = \mu_1^= = \mu\llcorner_{\{F_\mu^+ = F_\nu^+\}} = \mu\llcorner_{E^=}$. In the general case, the equality $\mu^= = \mu\llcorner_{E^=}$ does not hold.

Lemma 2.25. *1. The family*

$$\left(B_l^+ \cap B_r^+, B_l^+ \cap B_r^-, B_l^- \cap B_r^+, B_l^- \cap B_r^-, B_l^+ \setminus B_r, B_l^- \setminus B_r, B_r^+ \setminus B_l, B_r^- \setminus B_l\right)$$

is a family of disjoint set whose union is B.

2. The following inclusions are satisfied:

$$\begin{cases} B_l^+ \subset \{F_\mu^+ \geq F_\nu^+\} \\ B_r^+ \subset \{F_\mu^- \geq F_\nu^-\} \\ B_l^- \subset \{F_\mu^+ \leq F_\nu^+\} \\ B_r^- \subset \{F_\mu^- \leq F_\nu^-\} \\ B_r \setminus B_l \subset \{F_\mu^+ = F_\nu^+\} \\ B_l \setminus B_r \subset \{F_\mu^- = F_\nu^-\} \end{cases} . \tag{6}$$

3. For all $x \in B$,

$$\begin{cases} \mu^+(x) = \mathbb{1}_{B_l^+}(x)(F_\mu^+(x) - F_\nu^+(x)) \\ \nu^+(x) = \mathbb{1}_{B_r^+}(x)(F_\mu^-(x) - F_\nu^-(x)) \\ \mu^-(x) = \mathbb{1}_{B_r^-}(x)(F_\nu^-(x) - F_\mu^-(x)) \\ \nu^-(x) = \mathbb{1}_{B_l^-}(x)(F_\nu^+(x) - F_\mu^+(x)) \end{cases}. \tag{7}$$

4. We have $\mu_2^= = \nu_2^= = \sum_{x \in B} \left[\min\left(F_\mu^+(x), F_\nu^+(x)\right) - \max\left(F_\mu^-(x), F_\nu^-(x)\right)\right] \delta_x$.

Proof. 1. Since (B_l^+, B_l^-, B_l^c) and (B_r^+, B_r^-, B_r^c) are two families of disjoint sets covering $\mathbb{R}$, the family $\mathcal{F} = (B_l^+ \cap B_r^+, B_l^+ \cap B_r^-, B_l^- \cap B_r^+, B_l^- \cap B_r^-, B_l^+ \setminus B_r, B_l^- \setminus B_r, B_r^+ \setminus B_l, B_r^- \setminus B_l, B_l^c \cap B_r^c)$ is a family of disjoint sets that cover $\mathbb{R}$. Since the last element of $\mathcal{F}$ is $B_l^c \cap B_r^c = (B_l \cup B_r)^c = B^c$ and all the other terms are subsets of B,

$$B = B \cap \mathbb{R} = B \cap \biguplus_{C \in \mathcal{F}} C = \biguplus_{C \in \mathcal{F} \setminus \{B^c\}} B \cap C = \biguplus_{C \in \mathcal{F} \setminus \{B^c\}} C,$$

which proves Point 1.

2. For all $k \in \mathcal{K}^+$, $]a_k^+, b_k^+[\subset E^+ \subset \{F_\mu^+ \geq F_\nu^+\}$. Since F_μ^+ and F_ν^+ are right-continuous, we obtain $F_\mu^+(a_k^+) = \lim_{x \to a_k^+; x > a_k^+} F_\mu^+(x) \geq \lim_{x \to a_k^+; x > a_k^+} F_\nu^+(x) = F_\nu^+(a_k^+)$, establishing the first inclusion. The second, third and fourth inclusion are proved similarly. We now turn our attention to the fifth inclusion. Consider $x \in B_r \setminus B_l$. Since $x \in B_r \subset E^= \subset (E^+)^c$, if there exists $\varepsilon > 0$ such that $]x, x+\varepsilon[\subset E^+$, we would have $x \in B_l^+ \subset B_l$. As $x \notin B_l$, for all $\varepsilon > 0$, we get $]x, x+\varepsilon[\cap (E^+)^c$. Thus, there exists a sequence $(x_n)_{n \geq 1} \in \prod_{n \geq 1}]x, x+1/n]$ such that, for all $n \geq 1$, $F_\mu^-(x_n) \leq F_\nu^-(x_n)$ or $F_\mu^+(x_n) \leq F_\nu^+(x_n)$. Letting n go to $+\infty$, we obtain $F_\mu^+(x) \leq F_\nu^+(x)$. Hence $B_r \setminus B_l \subset \{F_\mu^+ \leq F_\nu^+\}$ and the reader may verify that the proof of $B_r \setminus B_l \subset \{F_\mu^+ \geq F_\nu^+\}$ is similar, thus establishing the fifth inclusion. The proof of the sixth inclusion is identical.

3. For all $x \in B$, $\mu^+(x) = \sum_{k \in \mathcal{K}^+} \mu_k^+(x) = \sum_{k \in \mathcal{K}^+} \mathbb{1}_{a_k^+ = x}(F_\mu^+(x) - F_\nu^+(x)) = \mathbb{1}_{B_l^+(x)}(F_\mu^+(x) - F_\nu^+(x))$. The proofs of the three other inequalities is similar.

4. For all $x \in B$, the values of $\mu^+(x)$ and $\mu^-(x)$ can be determined using the partition described in Point 1 and the expressions given in Equation (7). These computations are summarized in the two first columns of Table 1. The computation of $\mu_2^=(x) = \mu(x) - \mu^+(x) - \mu^-(x)$ is then obtained by subtracting the two first column to $\mu(x) = F_\mu^+(x) - F_\mu^-(x)$. The corresponding result is presented in the third column of Table 1. Based on the partition introduced in Point (1), and the expression from Equation (6), the values of $\min(F_\mu^+(x), F_\nu^+(x))$ and $\max(F_\mu^-(x), F_\nu^-(x))$ are computed and shown in Table 2. By subtracting the first column to the second, we recover the third column of Table 1, which proves that $\mu_2^=(x) = \mu(x) - \mu^+(x) - \mu^-(x) = \min(F_\mu^+(x), F_\nu^+(x)) - \max(F_\mu^-(x), F_\nu^-(x))$. The proof of $\nu_2^=(x) = \nu(x) - \nu^+(x) - \nu^-(x) = \min(F_\mu^+(x), F_\nu^+(x)) - \max(F_\mu^-(x), F_\nu^-(x))$ is similar.

□

Definition 2.26 (Components of elements of $\mathfrak{C}(\mu, \nu)$)**.** Consider $\pi \in \mathfrak{C}(\mu, \nu)$.

1. For all $k \in \mathcal{K}^+$, define $A_k^+ = [a_k^+, b_k^+[\times]a_k^+, b_k^+]$ and $\pi_k^+ = \pi_{\llcorner A_k^+}$.

$x \in \cdot$	$\mu^-(x)$	$\mu^+(x)$	$\mu(x) - \mu^-(x) - \mu^+(x)$
$B_l^+ \cap B_r^+$	0	$F_\mu^+(x) - F_\nu^+(x)$	$F_\nu^+(x) - F_\mu^-(x)$
$B_l^+ \cap B_r^-$	$F_\nu^-(x) - F_\mu^-(x)$	$F_\mu^+(x) - F_\nu^+(x)$	$\nu(x)$
$B_l^- \cap B_r^+$	0	0	$\mu(x)$
$B_l^- \cap B_r^-$	$F_\nu^-(x) - F_\mu^-(x)$	0	$F_\mu^+(x) - F_\nu^-(x)$
$B_l^+ \setminus B_r$	0	$F_\mu^+(x) - F_\nu^+(x)$	$F_\nu^+(x) - F_\mu^-(x)$
$B_l^- \setminus B_r$	0	0	$\mu(x)$
$B_r^+ \setminus B_l$	0	0	$\mu(x)$
$B_r^- \setminus B_l$	$F_\nu^-(x) - F_\mu^-(x)$	0	$F_\mu^+(x) - F_\nu^-(x)$

Table 1: Computations of $\mu(x) - \mu^+(x) - \mu^-(x)$

$x \in \cdot$	$\min(F_\mu^+(x), F_\nu^+(x))$	$\max(F_\mu^-(x), F_\nu^-(x))$
$B_l^+ \cap B_r^+$	$F_\nu^+(x)$	$F_\mu^-(x)$
$B_l^+ \cap B_r^-$	$F_\nu^+(x)$	$F_\nu^-(x)$
$B_l^- \cap B_r^+$	$F_\mu^+(x)$	$F_\mu^-(x)$
$B_l^- \cap B_r^-$	$F_\mu^+(x)$	$F_\nu^-(x)$
$B_l^+ \setminus B_r$	$F_\nu^+(x)$	$F_\mu^-(x)$
$B_l^- \setminus B_r$	$F_\mu^+(x)$	$F_\mu^-(x)$
$B_r^+ \setminus B_l$	$F_\mu^+(x)$	$F_\mu^-(x)$
$B_r^- \setminus B_l$	$F_\mu^+(x)$	$F_\nu^-(x)$

Table 2: Computations of $\min(F_\mu^+(x), F_\nu^+(x)) - \max(F_\mu^-(x), F_\nu^-(x))$

2. For all $k \in \mathcal{K}^-$, define $A_k^- =]a_k^-, b_k^-] \times [a_k^-, b_k^-[$ and $\pi_k^- = \pi_{\llcorner A_k^-}$.

3. Define $\pi_1^= = \pi_{\llcorner (B^c \cap E^=) \times \mathbb{R}}$, $\pi_2^= = \pi_{\llcorner \mathrm{D} \cap (B \times \mathbb{R})}$, $A^= = [(B^c \cap E^=) \times \mathbb{R}] \biguplus [\mathrm{D} \cap (B \times \mathbb{R})]$ and $\pi^= = \pi_{\llcorner A^=} = \pi_1^= + \pi_2^=$.

Remark 2.27. 1. Observe that the sets $\{A_k^+\}_{k \in \mathcal{K}^+} \cup \{A_k^-\}_{k \in \mathcal{K}^-} \cup \{A^=\}$ are pairwise disjoint, which implies that the corresponding components $\{\pi_k^+\}_{k \in \mathcal{K}^+} \cup \{\pi_k^-\}_{k \in \mathcal{K}^-} \cup \{\pi^=\}$ forms a family of mutually singular measures.

2. The inclusion or exclusion of boundary points when restricting π is crucial for ensuring $\pi_k^+ \in \mathrm{Marg}(\mu_k^+, \nu_k^+)$: since μ_k^+ is concentrated on $[a_k^+, b_k^+[$ and ν_k^+ on $]a_k^+, b_k^+]$, we must restrict π to $[a_k^+, b_k^+[\times]a_k^+, b_k^+]$. For instance, if $\mu = 2\delta_1 + \mathbb{1}_{[1,2]} \cdot \mathcal{L}^1 + \delta_2$ and $\nu = \delta_1 + \mathbb{1}_{[1,2]} \cdot \mathcal{L}^1 + 2\delta_2$, μ admits two components $\mu_1^+ = \delta_1 + \mathbb{1}_{[1,2]} \cdot \mathcal{L}^1$, $\mu^= = \delta_1 + \delta_2$, while ν admits two components $\nu_1^+ = \mathbb{1}_{[1,2]} \cdot \mathcal{L}^1 + \delta_2$ and $\nu^= = \delta_1 + \delta_2$. One may verify that, among the 16 possible combinations of boundary inclusion, only the chosen convention yields a restriction of optimal transport plans that lies in $\mathrm{Marg}(\mu_k^+, \nu_k^+)$.

3. Symmetry is also preserved at the level of cyclically monotone transport plans. Namely, by symmetry of $i : (x, y) \mapsto (y, x)$, the map $i_\# : \pi \in \mathcal{P}(\mathbb{R}^2) \mapsto i_\# \pi \in \mathcal{P}(\mathbb{R}^2)$ is a bijection from $\mathfrak{C}(\mu, \nu)$ to $\mathfrak{C}(\nu, \mu)$. Moreover, there is a correspondence between the components of a plan $\pi \in \mathfrak{C}(\mu, \nu)$ and the

components of $i_{\#}\pi \in \mathfrak{C}(\nu,\mu)$. More precisely, for all $k \in \mathcal{K}^+$, it follows from Point 2 of Remark 2.20, that $]a_k^+, b_k^+[$ is a component for $E^+(\mu,\nu)$ and a component for $E^-(\nu,\mu)$. Let $\varphi_k : \pi \in \mathfrak{C}(\mu,\nu) \mapsto \pi\llcorner_{]a_k^+,b_k^+[\times]a_k^+,b_k^+]}$ and $\phi_k : \pi \in \mathfrak{C}(\nu,\mu) \mapsto \pi\llcorner_{]a_k^+,b_k^+]\times[a_k^+,b_k^+[}$ denote the functions mapping every element of $\mathfrak{C}(\mu,\nu)$ and $\mathfrak{C}(\nu,\mu)$, respectively, to its component associated with $]a_k^+, b_k^+[$. Then, it is straightforward that $\phi_k = i_{\#} \circ \varphi_k \circ i_{\#}$. The same results holds for $k \in \mathcal{K}^-$. For the two equal components of the decomposition of (μ,ν) and (ν,μ), one can go from one to another by applying $i_{\#}$. This is a consequence of the equalities $E^=(\mu,\nu) = E^=(\nu,\mu)$ and $B(\mu,\nu) = B(\nu,\mu)$ established in Point 2 of Remark 2.20.

The following lemma quantifies how mass splits at border points, specifying the amount of mass that moves strictly to the left or strictly to the right. To quantify mass moving to the right, we use the disjoint covering $B = B_l^- \uplus B_l^+ \uplus (B_r \setminus B_l)$; for mass moving on the left we base our computations on the disjoint covering $B = B_r^- \uplus B_r^+ \uplus (B_l \setminus B_r)$.

Lemma 2.28. *Consider $(\mu,\nu) \in \mathcal{M}_+^2$, $x \in \mathbb{R}$ and $\pi \in \mathfrak{C}(\mu,\nu)$.*

1. *If $x \in B_l^- \cup (B_r \setminus B_l)$, then $\pi(\{x\}\times]x,+\infty[) = 0$.*
2. *If $x \in B_l^+$, $\pi(\{x\}\times]x,+\infty[) = F_\mu^+(x) - F_\nu^+(x)$.*
3. *If $x \in B_r^+ \cup (B_l \setminus B_r)$, then $\pi(\{x\}\times]-\infty,x[) = 0$.*
4. *If $x \in B_r^-$, $\pi(\{x\}\times]-\infty,x[) = F_\nu^-(x) - F_\mu^-(x)$.*

Proof. The proofs of Points 3 and 4 are analogous to those of Points 1 and 2, and are therefore omitted.

1. Equation (6), implies that $B_l^- \cup (B_r \setminus B_l) \subset \{F_\mu^+ \le F_\nu^+\}$. By Point 3 Proposition 2.15, it follows that $\pi(]-\infty,x]\times]x,+\infty[) = 0$. In particular, $\pi(\{x\}\times]x,+\infty[) = 0$.
2. Since $B_l^+ \subset E^=$, Lemma 2.21 implies $B_l^+ \subset \mathcal{B}$, so that $\pi(]-\infty,x[\times]x,+\infty[) = 0$. Hence,

$$\begin{aligned} F_\mu^+(x) &= \pi(]-\infty,x]\times\mathbb{R}) \\ &= \pi(]-\infty,x]^2) + \pi(]-\infty,x]\times]x,+\infty[) \\ &= \pi(]-\infty,x]^2) + \pi(]-\infty,x[\times]x+\infty[) + \pi(\{x\}\times]x,+\infty[) \\ &= \pi(]-\infty,x]^2) + \pi(\{x\}\times]x,+\infty[). \end{aligned}$$

   Moreover, from Equation (6), we have $B_l^+ \subset \{F_\mu^+ \ge F_\nu^+\}$. By Point 1 of Proposition 2.15, it follows that $\pi(]x,+\infty[\times]-\infty,x]) = 0$. Hence,

$$\begin{aligned} F_\nu^+(x) &= \pi(\mathbb{R}\times]-\infty,x]) \\ &= \pi(]-\infty,x]^2) + \pi(]x,+\infty[\times]-\infty,x]) \\ &= \pi(]-\infty,x]^2). \end{aligned}$$

   Therefore, $F_\mu^+(x) - F_\nu^+(x) = \pi(\{x\}\times]x,+\infty[)$. ∎

We now establish that the marginals of the components of a cyclically monotone transport are the components of our marginals. We then prove that a cyclically monotone transport plan equals the sum of its components.

Proposition 2.29. *Consider $(\mu, \nu) \in \mathcal{M}_+^2$ and $\pi \in \mathfrak{C}(\mu, \nu)$.*

1. *For all $k \in \mathcal{K}^+$, $\pi_k^+ \in \mathrm{Marg}(\mu_k^+, \nu_k^+)$.*
2. *For all $k \in \mathcal{K}^-$, $\pi_k^- \in \mathrm{Marg}(\mu_k^-, \nu_k^-)$.*
3. *Equalities $\pi_1^= = (\mathrm{id}, \mathrm{id})_\# \mu_1^= = (\mathrm{id}, \mathrm{id})_\# \nu_1^=$ are satisfied.*
4. *Equalities $\pi_2^= = (\mathrm{id}, \mathrm{id})_\# \mu_2^= = (\mathrm{id}, \mathrm{id})_\# \nu_2^=$ are satisfied.*
5. *Equality $\pi = \sum_{k \in \mathcal{K}^+} \pi_k^+ + \sum_{k \in \mathcal{K}^-} \pi_k^- + \pi^=$ is satisfied.*

Proof.

1. For all $t \in \mathbb{R}$,

$$\mu_k^+(]-\infty, t]) = \begin{cases} 0 & \text{if } t < a_k^+ \\ F_\mu^+(t) - F_\nu^+(a_k^+) & \text{if } t \in [a_k^+, b_k^+[\\ F_\mu^-(b_k^+) - F_\nu^+(a_k^+) & \text{if } t \geq b_k^+ \end{cases}$$

and

$$\pi_k^+(]-\infty, t] \times \mathbb{R}) = \begin{cases} 0 & \text{if } t < a_k^+ \\ \pi([a_k^+, t] \times]a_k^+, b_k^+]) & \text{if } t \in [a_k^+, b_k^+[\\ \pi([a_k^+, b_k^+[\times]a_k^+, b_k^+]) & \text{if } t \geq b_k^+ \end{cases}.$$

Since

$$\begin{cases} \lim_{t \to b_k^+, t < b_k^+} \pi([a_k^+, t] \times]a_k^+, b_k^+]) = \pi([a_k^+, b_k^+[\times]a_k^+, b_k^+]) \\ \lim_{t \to b_k^+, t < b_k^+} F_\mu^+(t) - F_\nu^+(a_k^+) = F_\mu^-(b_k^+) - F_\nu^+(a_k^+) \end{cases},$$

if, for all $t \in [a_k^+, b_k^+[$, $\pi([a_k^+, t] \times]a_k^+, b_k^+]) = F_\mu^+(t) - F_\nu^+(a_k^+)$ holds, then the first marginal of π_k^+ is μ_k^+. For every $t \in [a_k^+, b_k^+[$,

$$\begin{aligned} \pi([a_k^+, t] \times]a_k^+, b_k^+]) &= \pi([a_k^+, t] \times]a_k^+, +\infty[) \\ &= \mu([a_k^+, t]) - \pi([a_k^+, t] \times]-\infty, a_k^+]) \\ &= F_\mu^+(t) - F_\mu^-(a_k^+) - \pi(\{a_k^+\} \times]-\infty, a_k^+]) \\ &= F_\mu^+(t) - F_\mu^-(a_k^+) - (\mu(a_k^+) - \pi(\{a_k^+\} \times]a_k^+, +\infty[)) \\ &= F_\mu^+(t) - F_\nu^+(a_k^+), \end{aligned}$$

where the first equality comes from $b_k^+ \in E^= \subset \mathcal{B}$, the third equality comes from Point 1 of Proposition 2.15 (with $x = a_k^+$) and Equation (6), while the last equality comes from Point 2 of Lemma 2.28. Therefore, $p_{1\#} \pi_k^+ = \mu_k^+$, and a similar argument shows that $p_{2\#} \pi_k^+ = \nu_k^+$.

2. Same proof as for the previous point.

3. We first show the inclusion $\mathrm{spt}(\pi)\cap(E^=\times\mathbb{R})\cap(B^c\times\mathbb{R})\subset \mathrm{D}$. Assume for contradiction that there exists $(x,y)\in\mathrm{spt}(\pi)\cap(E^=\times\mathbb{R})\cap(B^c\times\mathbb{R})$ such that $x\neq y$. Since $x\in B^c$, $]\min(x,y),\max(x,y)[\subset E^-$ or $]\min(x,y),\max(x,y)[\subset E^+$ is prohibited, which implies there exists $z\in]\min(x,y),\max(x,y)[\cap E^=$. If $x<y$, $z\in E^=\subset\mathcal{B}$ implies $\pi(]-\infty,z[\times]z,+\infty[)=0$, and we have $(x,y)\in]-\infty,z[\times]z,+\infty[$. If $y<x$, $z\in E^=\subset\mathcal{B}$ implies $\pi(]z,+\infty[\times]-\infty,z[)=0$, and we have $(x,y)\in]z,+\infty[\times]-\infty,z[$. In both cases, this contradicts the fact that $(x,y)\in\mathrm{spt}(\pi)$, thereby proving the inclusion. By construction, the first marginal of $\pi_1^=$ is $\mu_1^=$ and $0\leq\pi_1^=(\mathrm{D}^c)=\pi([(B^c\cap E^=)\times\mathbb{R}]\cap\mathrm{D}^c)\leq\pi(\mathrm{D}\cap\mathrm{D}^c)=0$, which shows the first identity. Observe that $\pi_1^==\pi_{(E^=\cap B^c)\times(E^=\cap B^c)}$. The proof of $\pi\llcorner_{\mathbb{R}\times(E^=\cap B^c)}=(\mathrm{id},\mathrm{id})_\#\nu_1^=$ is similar. Therefore, $(\mathrm{id},\mathrm{id})_\#\mu_1^==\pi\llcorner_{(E^=\cap B^c)\times\mathbb{R}}=\pi\llcorner_{(E^=\cap B^c)^2}=\pi\llcorner_{\mathbb{R}\times(E^=\cap B^c)}=(\mathrm{id},\mathrm{id})_\#\nu_1^=$.

4. By definition, $\pi_2^==\sum_{x\in B}\pi(x,x)\delta_{(x,x)}$, so we have to show that, for all $x\in B$, $\pi(x,x)=\mu_2^=(x)=\mu(x)-\mu^-(x)-\mu^+(x)$. We compute using the partition of Point 1 of Lemma 2.25, together with Lemma 2.28. More precisely, the two first point of Lemma 2.28 give the value of first column of Table 3, while the two other points give the value of the second column of Table 3. The final column follows from the identity $\pi(x,x)=\mu(x)-\pi(\{x\}\times]x,+\infty[)-\pi(\{x\}\times]-\infty,x[)$. Since the last columns of Table 1 and Table 3 are the same, this finishes the proof.

5. Define $\tilde\pi=\sum_{k\in\mathcal{K}^+}\pi_k^++\sum_{k\in\mathcal{K}^-}\pi_k^-+\pi^=$. Since $\{A_k^+\ ;\ k\in\mathcal{K}^+\}\cup\{A_k^-\ ;\ k\in\mathcal{K}^-\}\cup\{A^=\}$ is a class of disjoint sets, we have $\tilde\pi=\pi\llcorner_{(\cup_{k\in\mathcal{K}^+}A_k^+)\cup(\cup_{k\in\mathcal{K}^-}A_k^-)\cup A^=}\leq\pi$. Furthermore, $\tilde\pi(\mathbb{R}^2)=\sum_{k\in\mathcal{K}^+}\mu_k^+(\mathbb{R})+\sum_{k\in\mathcal{K}^-}\mu_k^-(\mathbb{R})+\mu^=(\mathbb{R})=\mu(\mathbb{R})=\pi(\mathbb{R}^2)$. Thus, $\pi=\tilde\pi$. □

$x\in\cdot$	$\pi(\{x\}\times]x,+\infty[)$	$\pi(\{x\}\times]-\infty,x[)$	$\pi(\{x\}\times\{x\})$
$B_l^+\cap B_r^+$	$F_\mu^+(x)-F_\nu^+(x)$	0	$F_\nu^+(x)-F_\mu^-(x)$
$B_l^+\cap B_r^-$	$F_\mu^+(x)-F_\nu^+(x)$	$F_\nu^-(x)-F_\mu^-(x)$	$\nu(x)$
$B_l^-\cap B_r^+$	0	0	$\mu(x)$
$B_l^-\cap B_r^-$	0	$F_\nu^-(x)-F_\mu^-(x)$	$F_\mu^+(x)-F_\nu^-(x)$
$B_l^+\setminus B_r$	$F_\mu^+(x)-F_\nu^+(x)$	0	$F_\nu^+(x)-F_\mu^-(x)$
$B_l^-\setminus B_r$	0	0	$\mu(x)$
$B_r^+\setminus B_l$	0	0	$\mu(x)$
$B_r^-\setminus B_l$	0	$F_\nu^-(x)-F_\mu^-(x)$	$F_\mu^+(x)-F_\nu^-(x)$

Table 3: Values of $\pi(x,x)$

Definition 2.30. We denote by $\leq_{\mathbb{R}^2}$ the usual partial order on $\mathbb{R}^2$: for all $(x_1,y_1),(x_2,y_2)\in\mathbb{R}^2$, $(x_1,y_1)\leq_{\mathbb{R}^2}(x_2,y_2)$ if $x_1\leq x_2$ and $y_1\leq y_2$. This order naturally extends to subsets of $\mathbb{R}^2$ as follows: for all subsets A,B of $\mathbb{R}^2$, we write $A\leq_{\mathbb{R}^2}B$ if for all $(a,b)\in A\times B$, $a\leq_{\mathbb{R}^2}b$. Let $\mathfrak{OR}$ denote the set formed by the classes $\mathcal{C}$ of subsets of $\mathbb{R}^2$ such that,

$$\forall A,B\in\mathcal{C}: A\neq B\implies A\leq_{\mathbb{R}^2}B \text{ or } B\leq_{\mathbb{R}^2}A.$$

The reader may think of $\mathfrak{OR}$ as the set ordered classes of subset of $\mathbb{R}^2$. However, we do not require that classes $\mathcal{C}\in\mathfrak{OR}$ satisfy $A\leq_{\mathbb{R}^2}A$ for every $A\in\mathcal{C}$.

Lemma 2.31. *1. Let A and B denote two cyclically monotone subsets of $\mathbb{R}^2$. If $A\leq_{\mathbb{R}^2}B$, then $A\cup B$ is a cyclically monotone set.*

2. *If* Γ *is a cyclically monotone set, then* $\mathrm{D}\cup\Gamma$ *is also a cyclically monotone set*

Proof. 1. Consider $n\geq 1$ and $((x_k,y_k))_{k\in[\![1,n]\!]}\in(A\cup B)^n$. Define $I_A=\{i\in[\![1,n]\!]\ ;\ (x_i,y_i)\in A\}$ and $I_B=\{i\in[\![1,n]\!]\ ;\ (x_i,y_i)\in B\setminus A\}$. Observe that $I_A\uplus I_B=[\![1,n]\!]$ and define $k=\#I_A$. Let $(x_{(1)},\dots,x_{(k)})$, $(x_{(k+1)},\dots,x_{(n)})$, $(y_{(1)},\dots,y_{(k)})$, and $(y_{(k+1)},\dots,y_{(n)})$ denote the families $(x_i)_{i\in I_A}$, $(x_i)_{i\in I_B}$, $(y_i)_{i\in I_A}$, and $(y_i)_{i\in I_B}$, respectively, sorted in non-decreasing order. Since $A\leq_{\mathbb{R}^2}B$, there exists $x\in\mathbb{R}$ such that $A\subset]-\infty,x]^2$ and $B\subset[x,+\infty[^2$. Thus $x_{(1)}\leq\dots\leq x_{(k)}\leq x\leq x_{(k+1)}\leq\dots x_{(n)}$ and $y_{(1)}\leq\dots\leq y_{(k)}\leq x\leq y_{(k+1)}\leq\dots y_{(n)}$. Hence, $(x_{(1)},\dots,x_{(n)})$ and $(y_{(1)},\dots,y_{(n)})$ correspond to the sequences $(x_1,\dots,x_n)$ and $(y_1,\dots,y_n)$, respectively, sorted in non-decreasing order. Since the monotone transport plan $\sum_{i=1}^n\delta_{(x_{(i)},y_{(i)})}$ from $\sum_{i=1}^n\delta_{x_i}$ to $\sum_{i=1}^n\delta_{y_i}$ is optimal, we have $\sum_{i=1}^n|y_{(i)}-x_{(i)}|\leq\sum_{i=1}^n|y_{i+1}-x_i|$. Moreover, as $((x_i,y_i))_{i\in I_A}\in A^k$, $((x_i,y_i))_{i\in I_B}\in B^{n-k}$ and (A,B) is a pair of cyclically monotone sets, we deduce that:

$$\sum_{i=1}^n|y_i-x_i|=\sum_{i\in I_A}|y_i-x_i|+\sum_{i\in I_B}|y_i-x_i|\leq\sum_{i=1}^k|y_{(i)}-x_{(i)}|+\sum_{i=k+1}^n|y_{(i)}-x_{(i)}|.$$

Thus,

$$\sum_{i=1}^n|y_i-x_i|\leq\sum_{i=1}^n|y_{(i)}-x_{(i)}|\leq\sum_{i=1}^n|y_{i+1}-x_i|.$$

Therefore, $A\cup B$ is cyclically monotone.

2. Consider $((x_i,y_i))_{i\in[\![1,n]\!]}\in(\mathrm{D}\cup\Gamma)^n$ and define $I=\{i\in[\![1,n]\!]\ ;\ (x_i,y_i)\in\Gamma\}$. For all $i\in I$, set $\sigma(i)=\inf\{j>i\ ;\ (x_i,y_i)\in\Gamma\}$.[8] For every $i\in I$ and $j\in[\![i+1,\sigma(i)-1]\!]$, (x_j,y_j) belongs to D. As Γ is cyclically monotone, it follows that:

$$\sum_{i=1}^n|y_i-x_i|=\sum_{i\in I}|y_i-x_i|\leq\sum_{i\in I}|y_{\sigma(i)}-x_i|=\sum_{i\in I}\left|\sum_{j=i}^{\sigma(i)-1}y_{j+1}-x_j\right|\leq\sum_{i\in I}\sum_{j=i}^{\sigma(i)-1}|y_{j+1}-x_j|=\sum_{i=1}^n|y_{i+1}-x_i|.$$

Thus, $\Gamma\cup\mathrm{D}$ is cyclically monotone. □

Remark 2.32. Let $(C_i)_{i\in\mathcal{I}}$ be a family of ordered cyclically monotone set and define $\Gamma=\mathrm{D}\cup\left(\bigcup_{i\in\mathcal{I}}C_i\right)$. Then Γ is a cyclically monotone set. Indeed, by induction on Point 1 of Lemma 2.31, for any finite subset J of I, $\cup_{j\in J}C_j$ is cyclically monotone. As every finite sequence of element of $\cup_{i\in\mathcal{I}}C_i$ belongs to a such set, $\cup_{i\in\mathcal{I}}C_i$ is cyclically monotone. According to Point 2 of Lemma 2.31, Γ is cyclically monotone.

To state the properties of our marginal decomposition, we introduce a reinforced version of the stochastic order, due to Kellerer [15, Definition 1.17].

Definition 2.33 (Large[9] reinforced stochastic order of Kellerer)**.** Consider $(\gamma_1,\gamma_2)\in\mathcal{M}_+(\mathbb{R})^2$ and define

$$\begin{cases}T_+(\gamma_1,\gamma_2)=\left\{t\in\mathbb{R}\ ;\ F^+_{\gamma_1}(t)>0\text{ and }F^+_{\gamma_2}(t)<\gamma_2(\mathbb{R})\right\}\\ T_-(\gamma_1,\gamma_2)=\left\{t\in\mathbb{R}\ ;\ F^-_{\gamma_1}(t)>0\text{ and }F^-_{\gamma_2}(t)<\gamma_2(\mathbb{R})\right\}\end{cases}.$$

[8] Where we use the convention $(x_{n+j},y_{n+j})=(x_j,y_j)$.

[9] We added *large* to make a distinction between this order and a similar order — related to G instead of F — that will be introduced later in the article (Definition 3.18). For the sake of fluidity, we omit the word large in the following.

We say that γ_1 is smaller than γ_2 in the (large) reinforced stochastic order if $\gamma_1 \leq_{\text{st}} \gamma_2$, $T_+(\gamma_1,\gamma_2) \subset \{F^+_{\gamma_1} > F^+_{\gamma_2}\}$, and $T_-(\gamma_1,\gamma_2) \subset \{F^-_{\gamma_1} > F^-_{\gamma_2}\}$. In this case, we write $\gamma_1 \leq_{\text{F}} \gamma_2$.[10]

Remark 2.34. 1. Intuitively, if $\gamma_1 \leq_{\text{st}} \gamma_2$, then $\gamma_1 \leq_{\text{F}} \gamma_2$ means that the relations $F^+_{\gamma_1} \geq F^+_{\gamma_2}$ and $F^-_{\gamma_1} \geq F^-_{\gamma_2}$ given by the stochastic order are strict "wherever possible". Specifically, since $T_+(\gamma_1,\gamma_2)^c \subset \{F^+_{\gamma_1} = F^+_{\gamma_2} = 0\} \cup \{F^+_{\gamma_1} = F^+_{\gamma_2} = \gamma_1(\mathbb{R})\}$ and $T_-(\gamma_1,\gamma_2)^c \subset \{F^-_{\gamma_1} = F^-_{\gamma_2} = 0\} \cup \{F^-_{\gamma_1} = F^-_{\gamma_2} = \gamma_2(\mathbb{R})\}$, it follows that $\{F^+_{\gamma_1} > F^+_{\gamma_2}\} \subset T_+(\gamma_1,\gamma_2)$ and $\{F^-_{\gamma_1} > F^-_{\gamma_2}\} \subset T_-(\gamma_1,\gamma_2))$. Thus, the relation $\gamma_1 \leq_{\text{F}} \gamma_2$ precisely means that the inclusions $\{F^+_{\gamma_1} > F^+_{\gamma_2}\} \subset T_+(\gamma_1,\gamma_2)$ and $\{F^-_{\gamma_1} > F^-_{\gamma_2}\} \subset T_-(\gamma_1,\gamma_2)$ are actually equalities.

2. Observe that, for all $\gamma \in \mathcal{M}^+(\mathbb{R})$, $s_\gamma = \inf(\{F^-_\gamma > 0\})$ and $S_\gamma = \sup(\{F^+_\gamma < \gamma(\mathbb{R})\})$.

3. To state this point, recall the notation $\text{Atom}(\gamma)$ (see Sub-subsection 1.4, Point 3). By the previous point of the Remark, if $\gamma_1 \leq_{\text{st}} \gamma_2$,

$$T_+(\gamma_1,\gamma_2) = \begin{cases} [s_{\gamma_1}, S_{\gamma_2}[\text{ if } s_{\gamma_1} \in \text{Atom}(\gamma_1) \\ ]s_{\gamma_1}, S_{\gamma_2}[\text{ otherwise} \end{cases} \quad \text{and} \quad T_-(\gamma_1,\gamma_2) = \begin{cases}]s_{\gamma_1}, S_{\gamma_2}] \text{ if } S_{\gamma_2} \in \text{Atom}(\gamma_2) \\ ]s_{\gamma_1}, S_{\gamma_2}[\text{ otherwise} \end{cases} \tag{8}$$

The reader may verify that, if $F^+_{\gamma_1} \geq F^+_{\gamma_2}$, $]s_{\gamma_1}, S_{\gamma_2}[\subset E^+(\gamma_1,\gamma_2)$, $s_{\gamma_1} \notin \text{Atom}(\gamma_2)$, and $S_{\gamma_2} \notin \text{Atom}(\gamma_1)$, then $\gamma_1 \leq_{\text{F}} \gamma_2$.

We now examine the properties of the marginal decomposition introduced in Definition 2.22. Specifically, we establish that each pair is ordered in the reinforced stochastic order of Kellerer and that the fixed parts coincide.

Proposition 2.35. *Consider $(\mu,\nu) \in \mathcal{M}^2_+$.*

1. *The equalities $\mu = \sum_{k\in\mathcal{K}^+} \mu^+_k + \sum_{k\in\mathcal{K}^-} \mu^-_k + \mu^=$ and $\nu = \sum_{k\in\mathcal{K}^+} \nu^+_k + \sum_{k\in\mathcal{K}^-} \nu^-_k + \nu^=$ hold.*
2. *For all $k \in \mathcal{K}^+$ (resp. $\mathcal{K}^-$), $\mu^+_k \leq_{\text{F}} \nu^+_k$ (resp. $\nu^-_k \leq_{\text{F}} \mu^-_k$).*
3. *We have $\mu^=_1 = \nu^=_1$, $\mu^=_2 = \nu^=_2$, and $\mu^= = \nu^=$.*

Proof. 1. This result was established in Point 1 of Remark 2.24.

2. For all $k \in \mathcal{K}^+$ and $t \in \mathbb{R}$,

$$\begin{cases} F^+_{\mu^+_k}(t) = \mathbb{1}_{a^+_k \leq t < b^+_k}(F^+_\mu(t) - F^+_\nu(a^+_k)) + \mathbb{1}_{t \geq b^+_k}(F^-_\mu(b^+_k) - F^+_\nu(a^+_k)) \\ F^+_{\nu^+_k}(t) = \mathbb{1}_{a^+_k < t < b^+_k}(F^+_\nu(t) - F^+_\nu(a^+_k)) + \mathbb{1}_{t \geq b^+_k}((F^-_\mu(b^+_k) - F^+_\nu(a^+_k)) \end{cases} \tag{9}$$

and

$$\begin{cases} F^-_{\mu^+_k}(t) = \mathbb{1}_{a^+_k < t \leq b^+_k}(F^-_\mu(t) - F^+_\nu(a^+_k)) + \mathbb{1}_{t > b^+_k}(F^-_\mu(b^+_k) - F^+_\nu(a^+_k)) \\ F^-_{\nu^+_k}(t) = \mathbb{1}_{a^+_k < t \leq b^+_k}(F^-_\nu(t) - F^+_\nu(a^+_k)) + \mathbb{1}_{t > b^+_k}(F^-_\mu(b^+_k) - F^+_\nu(a^+_k)) \end{cases}, \tag{10}$$

[10] This notation, motivated by the associated Strassen-type theorem, is adopted from Kellerer [15].

which implies

$$\begin{cases} F^+_{\mu_k^+}(t) - F^+_{\nu_k^+}(t) = \mathbb{1}_{t=a_k^+}(F^+_\mu(a_k^+) - F^+_\nu(a_k^+)) + \mathbb{1}_{a_k^+<t<b_k^+}(F^+_\mu(t) - F^+_\nu(t)) \\ F^-_{\mu_k^+}(t) - F^-_{\nu_k^+}(t) = \mathbb{1}_{a_k^+<t\leq b_k^+}(F^-_\mu(t) - F^-_\nu(t)) \end{cases}. \tag{11}$$

We have $]a_k^+, b_k^+[\subset E^+$ and Equation (6) implies $a_k^+ \in \{F^+_\mu \geq F^+_\nu\}$. From Equation (11), it follows that $F^+_{\mu_k^+} \geq F^+_{\nu_k^+}$. Morever, for all $t \in]a_k^+, b_k^+[$, $F^-_\mu(t) > F^-_\nu(t) \geq F^+_\nu(a_k^+)$ and $F^+_\nu(t) < F^+_\mu(t) \leq F^-_\mu(b_k^+)$. By Equation (10) and (9), we respectively obtain that $\{F^-_{\mu_k^+} > 0\} =]a_k^+, +\infty[$ and $\{F^+_{\nu_k^+} < \nu_k^+(\mathbb{R})\}$. By Point 2 of Remark 2.34, we get $s_{\mu_k^+} = a_k^+$ and $S_{\nu_k^+} = b_k^+$ and Equation (11) implies $]s_{\mu_k^+}, S_{\nu_k^+}[\subset E^+(\mu_k^+, \nu_k^+)$. Since $\mu_k^+(S_{\nu_k^+}) = \nu_k^+(s_{\mu_k^+}) = 0$, by Point 3 of Remark 2.34, we get $\mu_k^+ \leq_{\mathrm{F}} \nu_k^+$. The proof of the inequality $\nu_k^- \leq_{\mathrm{F}} \mu_k^-$ is similar.

3. The Points 3 and 4 of Proposition 2.29 establish that $\mu_1^= = \nu_1^=$ and $\mu_2^= = \nu_2^=$. This shows that $\mu^= = \mu_1^= + \mu_2^= = \nu_1^= + \nu_2^= = \nu^=$.

□

We now state and prove the decomposition result for cyclically monotone transport plans associated with the marginal components introduced in Definition 2.22.

Theorem 2.36 (Decomposition of $\mathfrak{C}(\mu,\nu)$)**.** *Consider $(\mu,\nu) \in \mathcal{M}_+^2$. We use the notation for marginal components introduced in Definition 2.22 and the notation $((A_k^+)_{k\in\mathcal{K}^+}, (A_k^-)_{k\in\mathcal{K}^-}, A^=)$ introduced in Definition 2.26. Define also $\mathfrak{P} = \prod_{k\in\mathcal{K}^+}\mathfrak{C}(\mu_k^+,\nu_k^+) \times \prod_{k\in\mathcal{K}^-}\mathfrak{C}(\mu_k^-,\nu_k^-) \times \mathfrak{C}(\mu^=,\nu^=)$.*

1. *For all $k \in \mathcal{K}^+$ (resp. $\mathcal{K}^-$), $\mu_k^+ \leq_{\mathrm{F}} \nu_k^+$ (resp. $\nu_k^- \leq_{\mathrm{F}} \mu_k^-$). Moreover, $\mu^= = \nu^=$.*
2. *The map $\varphi : \mathfrak{P} \to \mathfrak{C}(\mu,\nu)$ defined by*

$$\varphi((\pi_k^+)_{k\in\mathcal{K}^+}, (\pi_k^-)_{k\in\mathcal{K}^-}, \pi^=) = \sum_{k\in\mathcal{K}^+}\pi_k^+ + \sum_{k\in\mathcal{K}^-}\pi_k^- + \pi^=$$

is a bijection. Furthermore, its inverse $\varphi^{-1} : \mathfrak{C}(\mu,\nu) \to \mathfrak{P}$ is given by

$$\varphi^{-1}(\pi) = \left(\left(\pi_{\llcorner A_k^+}\right)_{k\in\mathcal{K}^+}, \left(\pi_{\llcorner A_k^-}\right)_{k\in\mathcal{K}^-}, \pi_{\llcorner A^=}\right),. \tag{12}$$

Proof. The first point follows from Proposition 2.35. For the second point, we first establish that φ is indeed valued in $\mathfrak{C}(\mu,\nu)$. Consider a family $\mathcal{F} := ((\pi_k^+)_{k\in\mathcal{K}^+}, (\pi_k^-)_{k\in\mathcal{K}^-}, \pi^=) \in \mathfrak{P}$. By Point 1 of Proposition 2.35, $\pi^* := \sum_{k\in\mathcal{K}^+}\pi_k^+ + \sum_{k\in\mathcal{K}^-}\pi_k^- + \pi^=$ is a transport plan from μ to ν. By Proposition 2.11, π is concentrated on $\Gamma^* = \cup_{k\in\mathcal{K}^+}(\mathrm{F}\cap A_k^+) \cup \cup_{k\in\mathcal{K}^-}(\tilde{F}\cap A_k^-) \cup \mathrm{D}$. As F and $\tilde{\mathrm{F}}$ are cyclically monotone sets, by Remark 2.32, Γ^* is a cyclically monotone set. Thus, $\pi^* \in \mathfrak{C}(\mu,\nu)$ and φ is indeed valued in $\mathfrak{C}(\mu,\nu)$. We now establish that φ is injective. Consider $\pi \in \mathfrak{C}(\mu,\nu)$ and $\mathcal{F} := ((\pi_k^+)_{k\in\mathcal{K}^+}, (\pi_k^-)_{k\in\mathcal{K}^-}, \pi^=) \in \mathfrak{P}$ such that $\varphi(\mathcal{F}) = \pi$. Since $\pi^= \in \mathfrak{C}(\mu^=,\nu^=) = \{(\mathrm{id},\mathrm{id})_\#\mu^=\}$, $\pi^= = (\mathrm{id},\mathrm{id})_\#\mu^=$. In particular $\pi^=$ is concentrated on $A^=$. For all $k \in \mathcal{K}^+$(resp. $k \in \mathcal{K}^-$), π_k^+ (resp. π_k^-) is concentrated on A_k^+ (resp. A_k^-). As $\{A_k^+ ; k \in \mathcal{K}^+\} \cup \{A_k^- ; k \in \mathcal{K}^-\} \cup \{A^=\}$ is a class of disjoints set, for all $k \in \mathcal{K}^+$,

$$\pi_{\llcorner A_k^+} = \left(\sum_{k\in\mathcal{K}^+}\pi_j^+ + \sum_{j\in\mathcal{K}^-}\pi_j^- + \pi^=\right)_{\llcorner A_k^+} = \pi_k^+.$$

Similarly, for all $k \in \mathcal{K}^-$, $\pi_k^- = \pi_{\llcorner A_k^-}$, and $\pi^= = \pi_{A^=}$. Therefore, $\mathcal{F}$ is equal to the second term of Equation (12) and φ is injective. We establish now that φ is surjective. Consider $\pi \in \mathfrak{C}(\mu,\nu)$ and let $\mathcal{F}$ denote the right term of Equation (12). From Points (1–4) of Proposition 2.29 and Remark 2.7, if follows that $\mathcal{F} \in \mathfrak{P}$. By Point 5 of Proposition 2.29, it follows that $\varphi(\mathcal{F}) = \pi$. Therefore, φ is surjective and Formula (12) is satisfied. □

Using Notation 1.1, Point 2 gives the direct sum decomposition

$$\begin{aligned}\mathfrak{C}(\mu,\nu) &= \left(\bigoplus_{k\in\mathcal{K}^+} \mathfrak{C}(\mu_k^+,\nu_k^+)\right) \oplus \left(\bigoplus_{k\in\mathcal{K}^-} \mathfrak{C}(\mu_k^-,\nu_k^-)\right) \oplus \mathfrak{C}(\mu^=,\nu^=) \\ &= \left(\bigoplus_{k\in\mathcal{K}^+} \mathrm{Marg}_{\mathrm{F}}(\mu_k^+,\nu_k^+)\right) \oplus \left(\bigoplus_{k\in\mathcal{K}^-} \mathrm{Marg}_{\tilde{\mathrm{F}}}(\mu_k^-,\nu_k^-)\right) \oplus \{(\mathrm{id},\mathrm{id})_\#\eta\}.\end{aligned} \tag{13}$$

Remark 2.37. As previously noted, if $W_1(\mu,\nu) < +\infty$, then $\mathfrak{C}(\mu,\nu) = \mathcal{O}(\mu,\nu)$. However, in case $W_1(\mu,\nu) = +\infty$, one can not replace $\mathfrak{C}(\mu,\nu)$ with $\mathcal{O}(\mu,\nu)$ in our statements. For instance, define

$$\begin{cases}\mu = \sum_{k\geq 1} \frac{1}{2^k}\delta_{-k} + \delta_0 + \sum_{k\geq 1}\frac{1}{2^k}\delta_k \\ \nu = \sum_{k\geq 1} \frac{1}{2^k}\delta_{-k-2^k} + \delta_0 + \sum_{k\geq 1}\frac{1}{2^k}\delta_{k+2^k}\end{cases}.$$

As $\pi^* := \sum_{k\geq 1}\frac{1}{2^k}\delta_{(-k,-k-2^k)} + \delta_{(0,0)} + \sum_{k\geq 1}\frac{1}{2^k}\delta_{(k,k+2^k)} \in \mathrm{Marg}(\mu,\nu)$ is concentrated on $\left(\tilde{\mathrm{F}} \cap]-\infty,0]^2\right) \cup \left(\mathrm{F}\cap[0,+\infty[^2\right)$, by Lemma 2.31, π^* belongs to $\mathfrak{C}(\mu,\nu)$. By Theorem 2.6, $W_1(\mu,\nu) = J(\pi^*) = 2\sum_{k\geq 1}\frac{1}{2^k}2^k + 1 = +\infty$. In this example, the decomposition yields $E^+(\mu,\nu) =]0,+\infty[$, $E^-(\mu,\nu) =]-\infty,0[$, and $E^=(\mu,\nu) = \{0\}$. Hence $\mu_1^+ = \sum_{k\geq 1}\frac{1}{2^k}\delta_k$, $\nu_1^+ = \sum_{k\geq 1}\frac{1}{2^k}\delta_{k+2^k}$, $\mu_1^- = \sum_{k\geq 1}\frac{1}{2^k}d_{-k}$, $\nu_1^- = \sum_{k\geq 1}\frac{1}{2^k}\delta_{-k-2^k}$, and $\mu^= = \nu^= = \delta_0$. Clearly $\mu\otimes\nu \notin \mathcal{O}(\mu_1^+,\nu_1^+) \oplus \mathcal{O}(\mu_1^-,\nu_1^-) \oplus \mathcal{O}(\mu^=,\nu^=)$, and Theorem 2.36 does not hold for $\mathcal{O}(\mu,\nu)$ instead of $\mathfrak{C}(\mu,\nu)$.

Definition 2.38. We define the set of crossings as

$$\mathcal{C} = \left\{((x_1,y_1),(x_2,y_2)) \in \mathbb{R}^2\times\mathbb{R}^2 \; ; \; x_1 \leq x_2 \text{ and } y_2 \leq y_1\right\}.$$

A crossing $((x_1,y_1),(x_2,y_2)) \in \mathcal{C}$ is said to be free if $y_1 \leq x_1$ or $x_2 \leq y_2$. Let $\mathcal{C}_0$ denote the set of free crossing and define $\mathcal{C}_+ = \mathcal{C}\setminus\mathcal{C}_0$. We define the transport plans avoiding non-free crossing as

$$\mathrm{Marg}^{\mathcal{C}_+^c}(\mu,\nu) = \left\{\pi \in \mathrm{Marg}(\mu,\nu) \; ; \; \exists\Gamma \in \mathcal{B}(\mathbb{R}^2), \pi(\Gamma^c) = 0 \text{ and } \Gamma^2 \cap \mathcal{C}_+ = \emptyset\right\}.$$

Remark 2.39.

1. For notational simplicity, we adopt a non-symmetrical definition of crossing, i.e., a definition that does not include pairs $((x_1,y_1),(x_2,y_2)) \in \mathbb{R}^2\times\mathbb{R}^2$ such that $x_2 \leq x_1$ and $y_1 \leq y_2$. Had we chosen the symmetrical definition of $\mathcal{C}$ (and $\mathcal{C}_0$, $\mathcal{C}_+$), we would have obtained the same set $\mathrm{Marg}^{\mathcal{C}_+^c}(\mu,\nu)$.

2. Consider $(x_1,y_1),(x_2,y_2) \in \mathbb{R}^2$ with $x_1 \leq x_2$. Then, $((x_1,y_1),(x_2,y_2)) \in \mathcal{C}_+^c$ if and only if $|y_1-x_1| + |y_2-x_2| \leq |y_2-x_1| + |y_1-x_2|$. Therefore, $\mathrm{Marg}^{\mathcal{C}_+^c}(\mu,\nu)$ is the set of transport plan π concentrated on a set Γ satisfying the following condition:

$$\forall (x_1,y_1),(x_2,y_2) \in \Gamma, |y_1-x_1| + |y_2-x_2| \leq |y_2-x_1| + |y_1-x_2|. \tag{14}$$

As this condition is weaker than cyclical monotonicity, we have $\mathfrak{C}(\mu,\nu) \subset \mathrm{Marg}^{\mathcal{C}_+^c}(\mu,\nu)$.

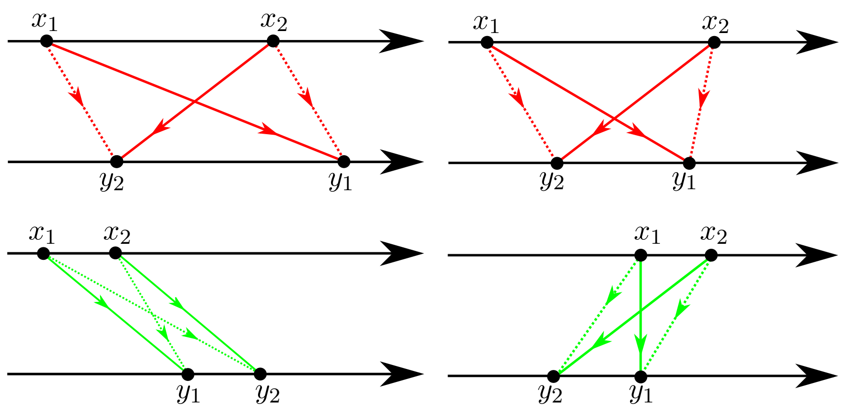

Figure 5: Free crossings (in green) and non-free crossings (in red).

Proposition 2.40. *Consider* $(\mu,\nu)\in\mathcal{M}_+^2$ *and define* $\Gamma=\left(\bigcup_{k\in\mathcal{K}^+}\mathrm{F}\cap A_k^+\right)\cup\left(\bigcup_{k\in\mathcal{K}^-}\tilde{\mathrm{F}}\cap A_k^-\right)\cup(A^{=}\cap\mathrm{D})$. *Then*

$$\mathfrak{C}(\mu,\nu)=\mathrm{Marg}^{\mathcal{C}_+^c}(\mu,\nu)=\mathrm{Marg}_\Gamma(\mu,\nu). \tag{15}$$

Proof. We prove the inclusions $\mathfrak{C}(\mu,\nu)\subset\mathrm{Marg}^{\mathcal{C}_+^c}(\mu,\nu)\subset\mathrm{Marg}_\Gamma(\mu,\nu)\subset\mathfrak{C}(\mu,\nu)$. The first inclusion corresponds to Point 2 of Remark 2.39. For the second inclusion, by Remark 2.16, the proof of Proposition 2.15 is only based on the fact that π is concentrated on a set satisfying Equation (14). By Point 2 of Remark 2.39, we may thus replace $\pi\in\mathfrak{C}(\mu,\nu)$ with $\pi\in\mathrm{Marg}^{\mathcal{C}_+^c}(\mu,\nu)$ in the statement of Proposition 2.15. Since the proof of Theorem 2.36 and its intermediate steps rely on Proposition 2.15, we may also replace $\mathfrak{C}(\mu,\nu)$ with the larger set $\mathrm{Marg}^{\mathcal{C}_+^c}(\mu,\nu)$ in the statement of Theorem 2.36. Thus, every element of $\mathrm{Marg}^{\mathcal{C}_+^c}(\mu,\nu)$ belongs to $\sum_{k\in\mathcal{K}^+}\mathrm{Marg}_\mathrm{F}(\mu_k^+,\nu_k^+)+\sum_{k\in\mathcal{K}^-}\mathrm{Marg}_{\tilde{\mathrm{F}}}(\mu_k^-,\nu_k^-)+\{(\mathrm{id},\mathrm{id})_\#\eta\}$. Therefore, elements of $\mathrm{Marg}^{\mathcal{C}_+^c}(\mu,\nu)$ are concentrated on Γ, which proves the second inclusion. The third inclusion is an immediate consequence of Remark 2.32. □

Remark 2.41. In the case of L^p transport with $p>1$, it is well known that a transport plan is optimal if and only if there exists a set Γ such that, for every $(x,y),(x',y')\in\Gamma$, $|y-x|^p+|y'-x'|^p\le|y-x'|^p+|y'-x|^p$. The first equality of Proposition 2.40 implies that this condition, sometimes referred as 2-cyclical monotonicity [22], is also equivalent to cyclical monotonicity when $p=1$. In case $p<1$, the situation is less clear: we refer the reader to [14, Section 2.1] for a discussion on the geometric consequences of the previous condition and to [11, Part 2] for general informations about optimal transportation for concave costs.

Remark 2.42 (The decomposition of Kellerer)**.** In [15, Proposition 1.20], Kellerer stated that if $\mu\le_{\mathrm{st}}\nu$ and (μ,ν) is reduced in the sense of Kellerer, then

$$\varphi:(\pi_k^+)_{k\in\mathcal{K}^+}\in\prod_{k\in\mathcal{K}^+}\mathrm{Marg}_\mathrm{F}\left(\mu\llcorner_{[a_k^+,b_k^+[},\nu\llcorner_{]a_k^+,b_k^+]}\right)\mapsto\sum_{k\in\mathcal{K}^+}\pi_k^+\in\mathrm{Marg}_\mathrm{F}(\mu,\nu).$$

is a bijection. However, by Proposition 2.11 and Remark 2.37, we know that $\mathrm{Marg}_\mathrm{F}(\mu,\nu)=\mathfrak{C}(\mu,\nu)$. Hence, this result is a particular case of Theorem 2.36. Observe that, in this case, $\eta=\mu^-=0$, and the issue of mass allocation at boundary points vanishes.

2.3 Admissible decompositions and characterisation of barrier points

In the previous subsection, we established that $\mathfrak{C}(\mu,\nu)$ can be written as sum of spaces $\mathfrak{C}(\gamma_1,\gamma_2)$, where $(\gamma_1,\gamma_2)\in\mathcal{M}_+^2$. This naturally raises the question of the existence of a better way to decompose $\mathfrak{C}(\mu,\nu)$. In the following, we provide a definition of admissible decomposition. Then, we give a way to compare admissible decompositions, and claim that the decomposition $\mathcal{D}_\mathcal{K}$ (see Definition 2.22) is the best admissible decomposition. The proof of this result is left to the appendix (Theorem C.17).

Definition 2.43. A set $\mathcal{D}\subset\mathcal{M}_+^2$ of pairs of finite measures is said to be an admissible decomposition of (μ,ν) if there exists a measure $\theta\in\mathcal{M}_+(\mathbb{R})$ and a countable family $((\mu_i,\nu_i))_{i\in\mathcal{I}}$ of non-equal pairs of $\mathcal{M}_+^2$ such that $\mathcal{D}=\{(\mu_i,\nu_i)\}_{i\in\mathcal{I}}\cup\{(\theta,\theta)\}$, and the two following conditions are satisfied:

1. The set $\mathcal{D}$ induces a decomposition of $\mathfrak{C}(\mu,\nu)$, that is,

$$\mathfrak{C}(\mu,\nu)=\left(\bigoplus_{i\in\mathcal{I}}\mathfrak{C}(\mu_i,\nu_i)\right)\oplus\{(\mathrm{id},\mathrm{id})_\#\theta\};\tag{16}$$

2. The family $(]a_i,b_i[)_{i\in\mathcal{I}}:=(]\min(s_{\mu_i},s_{\nu_i}),\max(S_{\mu_i},S_{\nu_i})[)_{i\in\mathcal{I}}$ is a family of disjoint sets, and θ is concentrated on $\cup_{i\in\mathcal{I}}]a_i,b_i[^c$.

We denote by $\mathcal{A}$ the set of admissible decompositions.

Example 2.44.

1. By Theorem 2.36, $\mathcal{D}_\mathcal{K}$ satisfies the first condition of Definition 2.43. Since the second condition is obviously satisfied, $\mathcal{D}_\mathcal{K}$ belongs to $\mathcal{A}$.

2. Of course, $\mathcal{D}_\mathcal{K}$ is not the only possible decomposition. Consider $\mu=\mathbb{1}_{[0,3]}\cdot\mathcal{L}^1$ and $\nu=2\mathbb{1}_{[0,1/2]\cup[3/2,2]\cup[3,7/2]}\cdot\mathcal{L}^1$. According to Theorem 2.36, we have

$$\begin{aligned}\mathfrak{C}(\mu,\nu)&=\mathfrak{C}(\mathbb{1}_{[0,1]}\cdot\mathcal{L}^1,2\mathbb{1}_{[0,1/2]}\cdot\mathcal{L}^1)\oplus\mathfrak{C}(\mathbb{1}_{[1,2]}\cdot\mathcal{L}^1,2\mathbb{1}_{[3/2,2]}\cdot\mathcal{L}^1)\oplus\mathfrak{C}(\mathbb{1}_{[2,3]}\cdot\mathcal{L}^1,2\mathbb{1}_{[3,7/2]}\cdot\mathcal{L}^1)\\&=\mathfrak{C}(\mathbb{1}_{[0,2]}\cdot\mathcal{L}^1,2\mathbb{1}_{[0,1/2]\cup[3/2,2]}\cdot\mathcal{L}^1)\oplus\mathfrak{C}(\mathbb{1}_{[2,3]}\cdot\mathcal{L}^1,2\mathbb{1}_{[3,7/2]}\cdot\mathcal{L}^1).\end{aligned}$$

As $\mathcal{D}=\{(\mathbb{1}_{[0,2]}\cdot\mathcal{L}^1,2\mathbb{1}_{[0,1/2]\cup[3/2,2]}\cdot\mathcal{L}^1),(\mathbb{1}_{[2,3]}\cdot\mathcal{L}^1,2\mathbb{1}_{[3,7/2]}\cdot\mathcal{L}^1),(0,0)\}$ also satisfies the second condition, $\mathcal{D}$ is admissible. The reader may verify that $\mathcal{D}':=\{(\mathbb{1}_{[0,1]\cup[2,3]}\cdot\mathcal{L}^1,2\mathbb{1}_{[0,1/2]\cup[7/2,3]}\cdot\mathcal{L}^1),(\mathbb{1}_{[1,2]}\cdot\mathcal{L}^1,2\mathbb{1}_{[3/2,2]}\cdot\mathcal{L}^1),(0,0)\}$ satisfies the first condition, but not the second.

We now define a way to compare admissible decompositions. Loosely speaking, $\mathcal{D}_1$ is better than $\mathcal{D}_2$ if $\mathcal{D}_1$ is obtained by decomposing $\mathcal{D}_2$. More precisely, if $\mathcal{D}_1$ is a better decomposition than $\mathcal{D}_2$, every part $\mathfrak{C}(\mu^2_{i_2},\nu^2_{i_2})$ of the decomposition of $\mathfrak{C}(\mu,\nu)$ induced by $\mathcal{D}_2$ is a sum of parts of the decomposition of $\mathfrak{C}(\mu,\nu)$ induced by $\mathcal{D}_1$ (and a measure concentrated on D).

Definition 2.45. Consider two admissible decompositions $\mathcal{D}_1=\{(\mu^1_{i_1},\nu^1_i)\}_{i_1\in\mathcal{I}_1}\cup\{(\theta_1,\theta_1)\}$ and $\mathcal{D}_2=\{(\mu^2_{i_2},\nu^2_{i_2})\}_{i_2\in\mathcal{I}_2}\cup\{(\theta_2,\theta_2)\}\in\mathcal{A}$. We say that $\mathcal{D}_1$ is a finer decomposition than $\mathcal{D}_2$ if there exists a partition $(J_{i_2})_{i_2\in\mathcal{I}_2}$ of $\mathcal{I}_1$ and a family $(\theta_{i_2})_{i_2\in\mathcal{I}_2}$ of measures such that, for all $i_2\in\mathcal{I}_2$:

$$\mathfrak{C}(\mu^2_{i_2},\nu^2_{i_2})=\bigoplus_{i_1\in J_{i_2}}\mathfrak{C}(\mu^1_{i_1},\nu^1_{i_1})\oplus\{(\mathrm{id},\mathrm{id})_\#\theta_{i_2}\}.\tag{17}$$

In this case, we write $\mathcal{D}_1\preceq_\mathcal{A}\mathcal{D}_2$.

Remark 2.46. If $\mathcal{D}_1 \preceq_{\mathcal{A}} \mathcal{D}_2$, then $\mu - \theta_2 = \sum_{i_2 \in \mathcal{I}_2} \mu_{i_2} = \sum_{i_2 \in \mathcal{I}_2} \theta_{i_2} + \sum_{i_1 \in J_{i_2}} \mu^1_{i_1} = \sum_{i_2 \in \mathcal{I}_2} \theta_{i_2} + \sum_{i_1 \in \mathcal{I}_1} \mu^1_{i_1} = \sum_{i_2 \in \mathcal{I}_2} \theta_{i_2} + \mu - \theta_1$. In particular, $\theta_2 \geq \theta_1$, meaning that $\mathcal{D}_1$ provides more information about the fixed parts of elements of $\mathfrak{C}(\mu, \nu)$ than $\mathcal{D}_2$.

We can now state our main result about admissible decompositions. The proof will be done in the appendix (Theorem C.17).

Theorem 2.47. *The relation $\preceq_{\mathcal{A}}$ defines a partial order on $\mathcal{A}$. Moreover, the poset $(\mathcal{A}, \preceq_{\mathcal{A}})$ admits a minimum, which is given by $\mathcal{D}_{\mathcal{K}} = \{(\mu^+_k, \nu^+_k)\}_{k \in \mathcal{K}^+} \cup \{(\mu^-_k, \nu^-_k)\}_{k \in \mathcal{K}^-} \cup \{(\mu^=, \mu^=)\}$.*

In other terms $\mathcal{D}_{\mathcal{K}}$ is the best admissible decomposition.

Remark 2.48.

1. The definition of admissible decomposition involves implicit choices that we want to discuss. Note that the condition associated to Equation (16) is the minimum one should require from a decomposition of (μ, ν) in the context of study of $\mathfrak{C}(\mu, \nu)$. Precisely, if $\mathcal{D} = \{(\mu_i, \nu_i)\}_{i \in \mathcal{I}}$ is a decomposition, one should have:
$$\mathfrak{C}(\mu, \nu) = \bigoplus_{i \in \mathcal{I}} \mathfrak{C}(\mu_i, \nu_i). \tag{18}$$
We aim to show that Condition (18) alone is insufficient to guarantee that the set of decompositions admits a minimum. Indeed, if we only require Condition (18), then one can indefinitely split an atom $\delta_{(1,1)}$ in smaller parts to obtain a finer decomposition. This counterexample could be avoided, if additionally to Condition (18), we require the existence of a family of disjoint set $(A_i)_{i \in \mathcal{I}}$ such that, for all $i \in \mathcal{I}$, elements of $\mathfrak{C}(\mu_i, \nu_i)$ are concentrated on A_i. The following example shows that this condition is still not strong enough for our set of admissible decomposition to admit a minimum. Let $\mathcal{A}'$ denote the set of decompositions $\mathcal{D}$ satisfying Condition (18), and for which there exists such family $(A_i)_{i \in \mathcal{I}}$. Now, define $\mu = \delta_0$, $\nu = \mathbb{1}_{[0,1]} \cdot \mathcal{L}^1$, let $\mathcal{D} = \{(\mu_i, \nu_i)\}_{i \in \mathcal{I}}$ be an element of $\mathcal{A}'$, and fix $j \in \mathcal{I}$. If m_j stands for the median of ν_j, then the pair (ν^1_j, ν^2_j) defined by $\nu^1_j = \nu_j \llcorner_{[0, m_j]}$ and $\nu^2_j = \nu_j \llcorner_{[m_j/2, 1]}$ satisfies $\nu_j = \nu^1_j + \nu^2_j$, $\nu^1_j(\mathbb{R}) = \nu^2_j(\mathbb{R}) = \nu_j(\mathbb{R})/2$, and (ν^1_j, ν^2_j) is a pair of singular measure. Thus, $\mathcal{D}' := \{(2^{-1}\mu_j, \nu^1_j)\} \cup \{(2^{-1}\mu_j, \nu^2_j)\} \cup \{(\mu_i, \nu_i)\}_{i \in \mathcal{I} \setminus \{j\}}$ belongs to $\mathcal{A}'$, and we have $\mathfrak{C}(\mu_j, \nu_j) = \{\delta_0 \otimes \nu_j\} = \{\delta_0 \otimes \nu^1_j\} \oplus \{\delta_0 \otimes \nu^2_j\} = \mathfrak{C}(2^{-1}\mu_j, \nu^1_j) \oplus \mathfrak{C}(2^{-1}\mu_j, \nu^2_j)$. Therefore $\mathcal{D}' \prec \mathcal{D}$, which proves that $(\mathcal{A}', \preceq)$ does not admit a minimum.

2. In the appendix, the class of admissible decompositions shall be enlarged to a class $\mathcal{A}^*$ by weakening the second condition of Definition 2.43: an analogue of Theorem 2.47 will be established (see Theorem C.15). For more details, we refer to the appendix (Definition C.2 and Remark C.4, Point 1).

3. By Remark 2.46 and Theorem 2.47, $\mu^=$ represents the largest fixed part in any admissible decomposition. The following example establishes that this information is not complete. Define $\mu = \delta_0 + 2\delta_1$ and $\nu = 2\delta_1 + \delta_2$. The reader may verify that for all $\pi \in \mathfrak{C}(\mu, \nu)$, $\pi \llcorner_{\mathrm{D}} = \delta_{(1,1)} > 0 = (\mathrm{id}, \mathrm{id})_{\#}\mu^=$. In the appendix, we give a refinement of Theorem 2.36 for which the associated decomposition $\mathcal{D}^*_{\mathcal{K}}$ has a fixed part η satisfying $(\mathrm{id}, \mathrm{id})_{\#}\eta(A) = \min_{\pi \in \mathfrak{C}(\mu,\nu)} \pi(\mathrm{D} \cap A)$. By relaxing the second condition of Definition 2.43, we shall introduce a larger class of admissible decompositions $\mathcal{A}^*$ and a richer relation $\preceq_{\mathcal{A}^*}$ on $\mathcal{A}^*$, for which the minimum is not $\mathcal{D}_{\mathcal{K}}$ but a refined version $\mathcal{D}^*_{\mathcal{K}}$.

A key ingredient in proving that $\mathcal{D}_{\mathcal{K}}$ is the minimum of $(\mathcal{A}, \preceq_{\mathcal{A}})$ is the equality $E^{=}(\mu,\nu) = \mathcal{B}(\mu,\nu)$. We already proved the inclusion $E^{=}(\mu,\nu) \subset \mathcal{B}(\mu,\nu)$, and now aim to prove the reverse inclusion $\mathcal{B}(\mu,\nu) \subset E^{=}(\mu,\nu)$. If this inclusion were not satisfied, then $\mathcal{D}_{\mathcal{K}}$ could not be the minimum of $(\mathcal{A}, \preceq_{\mathcal{A}})$. Indeed, one may verify that if there exists $x \in]a_k^+, b_k^+[$ that belongs to $\mathcal{B}(\mu,\nu)$, then $\mathfrak{C}(\mu_k^+, \nu_k^+) = \mathfrak{C}(\mu_x^1, \nu_x^1) \oplus \mathfrak{C}(\mu_x^2, \nu_x^2) \oplus \{(F_\mu^-(x) - F_\nu^+(x))\delta_{(x,x)}\}$, where $\mu_x^1 = (F_\mu^+(a_k^+) - F_\nu^+(a_k^+))\delta_{a_k^+} + \mu_{\llcorner]a_k^+, x[}$, $\nu_x^1 = \nu_{\llcorner]a_k^+, x[} + (F_\mu^-(x) - F_\nu^-(x))\delta_x$, $\mu_x^2 = (F_\mu^+(x) - F_\nu^+(x))\delta_x + \mu_{]x, b_k^+[}$, and $\nu_x^2 = \nu_{\llcorner]x, b_k^+[} + (F_\mu^-(b_k^+) - F_\nu^-(b_k^+))\delta_{b_k^+}$. Therefore, if we define $\theta = (F_\mu^-(x) - F_\nu^+(x))\delta_{(x,x)}$, then $\mathcal{D}' = \{(\mu_j^+, \nu_j^+)\}_{j \in \mathcal{K}^+ \setminus \{j\}} \cup \{(\mu_k^-, \nu_k^-)\}_{k \in \mathcal{K}^-} \cup \{(\mu_x^1, \nu_x^1), (\mu_x^2, \nu_x^2)\} \cup \{(\mu^= + \theta, \mu^= + \theta)\}$ belongs to $\mathcal{A}$ and $\mathcal{D}' \preceq_{\mathcal{A}} \mathcal{D}$. The proof of the equality $E^{=}(\mu,\nu) = \mathcal{B}(\mu,\nu)$ relies on the Strassen-type theorem for the order $\leq_{\mathrm{F}}$ (Definition 2.33), which involves the notion of strongly multiplicative measure.

Definition 2.49 (Large strong multiplicativity)**.** A measure $\pi \in \mathcal{M}_+(\mathbb{R}^2)$ is said to be (largely) strongly multiplicative if there exists $\eta_1, \eta_2 \in \mathcal{M}_+^\sigma(\mathbb{R})$ such that $\pi = (\eta_1 \otimes \eta_2)_{\llcorner \mathrm{F}}$.

For any pair (γ_1, γ_2) defined on the same measurable space, we write $\gamma_1 \ll \gamma_2$ if γ_1 is absolutely continuous with respect to γ_2. In this case, we denote by $\frac{d\gamma_1}{d\gamma_2} \cdot \gamma_2 = \gamma_1$ the Radon-Nikodym derivative of γ_1 with respect to γ_2. If $\gamma_1 \ll \gamma_2$ and $\gamma_2 \ll \gamma_1$, we say that γ_1 and γ_2 are equivalent and we write $\gamma_1 \sim \gamma_2$. We can now state the Strassen-type theorem of Kellerer associated to $\leq_{\mathrm{F}}$ [15, Theorem 3.6].

Theorem 2.50 (Strassen-type theorem for $\leq_{\mathrm{F}}$)**.** *Consider $(\gamma_1, \gamma_2) \in \mathcal{M}_+(\mathbb{R})^2$. The relation $\gamma_1 \leq_{\mathrm{F}} \gamma_2$ holds if and only if* $\mathrm{Marg}(\gamma_1, \gamma_2)$ *contains a strongly multiplicative measure,* i.e.*, there exists $\eta_1, \eta_2 \in \mathcal{M}_+^\sigma(\mathbb{R})$ such that*

$$(\eta_1 \otimes \eta_2)_{\llcorner \mathrm{F}} \in \mathrm{Marg}(\gamma_1, \gamma_2). \tag{19}$$

In this case, $\mathrm{Marg}(\gamma_1, \gamma_2)$ *contains exactly one strongly multiplicative measure. Moreover, one can assume $\eta_1 \sim \gamma_1$ and $\eta_2 \sim \gamma_2$ in Equation* (27)*.*

We denote by $\mathcal{K}_{\mathrm{F}}(\gamma_1, \gamma_2)$ the unique strongly multiplicative measure in $\mathrm{Marg}(\gamma_1, \gamma_2)$ when $\gamma_1 \leq_{\mathrm{F}} \gamma_2$ and call it Kellerer transport plan. Note that $\mathcal{K}_{\mathrm{F}}(\gamma_1, \gamma_2)$ belongs to $\mathrm{Marg}_F(\gamma_1, \gamma_2) = \mathfrak{C}(\gamma_1, \gamma_2)$. Our decomposition allows us to extend the definition of Kellerer transport plans to measures that are not in the reinforced large stochastic order.

Definition 2.51 (Generalized Kellerer transport plan)**.** Set $i : (x, y) \in \mathbb{R}^2 \mapsto (y, x) \in \mathbb{R}^2$. By Point 1 of Theorem 2.36 and Theorem 2.50, the transport plan $\mathcal{K}(\mu,\nu) := \sum_{k \in \mathcal{K}^+} \mathcal{K}_{\mathrm{F}}(\mu_k^+, \nu_k^+) + \sum_{k \in \mathcal{K}^-} i_\# \mathcal{K}_{\mathrm{F}}(\nu_k^-, \mu_k^-) + (\mathrm{id}, \mathrm{id})_\# \mu^=$ is well defined. By Equation (13), $\mathcal{K}(\mu,\nu)$ belongs to $\mathfrak{C}(\mu,\nu)$.

The following result is a characterization of barrier points in terms of cumulative distribution functions. Its proof relies on the Strassen-type result of Kellerer for $\leq_{\mathrm{F}}$ and our decomposition. We refer to Definition 2.26 for the definition of $E^{=}(\mu,\nu)$ and Definition 2.8 for the definition of $\mathcal{B}(\mu,\nu)$.

Proposition 2.52. *Consider $(\mu,\nu) \in \mathcal{M}_+^2$. Then $\mathcal{B}(\mu,\nu) = E^{=}(\mu,\nu)$.*

Proof. According to Lemma 2.21, $E^= \subset \mathcal{B}$. It remains to prove $\mathcal{B} \subset E^=$. Suppose, to derive a contradiction, that $\mathcal{B} \cap E^+$ is non-empty. Then, there exists $x \in \mathcal{B}$ and $k \in \mathcal{K}^+$ such that $x \in]a_k^+, b_k^+[$. Since $\mathcal{K}(\mu,\nu) \in \mathfrak{C}(\mu,\nu)$ and $x \in \mathcal{B}$, we have $\mathcal{K}(\mu,\nu)(] - \infty, x[\times]x, +\infty[) = 0$. This implies $\mathcal{K}_{\mathrm{F}}(\mu_k^+, \nu_k^+)(] - \infty, x[\times]x, +\infty[) = 0$. According to Theorem 2.50, there exists two σ-finite measures $\eta_1, \eta_2 \in \mathcal{M}_+^\sigma(\mathbb{R})$ such that $\mathcal{K}_{\mathrm{F}}(\mu_k^+, \nu_k^+) = (\eta_1 \otimes \eta_2)_{\llcorner \mathrm{F}}$, $\mu_k^+ \sim \eta_1$ and $\nu_k^+ \sim \eta_2$. Hence, $0 =$

$\mathcal{K}_{\mathrm{F}}(\mu_k^+, \nu_k^+)(]-\infty, x[\times]x, +\infty[) = (\eta_1 \otimes \eta_2)_{\lfloor \mathrm{F}}(]-\infty, x[\times]x, +\infty[) = \eta_1(]-\infty, x[)\eta_2(]x, +\infty[)$, which forces either $\eta_1(]-\infty, x[) = 0$ or $\eta_2(]x, +\infty[) = 0$. Thus, $\mu_k^+(]-\infty, x[) = 0$ or $\nu_k^+(]x, +\infty[) = 0$. Moreover, in the proof of Point 2 of Proposition 2.35, we established that $]a_k^+, b_k^+[=]s_{\mu_k^+}, S_{\nu_k^+}[$. By Point 2 of Remark 2.34, this is a contradiction. Hence, $\mathcal{B} \cap E^+ = \emptyset$. Similarly, $\mathcal{B} \cap E^- = \emptyset$. Since $\mathbb{R} = E^+ \uplus E^- \uplus E^=$, we get $\mathcal{B} \subset (E^+)^c \cap (E^-)^c = E^=$. □

Remark 2.53. As $E^= = (E^+)^c \cup (E^-)^c = \big(\{F_\mu^+ > F_\nu^+\} \cap \{F_\mu^- > F_\nu^-\}\big)^c \cup \big(\{F_\mu^+ < F_\nu^+\} \cap \{F_\mu^- < F_\nu^-\}\big)^c = \{F_\mu^+ = F_\nu^+\} \cup \{F_\mu^- = F_\nu^-\} \cup \{F_\nu^- \leq F_\mu^- < F_\mu^+ \leq F_\nu^+\} \cup \{F_\mu^- \leq F_\nu^- < F_\nu^+ \leq F_\mu^+\}$, Equation (4) follows from Proposition 2.52.

3 Convergence of the solutions to the entropically regularized problem

Throughout this section, we fix a pair $(\mu, \nu) \in \mathcal{M}_+^2$ of measures. Given a measurable space S and $(\gamma_1, \gamma_2) \in \mathcal{M}_+(S)^2$, we denote by $\mathrm{Ent}(\gamma_1|\gamma_2)$ the entropy of γ_1 relatively to γ_2 defined by:

$$\mathrm{Ent}(\gamma_1|\gamma_2) = \begin{cases} \int_S \log\left(\frac{d\gamma_1}{d\gamma_2}\right) \frac{\mathrm{d}\gamma_1}{\mathrm{d}\gamma_2}\, \mathrm{d}\gamma_2 & \text{if } \gamma_1 \ll \gamma_2 \\ +\infty & \text{otherwise} \end{cases},$$

where we recall that $\gamma_1 \ll \gamma_2$ means that γ_1 is absolutely continuous with respect to γ_2 and $\frac{d\gamma_1}{d\gamma_2}$ stands for the Radon-Nykodim derivative of γ_1 with respect to γ_2. We recall the $\mathrm{Ent}(\cdot|\gamma_2)$ is valued in $]-\infty, +\infty]$, strictly convex on $\mathcal{M}_+(S)$ and lower semi-continuous. For all $\varepsilon > 0$, let

$$J_\varepsilon : \pi \in \mathrm{Marg}(\mu, \nu) \mapsto \int_{\mathbb{R}^2} |y - x|\, \mathrm{d}\pi(x, y) + \varepsilon \mathrm{Ent}(\pi|\mu \otimes \nu)$$

denote the transport cost with regularization parameter ε, and define

$$W_1^\varepsilon(\mu, \nu) = \min_{\pi \in \mathrm{Marg}(\mu,\nu)} J_\varepsilon(\pi).$$

For all $\varepsilon > 0$, there exists a notion analogue to cyclical monotonicity in the original optimal transport problem. This notion was introduced by Bernton, Ghosal and Nutz in [5] for more general costs and characterizes the solution of the minimization problem associated to J_ε when $W_1^\varepsilon(\mu, \nu) < +\infty$.

Definition 3.1. Consider $\varepsilon > 0$ and $\pi \in \mathrm{Marg}(\mu, \nu)$. The transport plan $\pi \in \mathrm{Marg}(\mu, \nu)$ is said to be ε-cyclically invariant if there exists $f : \mathbb{R}^2 \to [0, +\infty[$ such that $\pi = f \cdot \mu \otimes \nu$, and, for all $((x_i, y_i))_{i \in [\![1,n]\!]} \in (\mathbb{R}^2)^n$,

$$\prod_{i=1}^n \exp\left(-\frac{1}{\varepsilon}|y_i - x_i|\right) f(x_i, y_i) = \prod_{i=1}^n \exp\left(-\frac{1}{\varepsilon}|y_{i+1} - x_i|\right) f(x_i, y_{i+1}),$$

where y_{n+1} stands for y_1.

As shown by the same authors in [12, Theorem 1.3], there exists a unique ε-cyclically monotone transport plan, even when $W_1^\varepsilon(\mu, \nu) = +\infty$. Analogously to Theorem 2.6 for the classic transport problem, they proved that it coincides with the unique minimizer of J_ε when $W_1^\varepsilon(\mu, \nu)$ is finite.

Theorem 3.2. *Consider $(\mu, \nu) \in \mathcal{M}_+^2$ and $\varepsilon > 0$.*

1. *There exists a unique ε-cyclically invariant transport plan from μ to ν. In the following, we denote this measure by $\pi_\varepsilon \in \mathrm{Marg}(\mu,\nu)$.*

2. *In case $W_1^\varepsilon(\mu,\nu)$ is finite, then J_ε admits a unique minimizer, and this minimizer is equal to π_ε.*

We now study the behaviour of $(\pi_\varepsilon)_{\varepsilon>0}$ when $\varepsilon \to 0^+$. This problem is related to the minimization of $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ among $\mathfrak{C}(\mu,\nu)$. Indeed, in the case of measure with finite support, every optimal transport plan has finite entropy, and by strict convexity of $\mathrm{Ent}(\cdot|\mu\otimes\nu)$, there exists a unique minimizer of $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ among $\mathfrak{C}(\mu,\nu)$. In this case, it is well known that $\lim_{\varepsilon\to 0^+}\pi_\varepsilon = \mathrm{argmin}_{\pi\in\mathfrak{C}(\mu,\nu)}\mathrm{Ent}(\pi|\mu\otimes\nu)$. In the following subsection, we verify that $\mathcal{K}(\mu,\nu)$ is a minimizer of $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ among $\mathfrak{C}(\mu,\nu)$ for every pair $(\mu,\nu)\in\mathcal{M}_+^2$. After, we establish that, if $W_1(\mu,\nu) < +\infty$ and there exists a optimal transport plan with finite entropy, then $\mathcal{K}(\mu,\nu) = \lim_{\varepsilon\to 0^+}\pi_\varepsilon$. Finally, we briefly discuss the convergence result of Di Marino and Louet.

3.1 $\mathcal{K}(\mu,\nu)$ minimizes $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ among $\mathfrak{C}(\mu,\nu)$.

We begin by we expressing the relative entropy of a cyclically monotone transport plan as the sum of the relative entropies of its components.

Proposition 3.3. *Consider $\pi\in\mathfrak{C}(\mu,\nu)$ and define the family $\left\{\pi_k^+ \ ; \ k\in\mathcal{K}^+\right\}\cup\left\{\pi_k^- \ ; \ k\in\mathcal{K}^-\right\}\cup\{\pi^=\}$ as in Definition 2.26. Then the entropy decomposes as*

$$\mathrm{Ent}(\pi|\mu\otimes\nu) = \sum_{k\in\mathcal{K}^+}\mathrm{Ent}(\pi_k^+|\mu\otimes\nu) + \sum_{k\in\mathcal{K}^-}\mathrm{Ent}(\pi_k^-|\mu\otimes\nu) + \mathrm{Ent}(\pi^=|\mu\otimes\nu). \tag{20}$$

Proof. According to Theorem 2.36, we have $\pi = \sum_{k\in\mathcal{K}^+}\pi_k^+ + \sum_{k\in\mathcal{K}^-}\pi_k^- + \pi^=$. Thus, $\pi \ll \mu\otimes\nu$ if and only if

$$\begin{cases} \forall k\in\mathcal{K}^+, \pi_k^+ \ll \mu\otimes\nu \\ \forall k\in\mathcal{K}^+, \pi_k^+ \ll \mu\otimes\nu \\ \pi^= \ll \mu\otimes\nu \end{cases}. \tag{21}$$

Thus, if $\pi\ll\mu\otimes\nu$ is not satisfied, then both sides of Equation (20) are equal to $+\infty$. Assume now $\pi\ll\mu\otimes\nu$ is satisfied, so that Equation (21) holds. Recall that for all $k\in\mathcal{K}^+$ (resp. $k\in\mathcal{K}^-$), π_k^+ (resp. π_k^-) in concentrated on A_k^+ (resp. A_k^-), that $\pi^=$ is concentrated on $A^=$, and that $\{A_k^+\}_{k\in\mathcal{K}^+}\cup\{A_k^-\}_{k\in\mathcal{K}^-}\cup\{A^=\}$ is a family of disjoint sets. In particular, for all $k\in\mathcal{K}^+$, $\pi-\pi_k^+ = \sum_{j\in\mathcal{K}^+\setminus\{k\}}\pi_j^+ + \sum_{j\in\mathcal{K}^-}\pi_j^- + \pi^=$ is concentrated on $\uplus_{j\in\mathcal{K}^+\setminus\{k\}}A_j^+\uplus\uplus_{j\in\mathcal{K}^-}A_j^-\uplus A^= \subset \left(A_k^+\right)^c$. Therefore, one can assume that $\frac{\mathrm{d}\pi-\pi_k^+}{\mathrm{d}\mu\otimes\nu}$ vanishes on A_k^+. Thus, we have

$$\int_{\mathbb{R}^2}\log\left(\frac{\mathrm{d}\pi}{\mathrm{d}\mu\otimes\nu}\right)\,\mathrm{d}\pi_k^+ = \int_{A_k^+}\log\left(\frac{\mathrm{d}\pi_k^+}{\mathrm{d}\mu\otimes\nu}+\frac{\mathrm{d}\pi-\pi_k^+}{\mathrm{d}\mu\otimes\nu}\right)\,\mathrm{d}\pi_k^+ = \int_{\mathbb{R}^2}\log\left(\frac{\mathrm{d}\pi_k^+}{\mathrm{d}\mu\otimes\nu}\right)\mathrm{d}\pi_k^+. \tag{22}$$

Similarly, for all $k\in\mathcal{K}^-$, $\int_{\mathbb{R}^2}\log\left(\frac{\mathrm{d}\pi}{\mathrm{d}\mu\otimes\nu}\right)\,\mathrm{d}\pi_k^- = \int_{\mathbb{R}^2}\log\left(\frac{\mathrm{d}\pi_k^-}{\mathrm{d}\mu\otimes\nu}\right)\,\mathrm{d}\pi_k^-$, and $\int_{\mathbb{R}^2}\log\left(\frac{\mathrm{d}\pi}{\mathrm{d}\mu\otimes\nu}\right)\,\mathrm{d}\pi^= =$

$\int_{\mathbb{R}^2} \log\left(\frac{\mathrm{d}\pi^=}{\mathrm{d}\mu\otimes\nu}\right)\,\mathrm{d}\pi^=$. Finally,

$$
\begin{aligned}
\mathrm{Ent}(\pi|\mu\otimes\nu) &= \sum_{k\in\mathcal{K}^+}\int_{\mathbb{R}^2}\log\left(\frac{\mathrm{d}\pi}{\mathrm{d}\mu\otimes\nu}\right)\,\mathrm{d}\pi_k^+ + \sum_{k\in\mathcal{K}^-}\int_{\mathbb{R}^2}\log\left(\frac{\mathrm{d}\pi}{\mathrm{d}\mu\otimes\nu}\right)\,\mathrm{d}\pi_k^- + \int_{\mathbb{R}^2}\log\left(\frac{\mathrm{d}\pi}{\mathrm{d}\mu\otimes\nu}\right)\,\mathrm{d}\pi^= \\
&= \sum_{k\in\mathcal{K}^+}\int_{\mathbb{R}^2}\log\left(\frac{\mathrm{d}\pi_k^+}{\mathrm{d}\mu\otimes\nu}\right)\,\mathrm{d}\pi_k^+ + \sum_{k\in\mathcal{K}^-}\int_{\mathbb{R}^2}\log\left(\frac{\mathrm{d}\pi_k^-}{\mathrm{d}\mu\otimes\nu}\right)\,\mathrm{d}\pi_k^- + \int_{\mathbb{R}^2}\log\left(\frac{d\pi^=}{d\mu\otimes\nu}\right)\,\mathrm{d}\pi^= \\
&= \sum_{k\in\mathcal{K}^+}\mathrm{Ent}(\pi_k^+|\mu\otimes\nu) + \sum_{k\in\mathcal{K}^-}\mathrm{Ent}(\pi_k^-|\mu\otimes\nu) + \mathrm{Ent}(\pi^=|\mu\otimes\nu).
\end{aligned}
$$

□

Now, to prove that $\mathcal{K}(\mu,\nu)$ minimizes $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ over $\mathfrak{C}(\mu,\nu)$, it suffices to show that, for every $k\in\mathcal{K}^+$, $\mathcal{K}(\mu,\nu)_{\llcorner A_k^+} = \mathcal{K}(\mu_k^+,\nu_k^+)$ is a minimizer of $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ over $\mathfrak{C}(\mu,\nu)$ (and similarly if $k\in\mathcal{K}^-$). Before establishing this result in Proposition 3.5, we state the following approximation result. For a proof, we refer to [10, Lemma 5.1]

Lemma 3.4 ([10], Lemma 5.1)**.** *Consider a measure $\pi\in\mathcal{M}_+(\mathbb{R}^2)$ and two Borel functions $a,b:\mathbb{R}^2\to\mathbb{R}$. For all $n\geq 1$, define $\varphi_n : t\in\mathbb{R}\mapsto \max(-n,\min(n,t))$. If $(a+b)_-\in L^1(\pi)$, then $\int_{\mathbb{R}^2}\varphi_n(a)+\varphi_n(b)\,\mathrm{d}\pi \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}^2} a+b\,\mathrm{d}\pi$.*

The following result shows that for measures $\theta_1,\theta_2\in\mathcal{M}_+(\mathbb{R})$ such that $\theta_1\leq_{\mathrm{F}}\theta_2$ (see Definition 2.33), the Kellerer transport plan $\mathcal{K}(\theta_1,\theta_2)$ (defined after Theorem 2.50) minimizes $\mathrm{Ent}(\cdot|\theta_1\otimes\theta_2)$ among cyclically monotone transport plan. In fact the statement is more general: $\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)$ minimizes $\mathrm{Ent}(\cdot|\gamma_1\otimes\gamma_2)$ for every pair $(\gamma_1,\gamma_2)\in\mathcal{M}_2(\mathbb{R})^2$ satisfying $\theta_1\ll\gamma_1$ and $\theta_2\ll\gamma_2$. This has already been proven by Di Marino and Louet in a more restricted context[11], but their proof holds in our more general setting. For ease of reading and precaution, we reproduce their proof.

Proposition 3.5. *Consider $(\gamma_1,\gamma_2)\in\mathcal{M}_+(\mathbb{R})^2$ and $\theta_1,\theta_2\in\mathcal{M}_+(\mathbb{R})$ such that $\theta_1\ll\gamma_1,\theta_2\ll\gamma_2$ and $\theta_1\leq_{\mathrm{F}}\theta_2$. Then*

$$\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)\in\mathrm{argmin}_{\pi\in\mathfrak{C}(\theta_1,\theta_2)}\mathrm{Ent}(\pi|\gamma_1\otimes\gamma_2).$$

Proof. Consider $\pi\in\mathfrak{C}(\theta_1,\theta_2)$. If $\pi\ll\gamma_1\otimes\gamma_2$ is not satisfied, then $\mathrm{Ent}(\pi|\gamma_1\otimes\gamma_2) = +\infty \geq \mathrm{Ent}(\mathcal{K}(\theta_1,\theta_2)|\gamma_1\otimes\gamma_2)$. We now assume $\pi\ll\gamma_1\otimes\gamma_2$, and define $\pi_* = \frac{\mathrm{d}\pi}{d\gamma_1\otimes\gamma_2}$. Let $\nu_1,\nu_2\in\mathcal{M}_+^\sigma(\mathbb{R})$ be such that $\nu_1\ll\theta_1$, $\nu_2\ll\theta_2$ and $\mathcal{K}(\gamma_1,\gamma_2) = (\nu_1\otimes\nu_2)_{\llcorner\mathrm{F}}$. We have

$$\frac{d\mathcal{K}(\theta_1,\theta_2)}{d\theta_1\otimes\theta_2} = \frac{d\mathcal{K}(\theta_1,\theta_2)}{d\theta_1\otimes\theta_2}\frac{d\theta_1\otimes\theta_2}{d\gamma_1\otimes\gamma_2} = \left(\mathbb{1}_{\mathrm{F}}\frac{\mathrm{d}\nu_1}{d\theta_1}\otimes\frac{\mathrm{d}\nu_2}{d\theta_2}\right)\left(\frac{d\theta_1}{d\gamma_1}\otimes\frac{d\theta_2}{d\gamma_2}\right) = \mathbb{1}_{\mathrm{F}} f\otimes g,$$

where f and g are defined by $f = \frac{\mathrm{d}\nu_1}{d\gamma_1}$ and $g = \frac{\mathrm{d}\nu_2}{d\gamma_2}$. Now, define $a = \log(f)$ and $b = \log(g)$. For all $n\geq 1$, we define $a_n := \varphi_n(a)$ and $b_n := \varphi_n(b)$. The reader may verify that, for all $(c,t)\in\mathbb{R}\times\mathbb{R}_+$, $t\log(t) - tc \geq -e^{c-1}$. For all $(x,y,n)\in\mathbb{R}^2\times\mathbb{N}^*$, by applying this inequality with $c = a_n(x)+b_n(y)+1$

[11] Precisely, they ask for γ_1 and γ_2 to be compactly supported and atomless. Their optimal transport minimizing entropy is not defined using the Strassen Theorem from Kellerer: in fact, they explicitly construct a strongly multiplicative transport plan, which by the uniqueness part of Theorem 2.50, is the Kellerer transport plan. Their proof of the Lemma does not use their specific construction, but only relies on the product form of the Kellerer transport plan.

and $t = \pi_*(x,y)$, we get $\pi_*(x,y)\log(\pi_*(x,y)) - \pi_*(x,y)(a_n(x)+b_n(y)+1) \geq -e^{a_n(x)+b_n(y)}$. Thus, as $\mathfrak{C}(\theta_1,\theta_2) = \mathrm{Marg}_{\mathrm{F}}(\theta_1,\theta_2)$, for all $n \geq 1$:

$$\begin{aligned}
\mathrm{Ent}(\pi|\gamma_1\otimes\gamma_2) &= \int_{\mathrm{F}} \pi_*(x,y)\log(\pi_*(x,y))\ \mathrm{d}\gamma_1\otimes\gamma_2(x,y)\\
&= \int_{\mathrm{F}} \pi_*(x,y)\log(\pi_*(x,y)) - \pi_*(x,y)(a_n(x)+b_n(y)+1)\ \mathrm{d}\gamma_1\otimes\gamma_2(x,y)\\
&\qquad + \int_{\mathrm{F}} \pi_*(x,y)(a_n(x)+b_n(y)+1)\ \mathrm{d}\gamma_1\otimes\gamma_2(x,y)\\
&\geq -\int_{\mathrm{F}} e^{a_n(x)+b_n(y)}\ \mathrm{d}\gamma_1\otimes\gamma_2(x,y) + \int_{\mathbb{R}^2}(a_n(x)+b_n(y)+1)\ \mathrm{d}\pi(x,y)\\
&= -\int_{\mathbb{R}^2} e^{a_n(x)+b_n(y)}\ \mathrm{d}\gamma_1\otimes\gamma_2(x,y) + \int_{\mathbb{R}^2}(a_n(x)+b_n(y)+1)\ \mathrm{d}\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)(x,y),
\end{aligned}$$

where we used that $\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)$ and π both belong to $\mathrm{Marg}(\theta_1,\theta_2)$ to establish the last equality. Since for all $(z_1,z_2)\in\mathbb{R}^2$, $(\varphi_n(z_1)+\varphi_n(z_2))_+ \leq (z_1+z_2)_+$, for all $(x,y)\in\mathbb{R}^2$, $0 \leq e^{a_n(x)+b_n(y)} \leq e^{(a(x)+b(y))_+} \leq \max(1, f(x)g(y)) \leq 1+f(x)g(y)$. As $\int_{\mathrm{F}} 1+f(x)g(y)\ \mathrm{d}\gamma_1\otimes\gamma_2(x,y) \leq (\gamma_1\otimes\gamma_2)(\mathbb{R}^2)+\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)(\mathbb{R}^2) < +\infty$ and the pointwise limits $a = \lim_{n\to+\infty} a_n$ and $b = \lim_{n\to+\infty} b_n$ hold, by dominated convergence

$$\int_{\mathrm{F}} e^{a_n(x)+b_n(y)}\ \mathrm{d}\gamma_1\otimes\gamma_2(x,y) \xrightarrow[n\to+\infty]{} \int_{\mathrm{F}} e^{a(x)+b(y)}\ \mathrm{d}\gamma_1\otimes\gamma_2(x,y) = \mathcal{K}(\theta_1,\theta_2)(\mathbb{R}^2).$$

As

$$\int_{\mathbb{R}^2}(a(x)+b(y))_-\ \mathrm{d}\mathcal{K}(\theta_1,\theta_2) = \int_{\mathrm{F}}(f(x)g(y)\log(f(x)g(y)))_-\ \mathrm{d}(\gamma_1\otimes\gamma_2)(x,y) \leq \int_{\mathbb{R}^2} 1\ \mathrm{d}\gamma_1\otimes\gamma_2 = \gamma_1\otimes\gamma_2(\mathbb{R}^2),$$

Lemma 3.4 applies. Thus,

$$\int_{\mathbb{R}^2} a_n(x)+b_n(y)\ \mathrm{d}\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)(x,y) \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}^2} a(x)+b(y)\ \mathrm{d}\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)(x,y).$$

Since $\int_{\mathbb{R}^2} a(x)+b(y)\ \mathrm{d}\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)(x,y) = \int_{\mathrm{F}}\log(f(x)g(y))\ \mathrm{d}\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)(x,y) = \mathrm{Ent}(\mathcal{K}_{\mathrm{F}}(\theta_1,\gamma_2)|\gamma_1\otimes\gamma_2)$, by gathering all our estimates together, we conclude $\mathrm{Ent}(\pi|\gamma_1\otimes\gamma_2) \geq -\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)(\mathbb{R}^2) + \mathrm{Ent}(\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)|\gamma_1\otimes\gamma_2) + \mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)(\mathbb{R}^2) = \mathrm{Ent}(\mathcal{K}_{\mathrm{F}}(\theta_1,\theta_2)|\gamma_1\otimes\gamma_2)$, which proves the lemma. ☐

We now prove that $\mathcal{K}(\mu,\nu)$ minimizes $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ among $\mathfrak{C}(\mu,\nu)$.

Theorem 3.6. *Consider $(\mu,\nu)\in\mathcal{M}_+^2$ and define $\mathcal{K}(\mu,\nu)$ as in Definition 2.51. Then*

$$\mathrm{Ent}(\mathcal{K}(\mu,\nu)|\mu\otimes\nu) = \min_{\pi\in\mathfrak{C}(\mu,\nu)} \mathrm{Ent}(\pi|\mu\otimes\nu). \tag{23}$$

Proof. Consider $\pi\in\mathfrak{C}(\mu,\nu)$, define $\mathcal{F} = ((\pi_k^+)_{k\in\mathcal{K}^+}, (\pi_k^-)_{k\in\mathcal{K}^-}, \pi^=)$ using the notation in Definition 2.26, and $\mathfrak{P}$ as in Theorem 2.36. As $\mathcal{F}\in\mathfrak{P}$, by Proposition 3.3 and Proposition 3.5, we obtain:

$$\begin{aligned}
\mathrm{Ent}(\pi|\mu\otimes\nu) &= \sum_{k\in\mathcal{K}^+}\mathrm{Ent}(\pi_k^+|\mu\otimes\nu) + \sum_{k\in\mathcal{K}^-}\mathrm{Ent}(\pi_k^-|\mu\otimes\nu) + \mathrm{Ent}(\pi^=|\mu\otimes\nu)\\
&\geq \sum_{k\in\mathcal{K}^+}\mathrm{Ent}(\mathcal{K}(\mu_k^+,\nu_k^+)|\mu\otimes\nu) + \sum_{k\in\mathcal{K}^-}\mathrm{Ent}(\mathcal{K}(\mu_k^-,\nu_k^-)|\mu\otimes\nu) + \mathrm{Ent}(\pi^=|\mu\otimes\nu)\\
&= \mathrm{Ent}(\mathcal{K}(\mu,\nu)|\mu\otimes\nu).
\end{aligned}$$

☐

3.2 Convergence of $(\pi_\varepsilon)_{\varepsilon>0}$ in two different settings.

In this part, we first prove that $\lim_{\varepsilon\to 0^+}\pi_\varepsilon = \mathcal{K}(\mu,\nu)$ when $W_1(\mu,\nu) < +\infty$ and $\mathrm{Ent}(\pi|\mu\otimes\nu) < +\infty$. Then, we briefly recall the result of Di Marino and Louet in the case of atomless marginals, when there exists $\pi \in \mathfrak{C}(\mu,\nu)$ such that $\mathrm{Ent}(\pi - (\mathrm{id},\mathrm{id})_\# \mu^= |\mu\otimes\nu)$ is finite. In this case, under some additional technical assumptions, Di Marino and Louet proved (using involved Gamma-convergence methods) that $(\pi_\varepsilon)_{\varepsilon>0}$ converges to $\mathcal{K}(\mu,\nu)$. In another part (Theorem (3.24)), we will prove a new convergence result. This result is by no mean generalizing the two precedent result, but is a new case of convergence. Under the hypotheses that non-equal components form pairs of mutually singular measures, we prove that $\mathcal{K}(\mu,\nu) = \lim_{\varepsilon\to 0^+}\pi_\varepsilon$. Observe that we do not require the existence of an optimal transport plan with finite entropy, nor — as in the case of Di Marino and Louet — the existence of a transport plan whose non-fixed part has finite entropy.

3.2.1 Convergence under the existence of an optimal transport plan with finite entropy

In case of measures with finite support, it is well known that $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ admits a unique minimizer π^* among $\mathcal{O}(\mu,\nu)$, and that $\pi^* = \lim_{\varepsilon\to 0^+}\pi_\varepsilon$. According to Theorem 3.6, we immediately get $\mathcal{K}(\mu,\nu) = \lim_{\varepsilon\to 0^+}\pi_\varepsilon$. We now extend this convergence result to a more general setting, namely when:

$$W_1(\mu,\nu) < +\infty \quad \text{and} \quad \inf_{\pi\in\mathfrak{C}(\mu,\nu)} \mathrm{Ent}(\pi|\mu\otimes\nu) < +\infty. \tag{24}$$

Note that under Condition (24), Theorem 2.6 ensures that $\mathfrak{C}(\mu,\nu) = \mathcal{O}(\mu,\nu)$. The proof follows a similar path as the one in the finite support case (which can be found in the monograph [23, Proposition 4.1]), but an additional result is required to guarantee that cluster points of $(\pi_\varepsilon)_{\varepsilon>0}$ are elements of $\mathfrak{C}(\mu,\nu)$. This result has been proven in [5, Proposition 3.2] for any continuous cost function.

Proposition 3.7. *Cluster points of $(\pi_\varepsilon)_{\varepsilon>0}$ at point 0^+ are cyclically monotone transport plans.*

Theorem 3.8. *Consider $(\mu,\nu) \in \mathcal{M}_+^2$ and assume Condition (24) holds. Then $\mathcal{K}_{\mathrm{F}}(\mu,\nu) = \lim_{\varepsilon\to 0^+}\pi_\varepsilon$.*

Proof. Let π^* be a cluster point of $(\pi_\varepsilon)_{\varepsilon>0}$. By Proposition 3.7, $\pi^* \in \mathfrak{C}(\mu,\nu)$. Fix $\varepsilon > 0$ and $\pi \in \mathfrak{C}(\mu,\nu)$. As $J(\pi_\varepsilon) + \varepsilon\mathrm{Ent}(\pi_\varepsilon|\mu\otimes\mu) = J_\varepsilon(\pi_\varepsilon) \le J_\varepsilon(\pi) = J(\pi) + \varepsilon\mathrm{Ent}(\pi|\mu\otimes\mu)$ and $J(\pi) \le J(\pi_\varepsilon)$, we get

$$\varepsilon\mathrm{Ent}(\pi_\varepsilon|\mu\otimes\nu) \le J(\pi_\varepsilon) + \varepsilon\mathrm{Ent}(\pi_\varepsilon|\mu\otimes\nu) - J(\pi) \le J(\pi) + \varepsilon\mathrm{Ent}(\pi|\mu\otimes\nu) - J(\pi) = \varepsilon\mathrm{Ent}(\pi|\mu\otimes\nu).$$

By lower semi-continuity of $\mathrm{Ent}(\cdot|\mu\otimes\nu)$, we get $\mathrm{Ent}(\pi^*|\mu\otimes\nu) \le \mathrm{Ent}(\pi|\mu\otimes\nu)$. Hence, π^* minimizes $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ among $\mathfrak{C}(\mu,\nu)$. By the strict convexity of $\mathrm{Ent}(\pi|\mu\otimes\nu)$ and Condition (24), we know that $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ has a unique minimizer among optimal transport plans. By Theorem 3.6 this minimizer is $\mathcal{K}(\mu,\nu)$. Therefore $\pi^* = \mathcal{K}(\mu,\nu)$. This establishes the convergence result. □

Observe that, when μ and ν have finite support, Condition (24) is satisfied, so that $\mathcal{K}(\mu,\nu) = \lim_{\varepsilon\to 0^+}\pi_\varepsilon$. In case of atomic measure (not necessarily with finite support), Condition (24) may be satisfied or not. For instance, if for some atomic measure θ, we have $\mu = \nu = \theta$, then $\mathfrak{C}(\mu,\nu) = \mathcal{O}(\mu,\nu) = \{(\mathrm{id},\mathrm{id})_\#\theta\}$, and $\mathrm{Ent}((\mathrm{id},\mathrm{id})_\#\theta|\mu\otimes\nu) = -\sum_{x\in\mathrm{spt}(\theta)}\theta(x)\log(\theta(x))$. Thus, for $\theta = \sum_{n\ge 2} n^{-2}\delta_n$ Condition (24) is satisfied, whereas for $\theta = \sum_{n\ge 2}(n\log(n)^2)^{-1}\delta_n$ Condition (24) is not satisfied.

3.2.2 The Di Marino–Louet case

In [10, Theorem 4.1], Di Marino and Louet proved the following convergence result of $(\pi_\varepsilon)_{\varepsilon>0}$.

Theorem 3.9. *Assume $\mu, \nu \in \mathcal{P}(\mathbb{R})$ satisfy the following conditions:*

1. *The measures μ, ν have compact support and satisfy $\mathrm{Ent}(\mu|\mathcal{L}^1) < +\infty$ and $\mathrm{Ent}(\nu|\mathcal{L}^1) < +\infty$;*
2. *The set $\mathrm{spt}(\mu) \cap \{F_\mu^+ = F_\nu^+\}$ has a Lebesgue-negligible boundary and the family $(]a_i, b_i[)_{i\in\mathcal{I}}$ of connected components of its interior satisfies:*

$$-\sum_{i\in\mathcal{I}} \int_{a_i}^{b_i} \log(\min(x-a_i, b_i-x))\,\mathrm{d}\mu(x) < +\infty;$$

3. $\min_{\pi\in\mathcal{O}(\mu,\nu)} \mathrm{Ent}\left(\pi\llcorner_{\{F_\mu^+=F_\nu^+\}^c\times\{F_\mu^+=F_\nu^+\}^c}|\mu\otimes\nu\right) < +\infty.$

Then $\lim_{\varepsilon\to+\infty} \pi_\varepsilon = \mathcal{K}(\mu,\nu)$.

Observe that the first assumption implies that μ and ν are atomless. We briefly give the idea of the proof of this result to highlight the difference with our own convergence result (under different assumptions on μ and ν, see Theorem 3.24). A classical method to establish the convergence of $(\pi_\varepsilon)_{\varepsilon>0}$ is to prove that $(J_\varepsilon)_{\varepsilon>0}$ converges to J in the sense of Gamma-convergence, which requires a technique called block approximation. [12] Using a classical result on Gamma-convergence, we obtain that cluster points of $(\pi_\varepsilon)_{\varepsilon>0}$ converge to minimizers of J, that is toward optimal transport plans. When $\mathcal{O}(\mu,\nu)$ is not a singleton, this is not sufficient to prove that π_ε converges. In this case, one can prove that the functional $(H_\varepsilon)_{\varepsilon>0}$ defined by $H_\varepsilon : \pi \in \mathrm{Marg}(\mu,\nu) \to \frac{1}{\varepsilon}(J(\pi) - W_1(\mu,\nu)) + \mathrm{Ent}(\pi|\mu\otimes\nu)$ Γ-converges to

$$H : \pi \in \mathrm{Marg}(\mu,\nu) \mapsto \begin{cases} \mathrm{Ent}(\pi|\mu\otimes\nu) & \text{if } \pi \in \mathcal{O}(\mu,\nu) \\ +\infty & \text{otherwise} \end{cases}.$$

If there exists $\pi \in \mathcal{O}(\mu,\nu)$ such that $\mathrm{Ent}(\pi|\mu\otimes\nu)$ if finite, this Γ-convergence results proves that $(\pi_\varepsilon)_{\varepsilon>0}$ converges toward the unique minimizer of $\mathrm{Ent}(\cdot|\mu\otimes\nu)$ among $\mathcal{O}(\mu,\nu)$. In the L^1 case, as $H \equiv +\infty$ in general, this Γ-convergence result does not provide any information on the convergence of π_ε. The idea of Di Marino and Louet is to consider the refined family of functional $(F_\varepsilon)_{\varepsilon>0}$ defined as $F_\varepsilon = H_\varepsilon + \log(2\varepsilon)\mu(\mathrm{D})$, and prove that this family Γ-converges toward the functional F defined as

$$F : \pi \in \mathrm{Marg}(\mu,\nu) \mapsto \begin{cases} \mathrm{Ent}\left(\pi\llcorner_{\{F_\mu^+=F_\nu^+\}^c\times\{F_\mu^+=F_\nu^+\}^c}|\mu\otimes\nu\right) + \mathrm{Ent}(\mu\llcorner_{\mathrm{D}}|\mathcal{L}^1) & \text{if } \pi \in \mathcal{O}(\mu,\nu) \\ +\infty & \text{otherwise} \end{cases}.$$

The last hypothesis of Theorem 3.9 is crucial, as it allows to go from the convergence of $(F_\varepsilon)_{\varepsilon>0}$ toward F to the convergence of $(\pi_\varepsilon)_{\varepsilon>0}$. Next, they establish the existence of a strongly multiplicative transport plan on each component [10, Proposition 5.1], which corresponds to Theorem 2.50 for atomless measure with compact support. They finally identify $\mathcal{K}(\mu,\nu)$ as the minimizer of F by applying Proposition 3.5.

[12] For more information on Gamma-convergence, we e.g. refer to [7]. For more details on block approximation, we refer the reader to [8].

In their article Di Marino and Louet provide a necessary and sufficient condition to the third assumption [10, Theorem 1.1], which is satisfied if and only if $-\int_{\{F_\mu^+\neq F_\nu^+\}}\log(|F_\mu^+(x)-F_\nu^+(x)|)\,\mathrm{d}\mu(x)$ is finite. This condition is not always satisfied, for instance when $\mu(\mathrm{d}x)=\frac{1}{x\log(x)^2}\mathbb{1}_{x\geq 2}\,\mathcal{L}^1(\mathrm{d}x)$ and ν is defined by the relation $F_\nu(x)=F_\nu(x)-\frac{1}{x}\mathbb{1}_{x\geq 2}$.

3.3 Properties of cluster points of $(\pi_\varepsilon)_{\varepsilon>0}$

In this section, we show that restriction of any cluster point of $(\pi_\varepsilon)_{\varepsilon>0}$ to a product set included in F or $\tilde{\mathrm{F}}$ is a product measure. This property, referred to as (large) weak multiplicativity, is a symmetrized version of the notion of weak F-multiplicativity introduced by Kellerer [15, Definition 5.1]. First, let $\mathcal{C}$ be the class $\left\{B_1\times B_2\ ;\ (B_1,B_2)\in\mathcal{B}(\mathbb{R})^2, B_1\times B_2\subset \mathrm{F} \text{ or } B_1\times B_2\subset\tilde{\mathrm{F}}\right\}$ of product sets included in F or $\tilde{\mathrm{F}}$.

Definition 3.10 (Large weak multiplicativity)**.** A measure $\pi\in\mathcal{M}_+(\mathbb{R}^2)$ is said to be (largely) weakly multiplicative if, for all $A\in\mathcal{C}$, $\pi\llcorner_A$ is a product measure.

Clearly, strongly multiplicative measure (see Definition 2.49) are weakly multiplicative. The converse is not true, as $\pi=\delta_{(0,0)}+\delta_{(1,1)}$ is weakly multiplicative without being strongly multiplicative. We begin by showing that weak multiplicativity is preserved under weak convergence. Then, we use a structure result on the measures π_ε to prove that cluster points of π_ε at 0^+ are weakly multiplicative.

3.3.1 Stability of weak multiplicativity

The following characterization is straightforward and left to the reader.

Lemma 3.11. *A measure $\pi\in\mathcal{M}_+(\mathbb{R}^2)$ is a product measure if and only if $\pi(\mathbb{R}^2)\pi=(p_{1\#}\pi)\otimes(p_{2\#}\pi)$.*

The following Lemma provides a useful characterization of weak multiplicativity. For every $(a,b)\in\mathbb{R}^2$ such that $b<a$, we shall use the convention $[a,b]=\emptyset$.

Proposition 3.12. *Consider $\pi\in\mathrm{Marg}(\mu,\nu)$ and let $\mathcal{C}'$ be the class of subsets of $\mathbb{R}^2$ defined by*

$$\mathcal{C}'=\{]-\infty,t]\times[t,+\infty[\ ;\ t\in\mathbb{R}\}\cup\{[t,+\infty[\times]-\infty,t]\ ;\ t\in\mathbb{R}\}\,.$$

The following statements are equivalent:

1. *π is weakly multiplicative;*
2. *The restriction of π to any element of $\mathcal{C}'$ is a product measure;*
3. *For all $(x,y,t)\in\mathbb{R}^3$:*

$$\pi(]-\infty,t]\times[t,+\infty[)\pi(]-\infty,t\wedge x]\times[t,y])=\pi(]-\infty,t\wedge x]\times[t,+\infty[)\pi(]-\infty,t]\times[t,y])\quad(25)$$

and

$$\pi([t,+\infty[\times]-\infty,t])\pi([t,x]\times]-\infty,t\wedge y])=\pi([t,x]\times]-\infty,t])\pi([t,+\infty[\times]-\infty,t\wedge y]).\quad(26)$$

Proof. The equivalence between the first two statements was established by Kellerer in [15, Theorem 5.2]. We now prove the equivalence between the second and the third statements. For every $x \in \mathbb{R}$, define $I_x =]-\infty, x]$, and let us fix $t \in \mathbb{R}$. According to Proposition 3.11, $\pi_{]-\infty,t]\times[t,+\infty[}$ is a product measure if and only if

$$\pi_{]-\infty,t]\times[t,+\infty[}(\mathbb{R}^2)\pi_{]-\infty,t]\times[t,+\infty[} = (p_{1\#}\pi_{]-\infty,t]\times[t,+\infty[}) \otimes (p_{2\#}\pi_{]-\infty,t]\times[t,+\infty[}),$$

that is, for all $(x, y) \in \mathbb{R}^2$,

$$\pi_{]-\infty,t]\times[t,+\infty[}(\mathbb{R}^2)\pi_{]-\infty,t]\times[t,+\infty[}(I_x \times I_y) = (p_{1\#}\pi_{]-\infty,t]\times[t,+\infty[}) \otimes (p_{2\#}\pi_{]-\infty,t]\times[t,+\infty[})(I_x \times I_y),$$

which means Equation (25) holds. Similarly, $\pi_{[t,+\infty[\times]-\infty,t]}$ is a product measure if and only if, for all $(x, y) \in \mathbb{R}^2$, Equation (26) is satisfied. The equivalence between the second and the third point is now straightforward. □

To prove that weak multiplicativity is stable with respect to weak convergence, we shall prove that the third formulation is stable with respect to weak convergence. This is achieved by applying a variant of the Portmanteau theorem adapted to transport plans. For a proof, we refer to [6, Proposition/Notation 3.7].

Lemma 3.13. *Consider $(\pi_n)_{n\geq 1} \in \mathrm{Marg}(\mu,\nu)^{\mathbb{N}^*}$ and $\pi \in \mathrm{Marg}(\mu,\nu)$. The following conditions are equivalent:*

1. *$(\pi_n)_{n\geq 1}$ weakly converges to π;*
2. *For every pair (I_1, I_2) of intervals, $\pi_n(I_1 \times I_2) \xrightarrow[n\to+\infty]{} \pi(I_1 \times I_2)$.*

We can now prove that the set of weakly multiplicative transport plans from μ to ν is closed under weak convergence.

Proposition 3.14. *Consider $(\pi_n)_{n\geq 1} \in \mathrm{Marg}(\mu,\nu)^{\mathbb{N}^*}$ and $\pi \in \mathrm{Marg}(\mu,\nu)$. If the measures $(\pi_n)_{n\geq 1}$ are weakly multiplicative and $(\pi_n)_{n\geq 1}$ converges weakly to π, then π is weakly multiplicative.*

Proof. According to Proposition 3.12, for all $(x, y, t, n) \in \mathbb{R}^3 \times \mathbb{N}^*$,

$$\pi_n(]-\infty,t]\times[t,+\infty[)\pi_n(]-\infty,t\wedge x]\times[t,y]) = \pi_n(]-\infty,t\wedge x]\times[t,+\infty[)\pi_n(]-\infty,t]\times[t,y])$$

and

$$\pi_n([t,+\infty[\times]-\infty,t])\pi_n([t,x]\times]-\infty,t\wedge y]) = \pi_n([t,x]\times]-\infty,t])\pi_n([t,+\infty[\times]-\infty,t\wedge y]).$$

By Lemma 3.13, we get

$$\pi(]-\infty,t\wedge x]\times[t,y])\pi(]-\infty,t]\times[t,+\infty[) = \pi(]-\infty,t\wedge x]\times[t,+\infty[)\pi(]-\infty,t]\times[t,y])$$

and

$$\pi([t,+\infty[\times]-\infty,t])\pi([t,x]\times]-\infty,t\wedge y]) = \pi([t,x]\times]-\infty,t])\pi([t,+\infty[\times]-\infty,t\wedge y]).$$

From Proposition 3.12, it follows that π is weakly multiplicative. □

3.3.2 Weak multiplicativity of cluster points of $(\pi_\varepsilon)_{\varepsilon>0}$

The following Lemma provides us with information on the global form of ε-cyclically invariant transport plans. Its proof is relatively straightforward and can be found in [20, Lemma 2.7] by taking $R(dx, dy) = \exp(-|y-x|/\varepsilon)\mu(dx)\nu(dy)$.

Lemma 3.15. *Consider $\pi \in \mathrm{Marg}(\mu,\nu)$ and $\varepsilon > 0$. The following conditions are equivalent:*

1. *π is ε-cyclically monotone;*
2. *There exists two measurable maps $\phi_\varepsilon : \mathbb{R} \to \mathbb{R}$ and $\psi_\varepsilon : \mathbb{R} \to \mathbb{R}$ such that*

$$\frac{\mathrm{d}\pi}{\mathrm{d}\mu \otimes \nu}(x,y) = \exp\left(\frac{\phi_\varepsilon(x) + \psi_\varepsilon(y) - |y-x|}{\varepsilon}\right) \qquad (\mu \otimes \nu)(dx, dy) - a.s. \ .$$

Theorem 3.16. *Consider $(\mu,\nu) \in \mathcal{M}_+^2$. If π^* is a cluster point of $(\pi_\varepsilon)_{\varepsilon>0}$, then π^* is a weakly multiplicative measure. From Proposition 3.7, it follows that cluster points of $(\pi_\varepsilon)_{\varepsilon>0}$ are weakly multiplicative elements of $\mathfrak{C}(\mu,\nu)$.*

Proof. Consider $t \in \mathbb{R}$ and $\varepsilon > 0$. By Lemma 3.15, there exists $\phi_\varepsilon : \mathbb{R} \to \mathbb{R}$ and $\psi_\varepsilon : \mathbb{R} \to \mathbb{R}$ measurable such that

$$\pi_\varepsilon(dx, dy) = \exp\left(\frac{\phi_\varepsilon(x) + \psi_\varepsilon(y) - |y-x|}{\varepsilon}\right)\mu(dx)\nu(dy).$$

Thus, for every $t \in \mathbb{R}$,

$$\mathbb{1}_{]-\infty,t]\times[t,+\infty[}\pi_\varepsilon = (f_1^t\mu) \otimes (f_2^t\nu)$$

and

$$\mathbb{1}_{[t,+\infty[\times]-\infty,t]}\pi_\varepsilon = (f_3^t\mu) \otimes (f_4^t\nu),$$

where the maps $f_1^t, f_2^t, f_3^t, f_4^t : \mathbb{R} \to \mathbb{R}$ are defined by

$$\begin{aligned}
f_1^t(x) &= \mathbb{1}_{]-\infty,t]}(x)\exp\left(\frac{\phi_\varepsilon(x) + x}{\varepsilon}\right)\\
f_2^t(y) &= \mathbb{1}_{[t,+\infty[}(y)\exp\left(\frac{\psi_\varepsilon(y) - y}{\varepsilon}\right)\\
f_3^t(x) &= \mathbb{1}_{[t,+\infty[}(x)\exp\left(\frac{\phi_\varepsilon(x) - x}{\varepsilon}\right)\\
f_4^t(y) &= \mathbb{1}_{]-\infty,t]}(y)\exp\left(\frac{\psi_\varepsilon(y) + y}{\varepsilon}\right).
\end{aligned}$$

From Lemma 3.12, it follows that π_ε is weakly multiplicative. By Proposition 3.14, every cluster point of $(\pi_\varepsilon)_{\varepsilon>0}$ is weakly multiplicative. ☐

3.4 Convergence of entropic minimizer when the marginal components are mutually singular

To handle convergence in presence of mutually singular measures, we introduce the notions of strict reinforced stochastic order, strict strong multiplicativity, and strict weak multiplicativity. The definition of this notions and their relations are analogues, for the strict half-planes $\mathrm{G} = (x, y) \in \mathbb{R}^2 : y > x$ and $\tilde{\mathrm{G}} = (x, y) \in \mathbb{R}^2 : y < x$, to (large) reinforced stochastic order (Definition 2.33), (large) strong multiplicativity (Definition 2.49), and (large) weak multiplicativity (Definition 3.10).

Definition 3.17 (Strict strong multiplicativity)**.** A measure $\pi \in \mathcal{M}_+(\mathbb{R}^2)$ is called strictly strongly multiplicative if there exists $\eta_1, \eta_2 \in \mathcal{M}_+^\sigma(\mathbb{R})$ such that $\pi = (\eta_1 \otimes \eta_2)_{\llcorner \mathrm{G}}$.

We now introduce the order $\leq_{\mathrm{G}}$, that will be related to strict strong multiplicativity by its associated Strassen theorem in the following.

Definition 3.18 (Strict reinforced stochastic order)**.** Consider $(\gamma_1, \gamma_2) \in \mathcal{M}_+(\mathbb{R})^2$ and define

$$T_*(\gamma_1, \gamma_2) = \left\{ t \in \mathbb{R} \ ; \ F_{\gamma_1}^-(t) > 0 \text{ and } F_{\gamma_2}^+(t) < \gamma_2(\mathbb{R}) \right\}.$$

We say that γ_1 is smaller than γ_2 in the reinforced strict stochastic order if $F_{\gamma_1}^- \geq F_{\gamma_2}^+$ and $T_*(\gamma_1, \gamma_2) \subset \{F_{\gamma_1}^- > F_{\gamma_2}^+\}$. In this case, we write $\gamma_1 \leq_{\mathrm{G}} \gamma_2$.[13]

Similarly to Theorem 2.50 for the order $\leq_{\mathrm{F}}$, the following theorem from Kellerer [15, Theorem 2.4] provides a Strassen-type result for $\leq_{\mathrm{G}}$.

Theorem 3.19 (Strassen-type theorem for $\leq_{\mathrm{G}}$)**.** *Consider* $(\gamma_1, \gamma_2) \in \mathcal{M}_+(\mathbb{R})^2$*. The relation* $\gamma_1 \leq_{\mathrm{G}} \gamma_2$ *is satisfied if and only if* $\mathrm{Marg}(\gamma_1, \gamma_2)$ *contains a stricly strongly multiplicative measure,* i.e.*, there exists* $\eta_1, \eta_2 \in \mathcal{M}_+^\sigma(\mathbb{R})$ *such that*

$$(\eta_1 \otimes \eta_2)_{\llcorner \mathrm{G}} \in \mathrm{Marg}(\gamma_1, \gamma_2). \tag{27}$$

In this case, $\mathrm{Marg}(\gamma_1, \gamma_2)$ *contains exactly one strictly strongly multiplicative measure.*

We denote by $\mathcal{K}_{\mathrm{G}}(\gamma_1, \gamma_2)$ the unique strongly multiplicative measure in $\mathrm{Marg}(\gamma_1, \gamma_2)$ when $\gamma_1 \leq_{\mathrm{G}} \gamma_2$. Similarly as for the definition of $\mathcal{C}$ (Definition 3.10), let $\mathcal{D}$ denote the class

$$\left\{ B_1 \times B_2 \ ; \ (B_1, B_2) \in \mathcal{B}(\mathbb{R})^2, B_1 \times B_2 \subset \mathrm{G} \text{ or } B_1 \times B_2 \subset \tilde{\mathrm{G}} \right\}$$

of product sets included in G or $\tilde{\mathrm{G}}$.

Definition 3.20 (Strict weak multiplicativity)**.** A measure $\pi \in \mathcal{M}_+(\mathbb{R}^2)$ is said to be strictly weakly multiplicative if, for all $A \in \mathcal{D}$, $\pi_{\llcorner A}$ is a product measure.

The following result states that for measures γ_1, γ_2 in the strict reinforced stochastic order, $\mathcal{K}_{\mathrm{G}}(\gamma_1, \gamma_2)$ is the only weakly multiplicative measure in $\mathrm{Marg}_{\mathrm{G}}(\gamma_1, \gamma_2)$. There is no similar result for the analogue large properties: when $\gamma_1 \leq_F \gamma_2$, in general, $\mathcal{K}_{\mathrm{F}}(\gamma_1, \gamma_2)$ is not the only weakly multiplicative element of $\mathrm{Marg}_{\mathrm{F}}(\gamma_1, \gamma_2)$. For the proof, we refer to [15, Proposition 4.3]

[13] This notation is motivated by the associated Strassen-type theorem.

Lemma 3.21. *Consider $(\gamma_1, \gamma_2) \in \mathcal{M}_+(\mathbb{R})^2$ such that $\gamma_1 \leq_{\mathrm{G}} \gamma_2$. If $\pi \in \mathrm{Marg}_{\mathrm{G}}(\gamma_1, \gamma_2)$ and π is weakly multiplicative, then $\pi = \mathcal{K}_{\mathrm{G}}(\gamma_1, \gamma_2)$.*

Proposition 3.22. *Assume $(\gamma_1, \gamma_2) \in \mathcal{M}_+(\mathbb{R})^2$ have no atoms in common and $\gamma_1 \leq_{\mathrm{F}} \gamma_2$. Then, $\gamma_1 \leq_{\mathrm{G}} \gamma_2$ and $\mathcal{K}_{\mathrm{F}}(\mu, \nu) = \mathcal{K}_{\mathrm{G}}(\mu, \nu)$.*

Proof. For all $t \in \mathbb{R}$, as γ_1 and γ_2 have no atom in common, $\gamma_1(t) = 0$ or $\gamma_2(t) = 0$. Therefore, $F^-_{\gamma_1}(t) - F^+_{\gamma_2}(t) \in \{F^+_{\gamma_1}(t) - F^+_{\gamma_2}(t), F^-_{\gamma_1}(t) - F^-_{\gamma_2}(t)\}$, and the inequality $F^-_{\gamma_1}(t) - F^+_{\gamma_2}(t) \leq \min\left(F^+_{\gamma_1}(t) - F^+_{\gamma_2}(t), F^-_{\gamma_1}(t) - F^-_{\gamma_2}(t)\right)$ is an equality. As $\gamma_1 \leq_{\mathrm{st}} \gamma_2$, we get $F^-_{\gamma_1}(t) - F^+_{\gamma_2}(t) = \min\left(F^+_{\gamma_1}(t) - F^+_{\gamma_2}(t), F^-_{\gamma_1}(t) - F^-_{\gamma_2}(t)\right) \geq 0$. Furthermore, the sets $T^*(\gamma_1, \gamma_2)$ (Definition 3.18), $T_+(\gamma_1, \gamma_2)$ and $T_-(\gamma_1, \gamma_2)$ (Definition 2.33) satisfy: $T_*(\gamma_1, \gamma_2) \subset T_+(\gamma_1, \gamma_2) \cap T_-(\gamma_1, \gamma_2) \subset \{F^+_{\gamma_1} > F^+_{\gamma_2}\} \cap \{F^-_{\gamma_1} > F^-_{\gamma_2}\}$. As $\{F^+_{\gamma_1} > F^+_{\gamma_2}\} \cap \{F^-_{\gamma_1} > F^-_{\gamma_2}\} = \{\min\left(F^+_{\gamma_1} - F^+_{\gamma_2}, F^-_{\gamma_1} - F^-_{\gamma_2}\right) > 0\}$, and, as we just saw, $F^-_{\gamma_1} - F^+_{\gamma_2} = \min\left(F^+_{\gamma_1} - F^+_{\gamma_2}, F^-_{\gamma_1} - F^-_{\gamma_2}\right)$, we get $T_*(\gamma_1, \gamma_2) \subset \{F^-_{\gamma_1} > F^+_{\gamma_2}(t)\}$. Therefore, $\gamma_1 \leq_{\mathrm{G}} \gamma_2$. Now, as $\gamma_1 \leq_F \gamma_2$, by Theorem 2.50, there exists two σ-finite measures $\nu_1 \ll \gamma_1$ and $\nu_2 \ll \gamma_2$ such that $\mathcal{K}_{\mathrm{F}}(\mu, \nu) = (\nu_1 \otimes \nu_2)_{\llcorner \mathrm{F}}$. As $\mathrm{Atom}(\nu_1) \cap \mathrm{Atom}(\nu_2) \subset \mathrm{Atom}(\gamma_1) \cap \mathrm{Atom}(\gamma_2) = \emptyset$ and $\mathrm{Atom}(\nu_1)$ is countable, we have $\nu_2(\mathrm{Atom}(\nu_1)) = 0$. Since $x \mapsto \nu_1(\{x\})$ is null outside $\mathrm{Atom}(\nu_1)$, we get

$$(\nu_1 \otimes \nu_2)(\mathrm{D}) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{1}_{x=y} \, \mathrm{d}\nu_1(x) \right) \, \mathrm{d}\nu_2(y) = \int_{\mathrm{Atom}(\nu_1)^c} \nu_1(y) \, \mathrm{d}\nu_2(y) = 0.$$

Thus, $\mathcal{K}_{\mathrm{F}}(\mu, \nu) = (\nu_1 \otimes \nu_2)_{\llcorner \mathrm{G}} + (\nu_1 \otimes \nu_2)_{\llcorner \mathrm{D}} = (\nu_1 \otimes \nu_2)_{\llcorner \mathrm{G}}$ belongs to $\mathrm{Marg}_{\mathrm{G}}(\gamma_1, \gamma_2)$ and is the restriction of a product measure to G. This establishes $\mathcal{K}_{\mathrm{F}}(\gamma_1, \gamma_2) = \mathcal{K}_{\mathrm{G}}(\gamma_1, \gamma_2)$. ☐

Before stating that $\mathcal{K}(\mu, \nu) = \lim_{\varepsilon \to 0^+} \pi_\varepsilon$ when the components of non-equal pairs of $\mathcal{D}_{\mathcal{K}}$ are singular, we prove that two measures $(\gamma_1, \gamma_2) \in \mathcal{M}_+^2$ forms a pair of singular measures if and only if there exists no transport plan from γ_1 to γ_2 that fixes mass. To prove this equivalence, we shall use that two measures are singular if and only if their common part is null. Let us recall the definition of common part of two measure. For every $(x, y) \in \mathbb{R}^2$, let $x \wedge y$ denote the minimum between x and y. Now, if we define $h_1 = \frac{d\gamma_1}{d(\gamma_1 + \gamma_2)}$ and $h_2 = \frac{d\gamma_2}{d(\gamma_1 + \gamma_2)}$, the common part $\gamma_1 \wedge \gamma_2$ between γ_1 and γ_2 is defined by $\gamma_1 \wedge \gamma_2 = (h_1 \wedge h_2) \cdot (\gamma_1 + \gamma_2)$.

Proposition 3.23. *For all $\gamma_1, \gamma_2 \in \mathcal{M}_+^2$, the following conditions are equivalent:*

1. *$\gamma_1 \wedge \gamma_2 = 0$;*
2. *(γ_1, γ_2) is a pair of singular measures;*
3. *For all $\pi \in \mathfrak{C}(\gamma_1, \gamma_2)$, $\pi(\mathrm{D}) = 0$.*

Proof. The equivalence between the two first points is classical and omitted. We just prove the implications (2) $\implies$ (3) and (3) $\implies$ (1), beginning with (2) $\implies$ (3). Consider E such that $\gamma_1(E) = \gamma_2(E^c) = 0$ and $\pi \in \mathfrak{C}(\gamma_1, \gamma_2)$. As $p_{1\#}\pi_{\llcorner \mathrm{D}} \leq \gamma_1$, $p_{1\#}\pi_{\llcorner \mathrm{D}} = p_{2\#}\pi_{\llcorner \mathrm{D}} \leq \gamma_2$ and $\mathrm{D} = [\mathrm{D} \cap (E \times \mathbb{R})] \uplus [\mathrm{D} \cap (E^c \cap \mathbb{R})]$, we have $0 \leq \pi(\mathrm{D}) = \left(p_{1\#}\pi_{\llcorner \mathrm{D}}\right)(E) + \left(p_{1\#}\pi_{\llcorner \mathrm{D}}\right)(E^c) \leq \gamma_1(E) + \gamma_2(E^c) = 0 + 0 = 0$. We now prove that (3) $\implies$ (1). To derive a contradiction, assume $\gamma_1 \wedge \gamma_2 \neq 0$. Then, consider $\pi_0 \in \mathfrak{C}(\gamma_1 - \gamma_1 \wedge \gamma_2, \gamma_2 - \gamma_1 \wedge \gamma_2)$, and define $\pi = \pi_0 + (\mathrm{id}, \mathrm{id})_\# \gamma_1 \wedge \gamma_2$. By construction $\pi \in \mathrm{Marg}(\gamma_1, \gamma_2)$ and, from Point 2 of Lemma 2.31, π is cyclically monotone. Hence, $\pi \in \mathfrak{C}(\gamma_1, \gamma_2)$, and $\pi(\mathrm{D}) \geq [(\mathrm{id}, \mathrm{id})_\# \gamma_1 \wedge \gamma_2](\mathrm{D}) = (\gamma_1 \wedge \gamma_2)(\mathbb{R}) > 0$. This is a contradiction, hence 3 $\implies$ 1. ☐

Recall the notation for $\{(\mu_k^+,\nu_k^+)\}_{k\in\mathcal{K}^+}\cup\{(\mu_k^-,\nu_k^-)\}_{k\in\mathcal{K}^-}\cup\{(\mu^=,\mu^=)\}$ introduced in Definition 2.22. We now show that under the hypothesis,

$$\begin{cases}\forall k\in\mathcal{K}^+, \mu_k^+\wedge\nu_k^+=0\\ \forall k\in\mathcal{K}^-, \mu_k^-\wedge\nu_k^-=0\end{cases}, \tag{28}$$

we have $\mathcal{K}(\mu,\nu)=\lim_{\varepsilon\to 0^+}\pi_\varepsilon$. Note that Condition (28) implies the following condition,

$$\begin{cases}\forall k\in\mathcal{K}^+, \mathrm{Atom}(\mu_k^+)\cap\mathrm{Atom}(\nu_k^+)=0\\ \forall k\in\mathcal{K}^-, \mathrm{Atom}(\mu_k^-)\cap\mathrm{Atom}(\nu_k^-)=0\end{cases}, \tag{29}$$

which will allow us to apply Proposition 3.22. The assumption $\mu\wedge\nu=0$ is sufficient condition for Assumption (29), but is not necessary. For instance, $\mu=\delta_1+\delta_3+\delta_4$ and $\nu=\delta_2+\delta_3+\delta_5$ satisfy Assumption (29), but we do not have $\mu\wedge\nu=0$.

Theorem 3.24. *Consider $(\mu,\nu)\in\mathcal{M}_+^2$. If Assumption (28) is satisfied, then*

$$\pi_\varepsilon \xrightarrow[\varepsilon\to 0^+]{} \mathcal{K}(\mu,\nu).$$

Proof. Consider a cluster point π of $(\pi_\varepsilon)_{\varepsilon>0}$. According to Theorem 3.16, π is weakly multiplicative and belongs to $\mathfrak{C}(\mu,\nu)$. We now define $\mathcal{F}=((\pi_k^+)_{k\in\mathcal{K}^+},(\pi_k^-)_{k\in\mathcal{K}^-},\pi^=)$ using the notation in Definition 2.26 and $\mathfrak{P}$ as in Theorem 2.36. From this theorem, it follows that $\mathcal{F}\in\mathfrak{P}$ and $\pi=\sum_{k\in\mathcal{K}^+}\pi_k^++\sum_{k\in\mathcal{K}^-}\pi_k^-+(\mathrm{id},\mathrm{id})_\#\mu^=$. Fix $k\in\mathcal{K}^+$. As Equation (29) is satisfied and $\mu_k^+\leq_{\mathrm{F}}\nu_k^+$, from Proposition 3.22, it follows that $\mu_k^+\leq_{\mathrm{G}}\nu_k^+$ and $\mathcal{K}_{\mathrm{F}}(\mu_k^+,\nu_k^+)=\mathcal{K}_{\mathrm{G}}(\mu_k^+,\nu_k^+)$. Moreover, as Equation (28) is satisfied and $\pi\in\mathfrak{C}(\mu_k^+,\nu_k^+)$, by Proposition 3.23, $\pi_k^+(\mathrm{D})=0$. As $\mathfrak{C}(\mu_k^+,\nu_k^+)=\mathrm{Marg}_{\mathrm{F}}(\mu_k^+,\nu_k^+)$, $\pi_k^+\in\mathrm{Marg}_{\mathrm{G}}(\mu_k^+,\nu_k^+)$. Now, consider $(B_1,B_2)\in\mathcal{C}$. As π is weakly multiplicative, there exists $\nu_1,\nu_2\in\mathcal{M}^+(\mathbb{R})$ such that $\pi\llcorner_{B_1\times B_2}=(\nu_1\otimes\nu_2)$. Thus, $\pi_k^+\llcorner_{B_1\times B_2}=(\mathbb{1}_{[a_k^+,b_k^+[}\cdot\nu_1)\otimes(\mathbb{1}_{]a_k^+,b_k^+]}\cdot\nu_2)$. Therefore π_k^+ is weakly multiplicative. As $\mu_k^+\leq_{\mathrm{G}}\nu_k^+$ and $\pi_k^+\in\mathrm{Marg}_{\mathrm{G}}(\mu_k^+,\nu_k^+)$, $\pi_k^+=\mathcal{K}_{\mathrm{G}}(\mu_k^+,\nu_k^+)=\mathcal{K}_{\mathrm{F}}(\mu_k^+,\nu_k^+)$. Similarly, for all $k\in\mathcal{K}^-$, $\pi_k^-=i_\#\mathcal{K}_{\mathrm{F}}(\nu_k^-,\mu_k^-)$. Hence $\pi=\sum_{k\in\mathcal{K}^+}\mathcal{K}_{\mathrm{F}}(\mu_k^+,\nu_k^+)+\sum_{k\in\mathcal{K}^-}i_\#\mathcal{K}_{\mathrm{F}}(\nu_k^-,\mu_k^-)+(\mathrm{id},\mathrm{id})_\#\mu^==\mathcal{K}(\mu,\nu)$. Therefore $(\pi_\varepsilon)_{\varepsilon>0}$ admits a unique cluster point π. As $\mathrm{Marg}(\mu,\nu)$ is compact, this establishes our result. □

A Refinement of the decomposition $\mathcal{D}_\mathcal{K}$

This appendix is devoted to a more thorough investigation of the optimality properties of decompositions. Recall that in Point 3 of Remark 2.48, we established that the measure $\mu^=$ does not provide the maximum information about fixed part by giving an example of a pair (μ,ν) such that the inequality

$$(\mathrm{id},\mathrm{id})_\#\mu^=(A)\leq\inf_{\pi\in\mathfrak{C}(\mu,\nu)}\pi(\mathrm{D}\cap A)$$

may be strict for some $A\in\mathcal{B}(\mathbb{R}^2)$. In the following, we refine our decomposition $\mathcal{D}_\mathcal{K}$ to provide a decomposition, denoted $\mathcal{D}_\mathcal{K}^*$ with a fixed part η satisfying

$$(\mathrm{id},\mathrm{id})_\#\eta(A)=\inf_{\pi\in\mathfrak{C}(\mu,\nu)}\pi(\mathrm{D}\cap A). \tag{30}$$

In this section we define $\mathcal{D}_\mathcal{K}^*$ and prove its associated decomposition result. The proof that its fixed part satisfies Equality (30) will be given later in the appendix (Lemma B.4). The following Lemma is the key ingredient of the refinement.

Lemma A.1. *Consider $(\gamma_1, \gamma_2) \in \mathcal{M}_+^2$ such that $\gamma_1 \le_{\mathrm{st}} \gamma_2$. The set $\{F_{\gamma_2}^+ > F_{\gamma_1}^-\}$ is countable, which ensures that $\lambda := \sum_{x \in \{F_{\gamma_2}^+ > F_{\gamma_1}^-\}} (F_{\gamma_2}^+(x) - F_{\gamma_1}^-(x))\delta_x$ is well defined.*

1. *The pair $(\gamma_1 - \lambda, \gamma_2 - \lambda)$ belongs to $\mathcal{M}_+(\mathbb{R})^2$ and $F_{\gamma_1-\lambda}^- \ge F_{\gamma_2-\lambda}^+$. Additionally, if $\gamma_1 \le_{\mathrm{F}} \gamma_2$, then $\gamma_1 - \lambda \le_{\mathrm{F}} \gamma_2 - \lambda$.*

2. *The map $\tilde{\theta} : \mathfrak{C}(\gamma_1 - \lambda, \gamma_2 - \lambda) \to \mathfrak{C}(\gamma_1, \gamma_2)$ defined by $\tilde{\theta}(\tilde{\pi}) = \tilde{\pi} + (\mathrm{id}, \mathrm{id})_\# \lambda$ is invertible and its inverse $\theta : \mathfrak{C}(\gamma_1, \gamma_2) \to \mathfrak{C}(\gamma_1 - \lambda, \gamma_2 - \lambda)$ is given by $\theta(\pi) = \pi - (\mathrm{id}, \mathrm{id})_\# \lambda$.*

Proof. If $x \in \{F_{\gamma_2}^+ > F_{\gamma_1}^-\}$, then $F_{\gamma_1}^+(x) \ge F_{\gamma_2}^+(x) > F_{\gamma_1}^-(x)$. Therefore $\{F_{\gamma_2}^+ > F_{\gamma_1}^-\}$ is a subset of the countable set $\mathrm{Atom}(\gamma_1) = \{x \in \mathbb{R} \ ; \ \gamma_1(x) > 0\}$ and is countable. We now prove our claims.

1. For all $x \in \mathbb{R}$, we have $\lambda(x) \le F_{\gamma_2}^+(x) - F_{\gamma_1}^-(x) \le F_{\gamma_1}^+(x) - F_{\gamma_1}^-(x) = \gamma_1(x)$ and $\lambda(x) \le F_{\gamma_2}^+(x) - F_{\gamma_1}^-(x) \le F_{\gamma_2}^+(x) - F_{\gamma_2}^-(x) = \gamma_2(x)$. Therefore, $(\gamma_1 - \lambda, \gamma_2 - \lambda) \in \mathcal{M}_+(\mathbb{R})^2$. We now establish that $F_{\gamma_1-\lambda}^- \ge F_{\gamma_2-\lambda}^+$, that is, $F_{\gamma_1}^- + \lambda(\cdot) \ge F_{\gamma_2}^+$. If $x \in \{F_{\gamma_1}^- \ge F_{\gamma_2}^+\}$, then $F_{\gamma_1}^-(x) + \lambda(x) \ge F_{\gamma_2}^+(x) + 0 = F_{\gamma_2}^+(x)$, while if $x \in \{F_{\gamma_1}^- < F_{\gamma_2}^+\}$, $F_{\gamma_1}^-(x) + \lambda(x) = F_{\gamma_1}^-(x) + (F_{\gamma_2}^+(x) - F_{\gamma_1}^-(x)) = F_{\gamma_2}^+(x) \ge F_{\gamma_2}^+(x)$. Hence, $F_{\gamma_1-\lambda}^- \ge F_{\gamma_2-\lambda}^+$. Assume now $\gamma_1 \le_{\mathrm{F}} \gamma_2$. As $F_{\gamma_1-\lambda}^+ - F_{\gamma_2-\lambda}^+ = F_{\gamma_1}^+ - F_{\gamma_2}^+$, we get $\{F_{\gamma_1}^+ > F_{\gamma_2}^+\} = \{F_{\gamma_1-\lambda}^+ > F_{\gamma_2-\lambda}^+\}$. Moreover, for all $t \in T_+(\gamma_1 - \lambda, \gamma_2 - \lambda)$, $F_{\gamma_1}^+(t) \ge F_{\gamma_1-\lambda}^+(t) > 0$ and $F_{\gamma_2}^+(t) \le F_{\gamma_2-\lambda}^+(t) + \lambda(\mathbb{R}) < (\gamma_2 - \lambda)(\mathbb{R}) + \lambda(\mathbb{R}) = \gamma_2(\mathbb{R})$. Therefore, $T_+(\gamma_1 - \lambda, \gamma_2 - \lambda) \subset T_+(\gamma_1, \gamma_2)$. As $\gamma_1 \le_{\mathrm{F}} \gamma_2$, we get $T_+(\gamma_1 - \lambda, \gamma_2 - \lambda) \subset T_+(\gamma_1, \gamma_2) \subset \{F_{\gamma_1^+} > F_{\gamma_2}^+\} = \{F_{\gamma_1-\lambda}^+ > F_{\gamma_2-\lambda}^+\}$. Similarly, $T_-(\gamma_1 - \lambda, \gamma_2 - \lambda) \subset \{F_{\gamma_1-\lambda}^- > F_{\gamma_2-\lambda}^-\}$, which proves the first point.

2. Since $\gamma_1 \le_{\mathrm{st}} \gamma_2$ and $\gamma_1 - \lambda \le_{\mathrm{st}} \gamma_2 - \lambda$, from Proposition 2.11, it follows that $\mathfrak{C}(\gamma_1 - \lambda, \gamma_2 - \lambda) = \mathrm{Marg}_{\mathrm{F}}(\gamma_1 - \lambda, \gamma_2 - \lambda)$ and $\mathfrak{C}(\gamma_1, \gamma_2) = \mathrm{Marg}_{\mathrm{F}}(\gamma_1, \gamma_2)$. Therefore, $\tilde{\theta}$ is indeed valued in $\mathfrak{C}(\gamma_1, \gamma_2)$. To prove that θ is indeed valued in $\mathfrak{C}(\gamma_1 - \lambda, \gamma_2 - \lambda)$, the main part is to prove that it is valued in $\mathcal{M}_+(\mathbb{R}^2)$. Consider $\pi \in \mathfrak{C}(\gamma_1, \gamma_2) = \mathrm{Marg}_{\mathrm{F}}(\gamma_1, \gamma_2)$. For all $x \in \mathbb{R}$, $\gamma_1(x) = \pi(\{x\} \times [x, +\infty[) = \pi(x, x) + \pi(\{x\} \times]x, +\infty[)$ and $\gamma_1(]x, +\infty[) = \pi(]x, +\infty[\times]x, +\infty[)$, which implies $\gamma_2(\mathbb{R}) - F_{\gamma_2}^+(x) = \gamma_2(]x, +\infty[) = \pi(\mathbb{R} \times]x, +\infty[) \ge \pi(\{x\} \times]x, +\infty[) + \pi(]x, +\infty[\times]x, +\infty[) = \gamma_1(x) - \pi(x, x) + \gamma_1(]x, +\infty[) = \gamma_1(\mathbb{R}) - F_{\gamma_1}^-(x) - \pi(x, x)$. Since $\gamma_1(\mathbb{R}) = \gamma_2(\mathbb{R})$, this inequality implies $\pi(x, x) \ge F_{\gamma_2}^+(x) - F_{\gamma_1}^-(x) \ge \lambda(x)$. Therefore, $\pi - (\mathrm{id}, \mathrm{id})_\# \lambda \in \mathcal{M}_+(\mathbb{R}^2)$. From Remark 2.7, it follows that $\pi - (\mathrm{id}, \mathrm{id})_\# \lambda \in \mathfrak{C}(\gamma_1 - \lambda, \gamma_2 - \lambda)$. The identities $\theta \circ \tilde{\theta} = \mathrm{id}_{\mathfrak{C}(\gamma_1 - \lambda, \gamma_2 - \lambda)}$ and $\tilde{\theta} \circ \theta = \mathrm{id}_{\mathfrak{C}(\gamma_1, \gamma_2)}$ follow directly, thereby establishing our point. ☐

We now define the components of our refined decomposition $\mathcal{D}_{\mathcal{K}}^*$.

Definition A.2 (Refined marginal components)**.** Consider $(\mu, \nu) \in \mathcal{M}_+^2$.

1. For all $k \in \mathcal{K}^+$, define $\eta_k^+ = \sum_{x \in]a_k^+, b_k^+[\cap \{F_\nu^+ > F_\mu^-\}} (F_\nu^+(x) - F_\mu^-(x))\delta_x$ and $\tilde{\mu}_k^+ = \mu_k^+ - \eta_k^+$. Set $\tilde{\mu}^+ = \sum_{k \in \mathcal{K}^+} \tilde{\mu}_k^+$.

2. For all $k \in \mathcal{K}^-$, define $\eta_k^- = \sum_{x \in]a_k^-, b_k^-[\cap \{F_\mu^+ > F_\nu^-\}} (F_\mu^+(x) - F_\nu^-(x))\delta_x$ and $\tilde{\mu}_k^- = \mu_k^- - \eta_k^-$. Set $\tilde{\mu}^- = \sum_{k \in \mathcal{K}^-} \tilde{\mu}_k^-$.

3. Define $\eta = \mu^= + \sum_{k \in \mathcal{K}^+} \eta_k^+ + \sum_{k \in \mathcal{K}^-} \eta_k^- = \nu^= + \sum_{k \in \mathcal{K}^+} \eta_k^+ + \sum_{k \in \mathcal{K}^-} \eta_k^-$

4. Define $\mathcal{D}_{\mathcal{K}}^* = \{(\mu_k^+, \nu_k^+)\}_{k \in \mathcal{K}^+} \cup \{(\tilde{\mu}_k^-, \tilde{\nu}_k^-)\}_{k \in \mathcal{K}^-} \cup \{(\eta, \eta)\}$.

Example A.3. In Example 2.23, we have $\tilde{\mu}_1^+ = \mu_1^+ - \frac{1}{2}\delta_1$, $\nu_1^+ = \nu_1^+ - \frac{1}{2}\delta_1$, $\tilde{\mu}_1^- = \mu_1^-$, $\nu_1^- = \nu_1^-$, $\tilde{\mu}_2^- = \mu_2^-$, $\nu_2^- = \nu_2^-$, $\eta = \mu^= + \frac{1}{2}\delta_1$.

Observe that the sum of the refined components reconstructs the original measure: $\sum_{k\in\mathcal{K}^+} \tilde{\mu}_k^+ + \sum_{k\in\mathcal{K}^-} \tilde{\mu}_k^+ + \eta = \sum_{k\in\mathcal{K}^+} \mu_k^+ - \eta_k^+ + \sum_{k\in\mathcal{K}^-} \mu_k^- - \eta_k^- + \mu^= + \sum_{k\in\mathcal{K}^+} \eta_k^+ + \sum_{k\in\mathcal{K}^-} \eta_k^- = \sum_{k\in\mathcal{K}^+} \mu_k^+ + \sum_{k\in\mathcal{K}^-} \mu_k^- + \mu^= = \mu$. Similarly, $\sum_{k\in\mathcal{K}^+} \tilde{\nu}_k^+ + \sum_{k\in\mathcal{K}^-} \tilde{\nu}_k^- + \eta = \nu$. While it is not obvious that our components are positive measures, this will be verified in the proof of Theorem A.4. This Theorem establishes an analogue of Theorem 2.36 for the refined decomposition $\mathcal{D}_{\mathcal{K}}^*$ by applying Lemma B.4 to every non-equal pair of $\mathcal{D}_{\mathcal{K}}$.

Theorem A.4 (Refined decomposition of $\mathfrak{C}(\mu,\nu)$)**.** *Consider $(\mu,\nu) \in \mathcal{M}_+^2$. We use the notation introduced in Definition A.2 and define $\tilde{\mathfrak{P}} = \prod_{k\in\mathcal{K}^+} \mathfrak{C}(\tilde{\mu}_k^+, \tilde{\nu}_k^+) \times \prod_{k\in\mathcal{K}^-} \mathfrak{C}(\tilde{\mu}_k^-, \tilde{\nu}_k^-) \times \mathfrak{C}(\eta,\eta)$.*

1. *For all $k \in \mathcal{K}^+$ (resp. $\mathcal{K}^-$), $\tilde{\mu}_k^+ \leq_{\mathrm{F}} \tilde{\nu}_k^+$ and $F^-_{\tilde{\mu}_k^+} \geq F^+_{\tilde{\nu}_k^+}$ (resp. $\tilde{\nu}_k^- \leq_{\mathrm{F}} \tilde{\mu}_k^-$ and $F^-_{\tilde{\nu}_k^-} \geq F^+_{\tilde{\mu}_k^-}$).*
2. *The map $\tilde{\varphi} : \prod_{k\in\mathcal{K}^+} \mathfrak{C}(\tilde{\mu}_k^+, \tilde{\nu}_k^+) \times \prod_{k\in\mathcal{K}^-} \mathfrak{C}(\tilde{\mu}_k^-, \tilde{\nu}_k^-) \times \mathfrak{C}(\eta,\eta) \to \mathfrak{C}(\mu,\nu)$ defined by*

$$\tilde{\varphi}((\tilde{\pi}_k^+)_{k\in\mathcal{K}^+}, (\tilde{\pi}_k^-)_{k\in\mathcal{K}^-}, \tilde{\pi}^=) = \sum_{k\in\mathcal{K}^+} \tilde{\pi}_k^+ + \sum_{k\in\mathcal{K}^-} \tilde{\pi}_k^- + \tilde{\pi}^=$$

is a bijection. Furthermore, its inverse $\tilde{\varphi}^{-1} : \mathfrak{C}(\mu,\nu) \to \prod_{k\in\mathcal{K}^+} \mathfrak{C}(\tilde{\mu}_k^+, \tilde{\nu}_k^+) \times \prod_{k\in\mathcal{K}^-} \mathfrak{C}(\tilde{\mu}_k^-, \tilde{\nu}_k^-) \times \mathfrak{C}(\eta,\eta)$ is given by

$$\tilde{\varphi}^{-1}(\pi) = \Big(\Big(\pi\llcorner_{A_k^+} - (\mathrm{id},\mathrm{id})_\#\eta_k^+\Big)_{k\in\mathcal{K}^+}, \Big(\pi\llcorner_{A_k^-} - (\mathrm{id},\mathrm{id})_\#\eta_k^-\Big)_{k\in\mathcal{K}^-}, (\mathrm{id},\mathrm{id})_\#\eta\Big). \qquad (31)$$

Proof. 1. Consider $k \in \mathcal{K}^+$ and let $\tilde{\eta}_k^+ = \sum_{x; F^+_{\nu_k^+}(x) > F^-_{\mu_k^+}(x)} \Big(F^+_{\nu_k^+}(x) - F^+_{\mu_k^+}(x)\Big)\delta_x$ denote the measure λ of Lemma A.1 with $(\gamma_1,\gamma_2) = (\mu_k^+, \nu_k^+)$. By Equations (9) and (10), $F^+_{\nu_k^+} - F^-_{\mu_k^+} = \mathbb{1}_{]a_k^+, b_k^+[}(F^+_\nu - F^-_\mu)$. Hence, $\{F^+_{\nu_k^+} > F^+_{\mu_k^-}\} =]a_k^+, b_k^+[\cap \{F^+_\nu > F^-_\mu\}$ and $\tilde{\eta}_k^+ = \eta_k^+$. Since $\tilde{\mu}_k^+ = \mu_k^+ - \eta_k^+ = \mu_k^+ - \tilde{\eta}_k^+$ and $\tilde{\nu}_k^+ = \nu_k^+ - \eta_k^+ = \nu_k^+ - \tilde{\eta}_k^+$, it follows from Point 1 of Lemma A.1 that $(\tilde{\mu}_k^+, \tilde{\nu}_k^+) \in \mathcal{M}_+(\mathbb{R})^2$ and $F^-_{\tilde{\mu}_k^+} \geq F^+_{\tilde{\nu}_k^+}$. From Point 1 of Theorem 2.36 and Point 1 of Lemma A.1, $\tilde{\mu}_k^+ \leq_{\mathrm{F}} \tilde{\nu}_k^+$. The proof is similar for $k \in \mathcal{K}^-$.

2. Define $\mathfrak{P}$ and φ as in the proof of Theorem 2.36. Let $\theta : \mathfrak{P} \to \tilde{\mathfrak{P}}$ denote the map defined by

$$\theta\left((\pi_k^+)_{k\in\mathcal{K}^+}, (\pi_k^-)_{k\in\mathcal{K}^-}, \pi^=\right) = \left((\pi_k^+ + (\mathrm{id},\mathrm{id})_\#\eta_k^+)_{k\in\mathcal{K}^+}, (\pi_k^- + (\mathrm{id},\mathrm{id})_\#\eta_k^-)_{k\in\mathcal{K}^-}, (\mathrm{id},\mathrm{id})_\#\mu^=\right)$$

and $\tilde{\theta} : \tilde{\mathfrak{P}} \to \mathfrak{P}$ the map defined by

$$\tilde{\theta}\left((\tilde{\pi}_k^+)_{k\in\mathcal{K}^+}, (\tilde{\pi}_k^-)_{k\in\mathcal{K}^-}, \tilde{\pi}^=\right) = \left((\tilde{\pi}_k^+ - (\mathrm{id},\mathrm{id})_\#\eta_k^+)_{k\in\mathcal{K}^+}, (\tilde{\pi}_k^- - (\mathrm{id},\mathrm{id})_\#\eta_k^-)_{k\in\mathcal{K}^-}, (\mathrm{id},\mathrm{id})_\#\eta\right).$$

From Proposition A.1, it follows that θ and $\tilde{\theta}$ are indeed respectively valued in $\tilde{\mathfrak{P}}$ and $\mathfrak{P}$. Furthermore $\theta \circ \tilde{\theta} = \mathrm{id}_{\tilde{\mathfrak{P}}}$ and $\tilde{\theta} \circ \theta = \mathrm{id}_{\mathfrak{P}}$. Therefore, θ is a bijection. As $\eta = \mu^= + \sum_{k\in\mathcal{K}^+} \eta_k^+ + \sum_{k\in\mathcal{K}^-} \eta_k -$, we have $\varphi = \tilde{\varphi} \circ \theta$. From Theorem 2.36, φ is a bijection. As $\tilde{\varphi}^{-1} = \theta \circ \varphi^{-1}$, from Equation (12), Equation (31) follows.

□

Theorem A.4 states that $\mathfrak{C}(\mu,\nu)$ is the direct sum of $\{\mathfrak{C}(\tilde{\mu}_k^+,\tilde{\nu}_k^+)\}_{k\in\mathcal{K}^+}\cup\{\mathfrak{C}(\tilde{\mu}_k^-,\tilde{\nu}_k^-)\}_{k\in\mathcal{K}^-}\cup\{\mathfrak{C}(\eta,\eta)\}$, which expresses as

$$\begin{aligned}\mathfrak{C}(\mu,\nu) &= \left(\bigoplus_{k\in\mathcal{K}^+}\mathfrak{C}(\tilde{\mu}_k^+,\tilde{\nu}_k^+)\right)\oplus\left(\bigoplus_{k\in\mathcal{K}^-}\mathfrak{C}(\tilde{\mu}_k^-,\tilde{\nu}_k^-)\right)\oplus\mathfrak{C}(\eta,\eta)\\ &= \left(\bigoplus_{k\in\mathcal{K}^+}\mathrm{Marg}_{\mathrm{F}}(\tilde{\mu}_k^+,\tilde{\nu}_k^+)\right)\oplus\left(\bigoplus_{k\in\mathcal{K}^-}\mathrm{Marg}_{\tilde{\mathrm{F}}}(\tilde{\mu}_k^-,\tilde{\nu}_k^-)\right)\oplus\{(\mathrm{id},\mathrm{id})_\#\eta\}.\end{aligned}\tag{32}$$

B Fixed part of $\mathfrak{C}(\mu,\nu)$ and minimal concentration squares.

B.1 Fixed part of $\mathfrak{C}(\mu,\nu)$

In this section, we show that Equality (30) holds for the fixed part of $\mathcal{D}^*_{\mathcal{K}}$. First, we need a result of Kellerer that gives a necessary and sufficient condition for the existence of a transport plan with mass strictly going forward. To this end, we introduce the notion of reduced measures (in the sense of Kellerer) [15, Definition 1.7].

Definition B.1 (Reduced measures)**.** A pair $(\gamma_1,\gamma_2)\in\mathcal{M}_+(\mathbb{R})^2$ such that $\gamma_1\leq_{\mathrm{st}}\gamma_2$ is said to be reduced if $\gamma_1(\{F^+_{\gamma_1}=F^+_{\gamma_2}\})=\gamma_2(\{F^-_{\gamma_1}=F^-_{\gamma_2}\})=0$.

Remark B.2. If $\gamma_1\leq_{\mathrm{F}}\gamma_2$, one may verify that (γ_1,γ_2) is reduced if and only if $\gamma_1(S_{\gamma_2})=\gamma_2(s_{\gamma_1})=0$.

The following result provides a necessary and sufficient condition for the existence of an optimal transport plan concentrated on $\mathrm{G}=\{(x,y)\in\mathbb{R}^2\ ;\ x<y\}$. For a proof, we refer to [15, Proposition 1.12].

Proposition B.3. *Consider $(\gamma_1,\gamma_2)\in\mathcal{M}_+^2$. Then, the following conditions are equivalent:*

1. *$F^-_{\gamma_1}\geq F^+_{\gamma_2}$ and (γ_1,γ_2) is reduced.*
2. $\mathrm{Marg}_{\mathrm{G}}(\gamma_1,\gamma_2)\neq\emptyset$.

Lemma B.4. *Consider $(\mu,\nu)\in\mathcal{M}_+^2$. Define $\{\tilde{\mu}_k^+\ ;\ k\in\mathcal{K}^+\}\cup\{\tilde{\mu}_k^-\ ;\ k\in\mathcal{K}^-\}\cup\{\eta\}$ as in Definition A.2, and let $\tilde{\mathfrak{P}}$ and φ be defined as in Theorem A.4. Now, define $\mathfrak{G}=\prod_{k\in\mathcal{K}^+}\mathrm{Marg}_{\mathrm{G}}(\tilde{\mu}_k^+,\tilde{\nu}_k^+)\times\prod_{k\in\mathcal{K}^-}\mathrm{Marg}_{\tilde{\mathrm{G}}}(\tilde{\mu}_k^-,\tilde{\nu}_k^-)\times\{(\mathrm{id},\mathrm{id})_\#\eta\}$ and $\mathfrak{C}(\mu,\nu)_{\llcorner\mathrm{D}}=\{\pi_{\llcorner\mathrm{D}}\ ;\ \pi\in\mathfrak{C}(\mu,\nu)\}$.*

1. *The set $\mathfrak{G}$ is non-empty, and, for all $\mathcal{G}\in\mathfrak{G}$, $\tilde{\varphi}(\mathcal{G})_{\llcorner\mathrm{D}}=(\mathrm{id},\mathrm{id})_\#\eta$.*
2. *The set $\mathfrak{C}(\mu,\nu)_{\llcorner\mathrm{D}}$ admits a minimum with respect to $\leq$, which is given by the measure $(\mathrm{id},\mathrm{id})_\#\eta$. In particular, for all $A\in\mathcal{B}(\mathbb{R}^2)$,*

$$(\mathrm{id},\mathrm{id})_\#\eta(A)=\inf_{\pi\in\mathfrak{C}(\mu,\nu)}\pi(\mathrm{D}\cap A).$$

Proof. 1. According to Point 1 of Proposition A.4, for all $k\in\mathcal{K}^+$, $F^-_{\tilde{\mu}_k^+}\geq F^+_{\tilde{\nu}_k^+}$. Moreover, since $\tilde{\nu}_k^+(a_k^+)=\tilde{\mu}_k^+(b_k^+)=0$, Remark B.2 ensures that $(\tilde{\mu}_k^+,\tilde{\nu}_k^+)$ is reduced. By Proposition B.3, for all

$k \in \mathcal{K}^+$, the set $\mathrm{Marg}_{\mathrm{G}}(\tilde{\mu}_k^+, \tilde{\nu}_k^+)$ is non-empty. Similarly, for all $k \in \mathcal{K}^-$, $\mathrm{Marg}_{\tilde{\mathrm{G}}}(\tilde{\mu}_k^-, \tilde{\nu}_k^-)$ in non-empty. Therefore, $\mathfrak{G}$ is non-empty. Moreover, for all $\mathcal{G} = ((\tilde{\pi}_k^+)_{k\in\mathcal{K}^+}, (\tilde{\pi}_k^-)_{k\in\mathcal{K}^-}, (\mathrm{id},\mathrm{id})_\#\eta) \in \mathfrak{G}$, we have $\tilde{\varphi}(\mathcal{G})_{\llcorner \mathrm{D}} = (\sum_{k\in\mathcal{K}^+} \tilde{\pi}_k^+ + \sum_{k\in\mathcal{K}^-} \tilde{\pi}_k^- + (\mathrm{id},\mathrm{id})_\#\eta)_{\llcorner \mathrm{D}} = \sum_{k\in\mathcal{K}^+} 0 + \sum_{k\in\mathcal{K}^-} 0 + (\mathrm{id},\mathrm{id})_\#\eta_{\llcorner \mathrm{D}} = (\mathrm{id},\mathrm{id})_\#\eta$, which establishes our point.

2. By Theorem A.4, $\tilde{\varphi}(\tilde{\mathfrak{P}}) = \mathfrak{C}(\mu,\nu)$. Hence $\tilde{\varphi}(\mathfrak{G}) \subset \tilde{\varphi}(\tilde{\mathfrak{P}}) = \mathfrak{C}(\mu,\nu)$, and from the previous point, it follows that $(\mathrm{id},\mathrm{id})_\#\eta \in \mathfrak{C}(\mu,\nu)_{\llcorner D}$. As for all $\mathcal{F} = ((\tilde{\pi}_k^+)_{k\in\mathcal{K}^+}, (\tilde{\pi}_k^-)_{k\in\mathcal{K}^-}, (\mathrm{id},\mathrm{id})_\#\eta) \in \tilde{\mathfrak{P}}$, we have $\tilde{\varphi}(\mathcal{F})_{\llcorner \mathrm{D}} = \sum_{k\in\mathcal{K}^+} (\tilde{\pi}_k^+)_{\llcorner \mathrm{D}} + \sum_{k\in\mathcal{K}^-} (\tilde{\pi}_k^-)_{\llcorner \mathrm{D}} + (\mathrm{id},\mathrm{id})_\#\eta_{\llcorner \mathrm{D}} \geq (\mathrm{id},\mathrm{id})_\#\eta$, the measure $(\mathrm{id},\mathrm{id})_\#\eta$ is a minimum of $\mathfrak{C}(\mu,\nu)_{\llcorner \mathrm{D}}$. The final equation follows directly.

□

Notation B.5. Given $(\gamma_1,\gamma_2) \in \mathcal{M}_+^2$, we denote by $H(\gamma_1,\gamma_2) \in \mathcal{M}_+(\mathbb{R}^2)$ the minimum of $\mathfrak{C}(\gamma_1,\gamma_2)_{\llcorner \mathrm{D}}$ with respect to $\leq$. Define $\eta(\gamma_1,\gamma_2) = p_{1\#}H(\gamma_1,\gamma_2)$.

Remark B.6. 1. By Lemma B.4, $\eta = 0$ if and only if there exists $\pi \in \mathfrak{C}(\mu,\nu)$ such that $\pi(\mathrm{D}) = 0$. Observe that this does not imply that $\pi(\mathrm{D}) = 0$ for every $\pi \in \mathfrak{C}(\mu,\nu)$. For instance, if $\mu = \delta_1 + \delta_2$, $\nu = \delta_2 + \delta_3$ and $\pi = \delta_{1,3} + \delta_{2,2}$, we have $\eta(\mu,\nu) = 0$, $\pi \in \mathfrak{C}(\mu,\nu)$ and $\pi(\mathrm{D}) > 0$.

2. By the previous point and Point 1 of Lemma B.4, for all $k \in \mathcal{K}^+$ (resp. $k \in \mathcal{K}^-$), $\eta(\tilde{\mu}_k^+, \tilde{\nu}_k^+) = 0$ (resp. $\eta(\tilde{\mu}_k^-, \tilde{\nu}_k^-) = 0$).

According to Lemma B.4, for η defined as in Definition A.2, we have $H(\mu,\nu) = (\mathrm{id},\mathrm{id})_\#\eta$ and $\eta(\mu,\nu) = \eta$. Thus, for all $B \in \mathcal{B}(\mathbb{R})$, $\eta(\mu,\nu)(B)$ is the minimal amount of mass of B fixed by any element in $\mathfrak{C}(\mu,\nu)$. As an application, the minimal amount $\eta(\mathbb{R})$ of mass fixed by a transport plan is given by the expression $\eta(\mathbb{R}) = \mu(\mathbb{R}) + \sum_{k\in\mathcal{K}^+} \eta_k^+(\mathbb{R}) + \sum_{k\in\mathcal{K}^-} \eta_k^-(\mathbb{R}) = \mu(\mathbb{R}) - \mu^+(\mathbb{R}) - \mu^-(\mathbb{R}) + \sum_{k\in\mathcal{K}^+} \eta_k^+(\mathbb{R}) + \sum_{k\in\mathcal{K}^-} \eta_k^-(\mathbb{R})$, that is,

$$\begin{aligned}\eta(\mathbb{R}) = \mu(\mathbb{R}) &- \sum_{k\in\mathcal{K}^+} \left(F_\mu^-(b_k^+) - F_\nu^+(a_k^+)\right) - \sum_{k\in\mathcal{K}^-} \left(F_\nu^-(b_k^-) - F_\mu^+(a_k^-)\right) \\ &+ \sum_{x\in\{F_\mu^+ > F_\nu^+ > F_\mu^- > F_\nu^-\}} \left(F_\nu^+(x) - F_\mu^-(x)\right) + \sum_{x\in\{F_\nu^+ > F_\mu^+ > F_\nu^- > F_\mu^-\}} \left(F_\mu^+(x) - F_\nu^-(x)\right). \end{aligned} \tag{33}$$

B.2 Minimal concentration squares

In this part we aim to characterize the regions where non-fixed parts of elements of $\mathfrak{C}(\mu,\nu)$ are concentrated, that is, where elements of $\mathfrak{C}(\mu - \eta(\mu,\nu), \nu - \eta(\mu,\nu))$ are concentrated. Define

$$\mathcal{SQ} = \{I \times I\ ;\ I \text{ is a closed interval with positive length}\}.$$

Definition B.7. Consider $(\mu,\nu) \in \mathcal{M}_+^2$. We use the notation $\mathfrak{OR}$ introduced in Definition 2.30.

1. We denote by $\mathfrak{OR}(\mu,\nu)$ all the classes $\mathcal{C}$ belonging to $\mathfrak{OR}$ such that
   - $\mathcal{C} \subset \mathcal{SQ}$
   - Every element of $\mathfrak{C}(\mu - \eta(\mu,\nu), \nu - \eta(\mu,\nu))$ is concentrated on $\cup_{C\in\mathcal{C}} C$.

2. For every classes $\mathcal{C}, \mathcal{C}'$ of subset of $\mathbb{R}^2$, we write $\mathcal{C} \preceq \mathcal{C}'$ if $\mathcal{C}$ is finer than $\mathcal{C}'$, i.e.,

$$\forall C \in \mathcal{C}, \exists C' \in \mathcal{C}', C \subset C'.$$

Remark B.8. As an exercise, the reader may verify that $(\mathfrak{OR}(\mu, \nu), \preceq)$ forms a partially ordered set.

In the following, for every set $A \subset \mathbb{R}^2$, we denote by $\mathrm{cl}(A)$ the closure of A. For every set $C = I \times I \in \mathcal{SQ}$, we define $C_* = C \setminus \{(\inf I, \inf I), (\sup I, \sup I)\}$.

Proposition B.9. *Define $\mathcal{C} = \left\{\mathrm{cl}(A_k^+) \ ; \ k \in \mathcal{K}^+\right\} \cup \left\{\mathrm{cl}(A_k^-) \ ; \ k \in \mathcal{K}^-\right\}$. Then $\mathcal{C} \in \mathfrak{OR}$, $\mathcal{C} \subset \mathcal{SQ}$, and every element of $\mathfrak{C}(\mu - \eta, \nu - \eta)$ is concentrated on $\cup_{C \in \mathcal{C}} C_*$. In particular, $\mathcal{C}$ belongs to $\mathfrak{OR}(\mu, \nu)$. Furthermore, $\mathcal{C}$ is the minimum of $(\mathfrak{OR}(\mu, \nu), \preceq)$.*

Proof. The fact that $\mathcal{C}$ belongs to $\mathfrak{OR}$ and the inclusion $\mathcal{C} \subset \mathcal{SQ}$ are straightforward. Moreover, by Theorem A.4, every element of $\mathfrak{C}(\mu - \eta, \nu - \eta)$ is concentrated on $\left(\bigcup_{k \in \mathcal{K}^+} A_k^+\right) \cup \left(\bigcup_{k \in \mathcal{K}^-} A_k^-\right) \subset \cup_{C \in \mathcal{C}} C_*$. Suppose, to derive a contradiction, that there exists $\mathcal{C}' \in \mathfrak{OR}(\mu, \nu)$ such that $\mathcal{C} \preceq \mathcal{C}'$ is not satisfied. Then, there exists $C \in \mathcal{C}$ such that, for all $C' \in \mathcal{C}'$, $C \not\subset C'$. Let (a, a) denote the lower-left corner of C and (b, b) its upper-right corner. Since for any $\pi \in \mathfrak{C}(\mu - \eta, \nu - \eta)$, $\pi(\cup_{C' \in \mathcal{C}'} C_* \cap C') = \pi(C_*) > 0$, there exists $C' \in \mathcal{C}'$ with lower-left corner (x, x) and upper-right corner (y, y) such that $C_* \cap C' \neq \emptyset$. This implies $x < b$ and $a < y$. If $x \leq a$ and $b \leq y$ happens simultaneously, then $C \subset C'$. Hence, either $a < x$ or $y < b$. Assume for instance $a < x$ and note that $\mathrm{C}(x) = (]-\infty, x[\times]x, +\infty[) \cup (]x, +\infty[\times]-\infty, x[) = \left(]-\infty, x]^2 \cup [x, +\infty[^2\right)^c$. As $\mathcal{C} \in \mathfrak{OR}$, for all $C'' \in \mathcal{C}'$, we have $C'' = C'$ or $C'' \leq_{\mathbb{R}^2} C'$ or $C' \leq_{\mathbb{R}^2} C''$. If $C'' \leq_{\mathbb{R}^2} C'$, then $C'' \subset]-\infty, x]^2 \subset \mathrm{C}(x)^c$, while if $C'' = C'$ or $C' \leq_{\mathbb{R}^2} C''$, then $C'' \subset [x, +\infty[^2 \subset \mathrm{C}(x)^c$. Thus, $\mathrm{C}(x) \subset \left(\bigcup_{C'' \in \mathcal{C}'} C''\right)^c$. Therefore, for all $\pi \in \mathfrak{C}(\mu, \nu)$, $\pi(\mathrm{C}(x)) = (\pi - (\mathrm{id}, \mathrm{id})_\# \eta)(\mathrm{C}(x)) + (\mathrm{id}, \mathrm{id})_\# \eta(\mathrm{C}(x)) = 0$. Hence, $x \in \mathcal{B}$, and by Proposition 2.52, x belongs to $E^=$. As $x \in]a, b[\subset E^+ \subset (E^=)^c$, this yields a contradiction. Similarly, $y < b$ leads to a contradiction. Therefore $\mathcal{C} \preceq \mathcal{C}'$. ☐

Definition B.10. For every $(\gamma_1, \gamma_2) \in \mathcal{M}_+^2$, denote by $\mathcal{C}(\gamma_1, \gamma_2)$ the minimum of $\mathfrak{OR}(\gamma_1, \gamma_2)$ with respect to $\preceq$. According to Proposition B.9 this minimum exists and every element of $\mathfrak{C}(\gamma_1 - \eta(\gamma_1, \gamma_2), \gamma_2 - \eta(\gamma_1, \gamma_2))$ is concentrated on $\bigcup_{C \in \mathcal{C}(\gamma_1, \gamma_2)} C_*$.

C The ordered set of admissible decompositions

In Subsection 2.3, we introduced a preliminary definition of admissible decomposition (Definition C.2): in this section, we present an alternative definition, which extends the previous class of admissible decompositions to a new one, denoted by $\mathcal{A}^*$. Next, we introduce a relation $\preceq_{\mathcal{A}^*}$ on $\mathcal{A}^*$ to compare the fineness of admissible decompositions (Definition C.7). Finally, we show that $\preceq_{\mathcal{A}^*}$ is an order on $\mathcal{A}^*$, and $\mathcal{D}^*_{\mathcal{K}}$ is the minimum element of $\mathcal{A}^*$ with respect to $\preceq_{\mathcal{A}^*}$ (Theorem C.15).

Notation C.1. Given a family $\mathcal{D} = ((\mu_i, \nu_i))_{i \in \mathcal{I}} \subset (\mathcal{M}_+^2)^{\mathcal{I}}$ of pairs of measures, let $\mathcal{I}^=$ denote the set $\{i \in \mathcal{I} \ ; \ \mu_i = \nu_i\}$. For all $i \in \mathcal{I}$, define $H_i = H(\mu_i, \nu_i)$ and $\eta_i = \eta(\mu_i, \nu_i)$ (see Definition B.5). For all $i \in \mathcal{I} \setminus \mathcal{I}^=$, define $\mathcal{C}_i = \mathcal{C}(\mu_i, \nu_i)$ (see Definition B.10), $A^i = \bigcup_{C \in \mathcal{C}_i} C$ and $A^i_* = \bigcup_{C \in \mathcal{C}_i} C_*$ (see the notation before Proposition B.9).

Note that, for every $\gamma \in \mathcal{M}_+(\mathbb{R})$, one has $\eta(\gamma, \gamma) = \gamma$. Therefore, $i \in \mathcal{I}^=$ if and only if $\mu_i = \nu_i = \eta_i$.

We now introduce the notion of admissible decomposition. In summary, an such decomposition must satisfy the following two requirements:

- It must induce a decomposition of $\mathfrak{C}(\mu,\nu)$, see Equation (34);
- For all $i \in \mathcal{I} \setminus \mathcal{I}^=$, the measures $\mu_i - \eta_i$ and $\nu_i - \eta_i$, $i \in \mathcal{I} \setminus \mathcal{I}^=$ must be supported on unions of intervals that pairwise, possibly overlapping only a (see Equation (37)).

Definition C.2 (Admissible decompositions)**.** We say that $\mathcal{D} = ((\mu_i,\nu_i))_{i\in\mathcal{I}} \subset \mathcal{M}_+^2$ is an admissible decomposition of (μ,ν) if the following two conditions hold:

- $$\forall \pi \in \mathfrak{C}(\mu,\nu), \exists (\pi_i)_{i\in\mathcal{I}} \in \prod_{i\in\mathcal{I}} \mathfrak{C}(\mu_i,\nu_i), \pi = \sum_{i\in\mathcal{I}} \pi_i; \tag{34}$$
- There exist index sets $(J_i)_{i\in\mathcal{I}\setminus\mathcal{I}^=}$ and corresponding families $((a_i^j, b_i^j))_{j\in J_i} \in (\mathbb{R}\times\mathbb{R})^{J_i}$ with $a_i^j < b_i^j$ such that, for all $i \in \mathcal{I} \setminus \mathcal{I}^=$, $\mu_i - \eta_i$ and $\nu_i - \eta_i$ are concentrated on $\bigcup_{j\in J_i} [a_i^j, b_i^j]$, and
$$\forall (i_1,i_2) \in (\mathcal{I}\setminus\mathcal{I}^=)^2, \forall (j_1,j_2) \in J_{i_1} \times J_{i_2} : (i_1,j_1) \neq (i_2,j_2) \implies (b_{i_2}^{j_2} \leq a_{i_1}^{j_1} \text{ or } b_{i_1}^{j_1} \leq a_{i_2}^{j_2}). \tag{35}$$

We denote by $\mathcal{A}^*$ the set of admissible decompositions.

Notation C.3. When considering multiple decompositions $\mathcal{D}_1$, $\mathcal{D}_2$, we write $\mathcal{D}_k = \{(\mu_i^k, \nu_i^k)\}_{i\in\mathcal{I}_k}$ instead of $\mathcal{D} = \{(\mu_i,\nu_i)\}_{i\in\mathcal{I}}$ and η_i^k instead of η_i.

Remark C.4.

1. Unlike in Definition 2.43, an admissible decomposition may include several pairs of the form (η,η), and a pair (η,η) can appear multiple times in the decomposition. Observe that Requirement (34) states that $\mathfrak{C}(\mu,\nu) \subset \sum_{i\in\mathcal{I}} \mathfrak{C}(\mu_i,\nu_i)$ (see Notation 1.1): this is weaker than the first Requirement of Definition 2.43. In fact, we will prove that both requirements are equivalent provided that Equation (35) holds (see Remark C.6, Point 1). Also note that if the family $(]\min(s_{\mu_i-\eta_i}, s_{\nu_i-\eta_i}), \max(S_{\mu_i-\eta_i}, S_{\nu_i-\eta_i})[)_{i\in\mathcal{I}\setminus\mathcal{I}^=}$ consists of disjoint sets, Requirement (35) is satisfied: take $J_i = \{1\}$, $a_i^1 = \min(s_{\mu_i-\eta_i}, s_{\nu_i-\eta_i})$ and $b_i = \max(S_{\mu_i-\eta_i}, S_{\nu_i-\eta_i})$. In particular, this shows that the second requirement in Definition 2.43 implies Requirement (37). Thus, we have $\mathcal{A}^* \subset \mathcal{A}$.
2. We now present two equivalent formulations of Requirement (35). The first, offered as an intuitive formulation, states as follows: Equation (35) holds if and only if there exists a family $(J_i)_{i\in\mathcal{I}\setminus\mathcal{I}^=}$ and families $(A_i^j)_{j\in J_i} \in \mathcal{SQ}^{J_i}$ of elements of $\mathcal{SQ}$ such that, for all $i \in \mathcal{I}\setminus\mathcal{I}^=$, elements of $\mathfrak{C}(\mu_i - \eta_i, \nu_i - \eta_i)$ are concentrated on $\bigcup_{j\in J_i} A_i^j$, and the following condition holds:
$$\forall (i_1,i_2) \in (\mathcal{I}\setminus\mathcal{I}^=)^2, \forall (j_1,j_2) \in J_{i_1} \times J_{i_2} : (i_1,j_1) \neq (i_2,j_2) \implies (A_{i_2}^{j_2} \leq_{\mathbb{R}^2} A_{i_1}^{j_1} \text{ or } A_{i_1}^{j_1} \leq_{\mathbb{R}^2} A_{i_2}^{j_2}). \tag{36}$$
Indeed, if Requirement (35) holds just take $A_i^j = [a_i^j, b_i^j]^2$, whereas if Condition (36) holds, just take $a_i^j = \inf(p_1(A_i^j))$ and $b_i^j = \sup(p_1(A_i^j))$. The second formulation will often be used in the remainder of this subsection: Equation (35) is satisfied if and only if $(\mathcal{C}_i)_{i\in\mathcal{I}\setminus\mathcal{I}^=}$ forms a family of disjoint classes such that
$$\biguplus_{i\in\mathcal{I}\setminus\mathcal{I}^=} \mathcal{C}_i \in \mathfrak{OR}. \tag{37}$$

To see that Condition (37) implies Condition (36), observe that, if $(A_i^j)_{j\in J_i}$ is an injective family satisfying $\mathcal{C}_i = \left\{A_i^j \ ; \ j \in J_i\right\}$, then the families $(A_i^j)_{j\in J_i}$ for $i \in \mathcal{I}$ satisfy Condition (36). The establish that Condition (36) implies Condition (37) observe that $(A_i^j)_{i\in J_i} \in \mathfrak{OR}(\mu_i, \nu_i)$ (see Definition B.7). The minimality property of $\mathcal{C}_i$ implies that $\mathcal{C}_i \preceq \{A_i^j\}_{j\in J_i}$ for every $i \in \mathcal{I} \setminus \mathcal{I}^=$. As Condition (36) holds, $\bigcup_{j\in J_i}\{A_i^j\}_{j\in J_i} \in \mathfrak{OR}$: Condition (37) then follows directly. This is a consequence of the fact that, for all $i \in \mathcal{I}$, $\mathcal{C}_i \preceq \left\{A_i^j \ ; \ j \in J_i\right\}$. Observe that Equation (37) implies the following:

$$\forall (i,j) \in (\mathcal{I} \setminus \mathcal{I}^=)^2, \forall (C, C') \in \mathcal{C}_i \times \mathcal{C}_j : i \neq j \implies C_* \cap C'_* = \emptyset. \tag{38}$$

This follows directly from the fact that, for any pair $(A, B) \in \mathcal{SQ}^2$, $A \leq_{\mathbb{R}^2} B$ implies $A_* \cap B_* = \emptyset$.

3. We already know that $\mathcal{D}_\mathcal{K}$ belongs to $\mathcal{A} \subset \mathcal{A}^*$. Similarly $\mathcal{D}^*_\mathcal{K} \in \mathcal{A}^*$: Requirement (34) follows from Theorem A.4, while Requirement (37) follows from the equalities $\mathcal{C}(\tilde{\mu}_k^+, \tilde{\nu}_k^+) = \mathcal{C}(\mu_k^+, \nu_k^+) = \{\mathrm{cl}(A_k^+)\}$ and $\mathcal{C}(\tilde{\mu}_k^-, \tilde{\nu}_k^-) = \mathcal{C}(\mu_k^-, \nu_k^-) = \{\mathrm{cl}(A_k^-)\}$. Therefore, we conclude $\mathcal{D}^*_\mathcal{K} \in \mathcal{A}^*$.

4. Consider $\mathcal{D} = ((\mu_i, \nu_i))_{i\in\mathcal{I}}$ admissible. From Requirement (34), we have $\mathfrak{C}(\mu, \nu) \subset \sum_{i\in\mathcal{I}} \mathfrak{C}(\mu_i, \nu_i)$: we now prove that $\sum_{i\in\mathcal{I}} \mathfrak{C}(\mu_i, \nu_i) \subset \mathfrak{C}(\mu, \nu)$. To establish it, consider a family $(\pi_i)_{i\in\mathcal{I}} \in \prod_{i\in\mathcal{I}} \mathfrak{C}(\mu_i, \nu_i)$ and define $\pi = \sum_{i\in\mathcal{I}} \pi_i$. Let $(\Gamma_i)_{i\in\mathcal{I}}$ denote a family of cyclically monotone set such that, for all $i \in \mathcal{I}\setminus\mathcal{I}^=$, π_i is concentrated on Γ_i. For all $i \in \mathcal{I}\setminus\mathcal{I}^=$, $\pi_i - H_i$ is concentrated on $\Gamma_i \cap \left(\bigcup_{C\in\mathcal{C}_i} C\right)$. Hence π is concentrated on $\Gamma := \mathrm{D} \cup \left(\bigcup_{i\in\mathcal{I}\setminus\mathcal{I}^=} \bigcup_{C\in\mathcal{C}_i} (\Gamma_i \cap C)\right)$. According to Equation (37), $\biguplus_{i\in\mathcal{I}\setminus\mathcal{I}^=} \mathcal{C}_i$ belongs to $\mathfrak{OR}$. Therefore, $(\Gamma_i \cap C)_{i\in\mathcal{I}, C\in\mathcal{C}_i}$ is an ordered family of cyclically monotone sets. According to Remark 2.32, Γ is cyclically monotone. As π clearly has marginals μ and ν, π belongs to $\mathfrak{C}(\mu, \nu)$, which establishes the inclusion $\sum_{i\in\mathcal{I}} \mathfrak{C}(\mu_i, \nu_i) \subset \mathfrak{C}(\mu, \nu)$. By Equation (34), we have $\mathfrak{C}(\mu, \nu) = \sum_{i\in\mathcal{I}} \mathfrak{C}(\mu_i, \nu_i)$.

5. Consider $\mathcal{D} = \{(\mu_i, \nu_i) \ ; \ i \in \mathcal{I}\}$ admissible. By Proposition B.9, for all $i \in \mathcal{I} \setminus \mathcal{I}^=$ and $\pi_i \in \mathfrak{C}(\mu_i, \nu_i)$, the measure $\pi_i - H_i$ is concentrated on A_*^i. According to Equation (38), $(A_*^i)_{i\in\mathcal{I}\setminus\mathcal{I}^=}$ is a family of disjoint subsets. Thus, for all $(\pi_i)_{i\in\mathcal{I}\setminus\mathcal{I}^=} \in \prod_{i\in\mathcal{I}\setminus\mathcal{I}^=} \mathfrak{C}(\mu_i, \nu_i)$, the family $(\pi_i - H_i)_{i\in\mathcal{I}\setminus\mathcal{I}^=}$ consists of mutually singular measures. Note, however, that $(\pi_i)_{i\in\mathcal{I}\setminus\mathcal{I}^=}$ is not generally composed of mutually singular measures.

Lemma C.5. *For all $\mathcal{D} = ((\mu_i, \nu_i))_{i\in\mathcal{I}} \in \mathcal{A}^*$, it holds that $H(\mu, \nu) = \sum_{i\in\mathcal{I}} H_i$ and $\eta(\mu, \nu) = \sum_{i\in\mathcal{I}} \eta_i$.*

Proof. Consider $\pi \in \mathfrak{C}(\mu, \nu)$. According to Equation (34), there exists $(\pi_i)_{i\in\mathcal{I}}$ such that $\pi = \sum_{i\in\mathcal{I}} \pi_i$. Then, $\pi\lfloor_\mathrm{D} = \sum_{i\in\mathcal{I}} (\pi_i - H_i)\lfloor_\mathrm{D} + \sum_{i\in\mathcal{I}} H_i \geq \sum_{i\in\mathcal{I}} H_i$. Since this inequality holds for all $\pi \in \mathfrak{C}(\mu, \nu)$, it follows that $H(\mu, \nu) \geq \sum_{i\in\mathcal{I}} H_i$. To obtain the reverse inequality, we apply Lemma B.4: for every $i \in \mathcal{I}$, there exists $\pi_i \in \mathfrak{C}(\mu_i, \nu_i)$ such that $\pi_i\lfloor_\mathrm{D} = H_i$. By Point 4 of Remark C.4, $\pi := \sum_{i\in\mathcal{I}} \pi_i$ belongs to $\mathfrak{C}(\mu, \nu)$. Therefore, $\sum_{i\in\mathcal{I}} H_i = \sum_{i\in\mathcal{I}} \pi_i\lfloor_\mathrm{D} = \pi\lfloor_\mathrm{D} \geq H(\mu, \nu)$. Finally, $H(\mu, \nu) = \sum_{i\in\mathcal{I}} H_i$. Applying the first marginal projection yields $\eta(\mu, \nu) = \sum_{i\in\mathcal{I}} \eta_i$. □

Remark C.6. Consider $\mathcal{D} = ((\mu_i, \nu_i))_{i\in\mathcal{I}} \in \mathcal{A}^*$.

1. We now prove that the equality $\mathfrak{C}(\mu, \nu) = \sum_{i\in\mathcal{I}} \mathfrak{C}(\mu_i, \nu)$ can be improved to

$$\mathfrak{C}(\mu, \nu) = \bigoplus_{i\in\mathcal{I}} \mathfrak{C}(\mu_i, \nu_i). \tag{39}$$

Assume $\pi \in \mathfrak{C}(\mu,\nu)$ and $(\pi_i)_{i\in\mathcal{I}} \in \prod_{i\in\mathcal{I}} \mathfrak{C}(\mu_i,\nu_i)$ satisfies $\pi = \sum_{i\in\mathcal{I}} \pi_i$. According to Lemma C.5, we have $\pi - H(\mu,\nu) = \sum_{i\in\mathcal{I}}(\pi_i - H_i) = \sum_{i\in\mathcal{I}\setminus\mathcal{I}^=}(\pi_i - H_i)$. For all $i \in \mathcal{I}\setminus\mathcal{I}^=$, restricting the previous equality to A_i^* and applying Point 5 of Remark C.4, we obtain

$$(\pi - H(\mu,\nu))\llcorner_{A_*^i} = \sum_{j\in\mathcal{I}\setminus\mathcal{I}^=} (\pi_j - H_j)\llcorner_{A_*^i} = \pi_i - H_i.$$

Hence, $\pi_i = (\pi - H(\mu,\nu))\llcorner_{A_*^i} + H_i$. If $i \in \mathcal{I}^=$, $\pi_i = H_i$. Thus, uniqueness holds, and the conclusion of Point 4 of Remark C.4 yields the identity $\mathfrak{C}(\mu,\nu) = \bigoplus_{i\in\mathcal{I}} \mathfrak{C}(\mu_i,\nu_i)$.

2. We have $\biguplus_{i\in\mathcal{I}\setminus\mathcal{I}^=} \mathcal{C}_i \in \mathfrak{DR}(\mu,\nu)$. Since for all $i \in \mathcal{I}\setminus\mathcal{I}^=$, $\mathcal{C}_i \subset \mathcal{SQ}$, it follows that $\biguplus_{i\in\mathcal{I}\setminus\mathcal{I}^=} \mathcal{C}_i \subset \mathcal{SQ}$. By Equation (37), $\biguplus_{i\in\mathcal{I}\setminus\mathcal{I}^=} \mathcal{C}_i \in \mathfrak{DR}$. Now, for all $\pi \in \mathfrak{C}(\mu,\nu)$, there exists $(\pi_i)_{i\in\mathcal{I}} \in \prod_{i\in\mathcal{I}} \mathfrak{C}(\mu_i,\nu_i)$ such that $\pi = \sum_{i\in\mathcal{I}} \pi_i$. Then $\pi - H(\mu,\nu) = \sum_{i\in\mathcal{I}}(\pi_i - H_i)$, and for all $j \in \mathcal{I}$, $\pi_j - H_j \in \mathfrak{C}(\mu_j - \eta_j, \nu_j - \eta_j)$ is concentrated on $\bigcup_{C\in\mathcal{C}_j} C \subset \bigcup_{C\in\bigcup_{i\in\mathcal{I}\setminus\mathcal{I}^=}\mathcal{C}_i} C$. Therefore, $\pi - H(\mu,\nu)$ is concentrated on $\bigcup_{C\in\biguplus_{i\in\mathcal{I}\setminus\mathcal{I}^=}\mathcal{C}_i} C$. Thus, we obtain $\bigcup_{i\in\mathcal{I}\setminus\mathcal{I}^=} \mathcal{C}_i \in \mathfrak{DR}(\mu,\nu)$.

We now define a relation $\preceq_{\mathcal{A}^*}$ on $\mathcal{A}^*$ to compare different admissible decompositions. This definition, formulated purely in terms of measures, admits a more intuitive counterpart at the level of cyclically monotone transport plans (see Proposition C.8)

Definition C.7 (Order on decompositions)**.** Let $\mathcal{D}_1 = ((\mu_i^1,\nu_i^1))_{i\in\mathcal{I}_1}$ and $\mathcal{D}_2 = ((\mu_i^2,\nu_i^2))_{i\in\mathcal{I}_2}$ be elements of $\mathcal{A}^*$. We say that $\mathcal{D}_1$ is a finer decomposition than $\mathcal{D}_2$ if the following two conditions hold:

1. There exists a disjoint covering $(J_{i_1})_{i_1\in\mathcal{I}_1^=}$ of $\mathcal{I}_2^=$ such that:

$$\forall i_1 \in \mathcal{I}_1^=, \eta_{i_1}^1 \geq \sum_{j_2\in J_{i_1}} \eta_{j_2}^2. \tag{40}$$

2. There exists a partition $(J_{i_2})_{i_2\in\mathcal{I}_2\setminus\mathcal{I}_2^=}$ of $\mathcal{I}_1\setminus\mathcal{I}_1^=$ such that:

$$\forall i_2 \in \mathcal{I}_2\setminus\mathcal{I}_2^=, \begin{cases} \mu_{i_2}^2 - \eta_{i_2}^2 = \sum_{j_1\in J_{i_2}} \mu_{j_1}^1 - \eta_{j_1}^1 \\ \nu_{i_2}^2 - \eta_{i_2}^2 = \sum_{j_1\in J_{i_2}} \nu_{j_1}^1 - \eta_{j_1}^1 \\ \eta_{i_2}^2 \geq \sum_{j_1\in J_{i_2}} \eta_{j_1}^1 \end{cases}. \tag{41}$$

In this case, we write $\mathcal{D}_1 \preceq_{\mathcal{A}^*} \mathcal{D}_2$.

The following proposition provides a counterpart of Definition C.7 formulated in terms of cyclically monotone transport plans: $\mathcal{D}_1 \preceq_{\mathcal{A}^*} \mathcal{D}_2$ if and only if each non-equal component $\mathfrak{C}(\mu_{i_2}^2,\nu_{i_2}^2)$ for $\mathcal{D}_2$ can be expressed as a sum of non-equal components of $\mathfrak{C}(\mu_{i_1}^1,\nu_{i_1}^1)$ plus a fixed part, while equal components $\mathfrak{C}(\mu_{i_1}^1,\nu_{i_1}^1)$ for $i_1 \in \mathcal{I}_1^=$ can be expressed as the sum of equal components $\mathfrak{C}(\mu_{i_2}^2,\nu_{i_2}^2)$ with $i_2 \in \mathcal{I}_2^=$ plus a fixed part.

Proposition C.8. *Let $\mathcal{D}_1 = \{(\mu_{i_1}^1,\nu_{i_1}^1)\}_{i_1\in\mathcal{I}_1}$ and $\mathcal{D}_2 = \{(\mu_{i_2}^2,\nu_{i_2}^2)\}_{i_2\in\mathcal{I}_2}$ be two admissible decompositions. Then $\mathcal{D}_1 \preceq_{\mathcal{A}^*} \mathcal{D}_2$ if and only if the following two conditions hold:*

1. *There exists a disjoint covering* $(J_{i_1})_{i_1 \in \mathcal{I}_1^=}$ *of* $\mathcal{I}_2^=$ *and a family of measures* $(\theta^1_{i_1})_{i_1 \in \mathcal{I}_1^=}$ *such that*

$$\forall i_1 \in \mathcal{I}_1^=, \mathfrak{C}(\mu^1_{i_1}, \nu^1_{i_1}) = \bigoplus_{i_2 \in J_{i_1}} \mathfrak{C}(\mu^2_{i_2}, \nu^2_{i_2}) \oplus \mathfrak{C}(\theta^1_{i_1}, \theta^1_{i_1}). \tag{42}$$

2. *There exists a partition* $(J_{i_2})_{i_2 \in \mathcal{I}_2 \setminus \mathcal{I}_2^=}$ *of* $\mathcal{I}_1 \setminus \mathcal{I}_1^=$*, and a family of measures* $(\theta^2_{i_2})_{i_2 \in \mathcal{I}_2 \setminus \mathcal{I}_2^=}$ *such that,*

$$\forall i_2 \in \mathcal{I}_2 \setminus \mathcal{I}_2^=, \mathfrak{C}(\mu^2_{i_2}, \nu^2_{i_2}) = \bigoplus_{i_1 \in J_{i_2}} \mathfrak{C}(\mu^1_{i_1}, \nu^1_{i_1}) \oplus \mathfrak{C}(\theta^2_{i_2}, \theta^2_{i_2}). \tag{43}$$

Proof. Assume that $\mathcal{D}_1 \preceq_{\mathcal{A}^*} \mathcal{D}_2$ and let $(J_{i_1})_{i_1 \in \mathcal{I}_1^=}$ and $(J_{i_2})_{i_2 \in \mathcal{I}_2 \setminus \mathcal{I}_2^=}$ be as in Definition C.7. First, note that Equation (42) is equivalent to $\eta^1_{i_1} = \sum_{i_2 \in J_{i_2}} \eta^2_{i_2} + \theta^1_{i_1}$. Therefore Equation (42) follows directly from Requirement (40). To prove Equation (43), define

$$E_{i_2} = \begin{cases} \mathfrak{C}(\eta^2_{i_2}, \eta^2_{i_2}) \oplus \left(\bigoplus_{i_1 \in J_{i_2}} \mathfrak{C}(\mu^1_{i_1} - \eta^1_{i_1}, \nu^1_{i_1} - \eta^1_{i_1}) \right) & \text{if } i_2 \in \mathcal{I}_2 \setminus \mathcal{I}_2^= \\ \mathfrak{C}(\mu^2_{i_2}, \nu^2_{i_2}) & \text{otherwise} \end{cases}. \tag{44}$$

By Lemma C.5, we obtain:

$$\begin{aligned} \mathfrak{C}(\mu, \nu) &= \left(\bigoplus_{i_1 \in \mathcal{I}_1} \mathfrak{C}(\mu^1_{i_1}, \nu^1_{i_1}) \right) - \left(\bigoplus_{i_1 \in \mathcal{I}_1} \mathfrak{C}(\eta^1_{i_1}, \eta^1_{i_1}) \right) + \left(\bigoplus_{i_2 \in \mathcal{I}_2} \mathfrak{C}(\eta^2_{i_2}, \eta^2_{i_2}) \right) \\ &= \left(\bigoplus_{i_1 \in \mathcal{I}_1 \setminus \mathcal{I}_1^=} \mathfrak{C}(\mu^1_{i_1} - \eta^1_{i_1}, \nu^1_{i_1} - \eta^1_{i_1}) \right) \oplus \left(\bigoplus_{i_2 \in \mathcal{I}_2} \mathfrak{C}(\eta^2_{i_2}, \eta^2_{i_2}) \right) \\ &= \bigoplus_{i_2 \in \mathcal{I}_2 \setminus \mathcal{I}_2^=} \left(\mathfrak{C}(\eta^2_{i_2}, \eta^2_{i_2}) \oplus \bigoplus_{i_1 \in J_{i_2}} \mathfrak{C}(\mu^1_{i_1} - \eta^1_{i_1}, \nu^1_{i_1} - \eta^1_{i_1}) \right) \oplus \bigoplus_{i_2 \in \mathcal{I}_2^=} \mathfrak{C}(\eta^2_{i_2}, \eta^2_{i_2}) \\ &= \bigoplus_{i_2 \in \mathcal{I}_2} E_{i_2}. \end{aligned}$$

From Equation (41), it follows that $E_{i_2} \subset \mathfrak{C}(\mu^2_{i_2}, \nu^2_{i_2})$. As $\mathfrak{C}(\mu, \nu) = \bigoplus_{i_2 \in \mathcal{I}_2} \mathfrak{C}(\mu^2_{i_2}, \nu^2_{i_2})$, it is straightforward that, for all $i_2 \in \mathcal{I}_2$, $E_{i_2} = \mathfrak{C}(\mu^2_{i_2}, \nu^2_{i_2})$. Define, for all $i_2 \in \mathcal{I}_2 \setminus \mathcal{I}_2^=$, $\theta_{i_2} = \eta^2_{i_2} - \sum_{i_1 \in J_{i_2}} \eta^1_{i_1}$. Equation (43) then follows from $\mathfrak{C}(\mu^2_{i_2}, \nu^2_{i_2}) = E_{i_2}$. The implication from Conditions (42) and (43) to $\mathcal{D}_1 \preceq_{\mathcal{A}^*} \mathcal{D}_2$ is straightforward and left to the reader. □

Remark C.9. 1. For any $\mathcal{D} = ((\mu_i, \nu_i))_{i \in \mathcal{I}}$, the measure $\sum_{i \in \mathcal{I}^=} \eta_i$ represents the information that $\mathcal{D}$ provides about the fixed part of elements of $\mathfrak{C}(\mu, \nu)$. Observe that by taking $(J_{i_1})_{i_1 \in \mathcal{I}_1^=}$ as in Equation (40), we obtain:

$$\sum_{i_1 \in \mathcal{I}_1^=} \eta^1_{i_1} \geq \sum_{i_1 \in \mathcal{I}_1^=} \sum_{i_2 \in J_{i_1}} \eta^2_{i_2} = \sum_{i_2 \in \mathcal{I}_2^=} \eta^2_{i_2}. \tag{45}$$

Requirement (40) is strictly stronger than the inequality (45), as it favours decompositions in which the fixed parts are grouped: (δ_0, δ_0) is a better decomposition than $(2^{-1}(\delta_0, \delta_0), 2^{-1}(\delta_0, \delta_0))$.

2. The last line of Equation (41) also requires that the fixed part unaccounted for by $\{(\mu^1_{i_1}, \nu^1_{i_1})\}_{i_1 \in J_{i_2}}$ is less than that of $\{(\mu^2_{i_2}, \nu^2_{i_2})\}$.

The following definition makes use of the notation A^+_k and A^-_k introduced in Definition 2.26, as well as A^i introduced in Notation C.1.

Definition C.10. Let $\mathcal{D} = ((\mu_i, \nu_i))_{i \in \mathcal{I}}$ be an admissible decomposition. For all $i \in \mathcal{I} \setminus \mathcal{I}^=$, define:

$$K^+_i = \left\{k \in \mathcal{K}^+ \ ; \ A^+_k \subset A^i\right\} \quad \text{and} \quad K^-_i = \left\{k \in \mathcal{K}^- \ ; \ A^-_k \subset A^i\right\}.$$

Remark C.11.
1. According to Equation (37), for all $(i, j) \in \mathcal{I} \setminus \mathcal{I}^=$, if $i \neq j$, then $A^i \cap A^j \subset \mathrm{D}$. Therefore, $(K^+_i)_{i \in \mathcal{I} \setminus \mathcal{I}^=}$ and $(K^-_i)_{i \in \mathcal{I} \setminus \mathcal{I}^=}$ form disjoint family of index sets.
2. As $\biguplus_{i \in \mathcal{I} \setminus \mathcal{I}^=} \mathcal{C}_i \in \mathfrak{D}\mathfrak{R}(\mu, \nu)$, we have $\mathcal{C}(\mu, \nu) \preceq \biguplus_{i \in \mathcal{I} \setminus \mathcal{I}^=} \mathcal{C}_i$. Thus, for all $k \in \mathcal{K}^+$, there exists $i \in \mathcal{I} \setminus \mathcal{I}^=$ and $C \in \mathcal{C}_i$ such that $\mathrm{cl}(A^+_k) \subset C$. Hence, $A^*_k \subset \mathrm{cl}(A^+_k) \subset C \subset A^i$, which yields $k \in K^+_i$. Therefore, $\mathcal{K}^+ \subset \bigcup_{i \in \mathcal{I} \setminus \mathcal{I}^=} K^+_i$. Similarly, $\mathcal{K}^- \subset \bigcup_{i \in \mathcal{I} \setminus \mathcal{I}^=} K^-_i$. By the previous point, we have

$$\begin{cases} \mathcal{K}^+ = \biguplus_{i \in \mathcal{I} \setminus \mathcal{I}^=} K^+_i \\ \mathcal{K}^- = \biguplus_{i \in \mathcal{I} \setminus \mathcal{I}^=} K^-_i \end{cases}. \tag{46}$$

Lemma C.12. *Consider $\mathcal{D} = ((\mu_i, \nu_i))_{i \in \mathcal{I}} \in \mathcal{A}^*$. Then, for all $i \in \mathcal{I}$:*

$$\begin{cases} \mu_i - \eta_i = \sum_{k \in K^+_i} \tilde{\mu}^+_k + \sum_{k \in K^-_i} \tilde{\mu}^-_k \\ \nu_i - \eta_i = \sum_{k \in K^+_i} \tilde{\nu}^+_k + \sum_{k \in K^-_i} \tilde{\nu}^-_k \end{cases}.$$

Proof. Consider $\pi \in \mathfrak{C}(\mu, \nu)$, define $\tilde{\varphi}$ and $\tilde{\mathfrak{P}}$ as in Theorem A.4, and let $((\tilde{\pi}^+_k)_{k \in \mathcal{K}^+}, (\tilde{\pi}^-_k)_{k \in \mathcal{K}^-}, (\mathrm{id}, \mathrm{id})_\# \eta)$ denote $\tilde{\varphi}^{-1}(\pi) \in \tilde{\mathfrak{P}}$. By Theorem A.4, $\pi - H(\mu, \nu) = \sum_{k \in \mathcal{K}^+} \tilde{\pi}^+_k + \sum_{k \in \mathcal{K}^-} \tilde{\pi}^-_k$. According to Equation (46), $\pi - H(\mu, \nu) = \sum_{i \in \mathcal{I} \setminus \mathcal{I}^=} \left(\sum_{k \in K^+_i} \tilde{\pi}^+_k + \sum_{k \in K^-_i} \tilde{\pi}^-_k\right)$. Let $(\pi_i)_{i \in \mathcal{I}} \in \prod_{i \in \mathcal{I}} \mathfrak{C}(\mu_i, \nu_i)$ be the family satisfying $\pi = \sum_{i \in \mathcal{I}} \pi_i$. By Lemma C.5, $\pi - H(\mu, \nu) = \sum_{i \in \mathcal{I} \setminus \mathcal{I}^=} \pi_i - H_i$. Therefore,

$$\sum_{i \in \mathcal{I} \setminus \mathcal{I}^=} \pi_i - H_i = \sum_{i \in \mathcal{I} \setminus \mathcal{I}^=} \left(\sum_{k \in K^+_i} \tilde{\pi}^+_k + \sum_{k \in K^-_i} \tilde{\pi}^-_k \right). \tag{47}$$

Moreover, for all $i \in \mathcal{I} \setminus \mathcal{I}^=$, $\pi_i - H_i$ is concentrated on A^i_* and $\sum_{k \in K^+_i} \tilde{\pi}^+_k + \sum_{k \in K^-_i} \tilde{\pi}^-_k$ is concentrated on $\left(\bigcup_{k \in K^+_i} A^+_k\right) \cup \left(\bigcup_{k \in K^-_i} A^-_k\right) \subset A^i_*$. As $(A^i_*)_{i \in \mathcal{I} \setminus \mathcal{I}^=}$ is a family of disjoint set, by restricting Equation (47) to A^i_*, we obtain

$$\pi_i - H_i = \sum_{k \in K^+_i} \tilde{\pi}^+_k + \sum_{k \in K^-_i} \tilde{\pi}^-_k.$$

Projecting onto the first and second marginals yields the desired equalities. □

We now prove that $\preceq_{\mathcal{A}^*}$ defines a partial order on $\mathcal{A}^*$. Strictly speaking, this is false. If $\mathcal{D} = ((\mu_i, \nu_i))_{i \in \mathcal{I}} \in \mathcal{A}^*$ and ϕ is a permutation of $\mathcal{I}$, then the family $\mathcal{D}'$ defined as $\mathcal{D}' = ((\mu_{\phi(i)}, \nu_{\phi(i)}))_{i \in \mathcal{I}}$ belongs to $\mathcal{A}^*$: we have $\mathcal{D} \preceq_{\mathcal{A}^*} \mathcal{D}'$ and $\mathcal{D}' \preceq_{\mathcal{A}^*} \mathcal{D}$, whereas $\mathcal{D}$ and $\mathcal{D}'$ are distinct. We overcome this obstacle by introducing an equivalence relation $\sim$ as follows.

Definition C.13. 1. Let E be an arbitrary set and $(x_i)_{i\in\mathcal{I}}$, $(y_j)_{j\in\mathcal{J}}$ two families of elements of E. We say that $(y_j)_{j\in\mathcal{J}}$ is a reordering of $(x_i)_{i\in\mathcal{I}}$ if there exists a bijection $\phi:\mathcal{I}\to\mathcal{J}$ such that $x_{\phi(i)}=y_i$ for all $i\in\mathcal{I}$. In this case, we write $(x_i)_{i\in\mathcal{I}}\sim(y_j)_{j\in\mathcal{J}}$. The reader may easily verify that $\sim$ is an equivalence relation on families of element of E.

2. Let $\overline{\mathcal{A}^*}$ be the quotient set of $\mathcal{A}^*$ by the previously defined relation $\sim$. For all $\mathcal{D}\in\mathcal{A}^*$ we denote $\overline{\mathcal{D}}$ the equivalence class of $\mathcal{D}$ under $\sim$.

3. Given $\overline{\mathcal{D}_1},\overline{\mathcal{D}_2}\in\mathcal{A}^*$, we write $\overline{\mathcal{D}_1}\preceq_{\overline{\mathcal{A}^*}}\overline{\mathcal{D}_2}$ if $\mathcal{D}_1\preceq_{\mathcal{A}^*}\mathcal{D}_2$. Observe that this relation is well defined as $\preceq_{\mathcal{A}^*}$ is class invariant: if $\mathcal{D}_1\preceq_{\mathcal{A}^*}\mathcal{D}_2$, $\mathcal{D}_1\sim\mathcal{D}_1'$ and $\mathcal{D}_2\sim\mathcal{D}_2'$, then $\mathcal{D}_1'\preceq_{\mathcal{A}^*}\mathcal{D}_2'$.

In order to prove that $\preceq_{\overline{\mathcal{A}^*}}$ defines a partial order on $\overline{\mathcal{A}^*}$, we need a technical lemma.

Lemma C.14. *Assume that $(\gamma_i)_{i\in\mathcal{I}}\in(\mathcal{M}_+(\mathbb{R})\setminus\{0\})^{\mathcal{I}}$ is such that $\sum_{i\in\mathcal{I}}\gamma_i(\mathbb{R})<+\infty$ and there exists a partition $(J_i)_{i\in\mathcal{I}}$ of $\mathcal{I}$ satisfying the following condition:*

$$\forall i\in\mathcal{I},\gamma_i=\sum_{j\in J_i}\gamma_j. \tag{48}$$

Then, there exists a bijection $\phi:\mathcal{I}\to\mathcal{I}$ such that, for all $i\in\mathcal{I}$, $J_i=\{\phi(i)\}$.

Proof. Since $(J_i)_{i\in\mathcal{I}}$ is a partition of $\mathcal{I}$, there exists $\varphi:\mathcal{I}\to\mathcal{I}$ such that, for all $i\in\mathcal{I}$, $i\in J_{\varphi(i)}$. To derive a contradiction, assume there exists $i\in\mathcal{I}$ such that $\gamma_i\neq\gamma_{\varphi(i)}$. Since $\gamma_\varphi(l)=\sum_{j\in\varphi(l)}\gamma_j\geq\gamma_l$ for each $l\in\mathcal{I}$, we obtain $+\infty>\sum_{j\in\{\varphi^l(i)\ ;\ l\geq1\}}\gamma_j(\mathbb{R})\geq\operatorname{card}(\{\varphi^l(i)\ ;\ l\geq1\})\gamma_i(\mathbb{R})$. Thus $\operatorname{card}(\{\varphi^l(i)\ ;\ l\geq1\})\gamma_i(\mathbb{R})<+\infty$, which implies that $k:=\inf\left(\{l\geq0\ ;\ \varphi^{l+1}(i)\in\{i,\ldots,\varphi^l(i)\}\}\right)$ is finite. Let $j\geq0$ be such that $j\in[\![0,k-1]\!]$ and $\varphi^{j+1}(i)=\varphi^{k+1}(i)$. Since $\varphi^k(i)\neq\varphi^j(i)$ and $(\varphi^k(i),\varphi^j(i))\in J_{\varphi^{k+1}(i)}\times J_{\varphi^{j+1}(i)}=J^2_{\varphi^{j+1}(i)}$, we obtain $\gamma_{\varphi^{j+1}(i)}=\sum_{l\in J_{\varphi^{j+1}(i)}}\gamma_l\geq\gamma_{\varphi^j(i)}+\gamma_{\varphi^k(i)}\geq\gamma_{\varphi^j(i)}+\gamma_{\varphi^{j+1}(i)}$. This implies $\gamma_{\varphi^j(i)}=0$, which is a contradiction. Therefore $\gamma_i=\gamma_{\varphi(i)}$, which implies $\gamma_i=\gamma_i+\sum_{l\in J_{\varphi(i)\setminus\{i\}}}\gamma_l$. Thus $J_{\varphi(i)}\setminus\{i\}=\emptyset$, i.e., $J_{\varphi(i)}=\{i\}$. It is now sufficient to prove that φ is a bijection: surjectivity is a consequence from $J_i\neq\emptyset$, while injectivity follows from $J_{\varphi(i)}=\{i\}$. □

Theorem C.15. *The relation $\preceq_{\overline{\mathcal{A}^*}}$ constitutes a partial order on $\overline{\mathcal{A}^*}$. Furthermore, $(\overline{\mathcal{A}^*},\preceq_{\overline{\mathcal{A}^*}})$ admits a minimum element, given by $\overline{\mathcal{D}^*_{\mathcal{K}}}=\overline{\{(\tilde\mu_k^+,\tilde\nu_k^+)\}_{k\in\mathcal{K}^+}\cup\{(\tilde\mu_k^-,\tilde\nu_k^-)\}_{k\in\mathcal{K}^-}\cup\{(\eta,\eta)\}}$.*

Proof. The reflexivity and transitivity of $\preceq_{\mathcal{A}^*}$ follow directly from the definitions and are therefore omitted. Reflexivity and transitivity of $\leq_{\overline{\mathcal{A}^*}}$ directly follows. We now establish that $\preceq_{\mathcal{A}^*}$ is antisymmetric. Let $\mathcal{D}_1=\{(\mu_i^1,\nu_i^1)\}_{i\in\mathcal{I}_1}$ and $\mathcal{D}_2=\{(\mu_i^2,\nu_i^2)\}_{i\in\mathcal{I}_2}$ be two admissible decompositions such that $\mathcal{D}_1\preceq_{\mathcal{A}^*}\mathcal{D}_2$ and $\mathcal{D}_2\preceq_{\mathcal{A}^*}\mathcal{D}_1$. Let $(J^2_{i_2})_{i_2\in\mathcal{I}_2\setminus\mathcal{I}_2^=}$ be a partition of $\mathcal{I}_1\setminus\mathcal{I}_1^=$ and $(J^1_{i_1})_{i_1\in\mathcal{I}_1\setminus\mathcal{I}_1^=}$ a partition of $\mathcal{I}_2\setminus\mathcal{I}_2^=$ such that, for all $(i_1,i_2)\in(\mathcal{I}_1\setminus\mathcal{I}_1^=)\times(\mathcal{I}_2\setminus\mathcal{I}_2^=)$,

$$\begin{cases}\mu^2_{i_2}-\eta^2_{i_2}=\sum_{j_1\in J^2_{i_2}}\mu^1_{j_1}-\eta^1_{j_1}\\ \nu^2_{i_2}-\eta^2_{i_2}=\sum_{j_1\in J^2_{i_2}}\nu^1_{j_1}-\eta^1_{j_1}\\ \eta^2_{i_2}\geq\sum_{j_1\in J^2_{i_2}}\eta^1_{j_1}\end{cases}\quad\text{and}\quad\begin{cases}\mu^1_{j_1}-\eta^1_{j_1}=\sum_{j_2\in J^1_{j_1}}\mu^2_{j_2}-\eta^2_{j_2}\\ \nu^1_{j_1}-\eta^1_{j_1}=\sum_{j_2\in J^1_{j_1}}\nu^2_{j_2}-\eta^2_{j_2}\\ \eta^1_{j_1}\geq\sum_{j_2\in J^1_{j_1}}\eta^2_{j_2}\end{cases}. \tag{49}$$

Consider $i_2 \in \mathcal{I}_2 \setminus \mathcal{I}_2^=$ and define $\tilde{J}_{i_2} = \biguplus_{j_1 \in J_{i_2}^2} J_{j_1}^1$. By Equation (49),

$$\begin{cases} \mu_{i_2}^2 - \eta_{i_2}^2 = \sum_{j_2 \in \tilde{J}_{i_2}} \mu_{j_2}^2 - \eta_{j_2}^2 \\ \nu_{i_2}^2 - \eta_{i_2}^2 = \sum_{j_2 \in \tilde{J}_{i_2}} \nu_{j_2}^2 - \eta_{j_2}^2 \end{cases} \tag{50}$$

and

$$\eta_{i_2}^2 \geq \sum_{j_1 \in J_{i_2}} \eta_{j_1}^1 \geq \sum_{j_2 \in \tilde{J}_{i_2}} \eta_{j_2}^2. \tag{51}$$

From Proposition C.8, it follows

$$\mathfrak{C}(\mu_{i_2}^2 - \eta_{i_2}^2, \nu_{i_2}^2 - \eta_{i_2}^2) = \bigoplus_{j_2 \in \tilde{J}_{i_2}} \mathfrak{C}(\mu_{j_2}^2 - \eta_{j_2}^2, \nu_{j_2}^2 - \eta_{j_2}^2).$$

By Point 5 of Remark C.4, $\tilde{J}_{i_2} = \{i_2\}$. Since $(J_{j_1}^1)_{j_1 \in J_{i_2}}$ is a family of non-empty disjoint subsets, there exists $\phi(i_2) \in \mathcal{I}_1 \setminus \mathcal{I}_1^=$ such that $J_{i_2}^2 = \{\phi(i_2)\}$ and $J_{\phi(i_2)}^1 = \{i_2\}$. By the two first lines of Equation (49), we obtain

$$\begin{cases} \mu_{i_2}^2 - \eta_{i_2}^2 = \mu_{\phi(i_2)}^1 - \eta_{\phi(i_2)}^1 \\ \nu_{i_2}^2 - \eta_{i_2}^2 = \nu_{\phi(i_2)}^1 - \eta_{\phi(i_2)}^1 \end{cases},$$

while Equation (51) gives $\eta_{i_2}^2 \geq \eta_{\phi(i_2)}^1 \geq \eta_{i_2}^2$. Finally, we conclude that $(\mu_{i_2}^2, \nu_{i_2}^2) = (\mu_{\phi(i_2)}^1, \nu_{\phi(i_2)}^1)$. Observe that ϕ defines a bijection from $\mathcal{I}_2 \setminus \mathcal{I}_2^=$ to $\mathcal{I}_1 \setminus \mathcal{I}_1^=$: this follows from the fact that $(J_{i_2}^2)_{i_2 \setminus \mathcal{I}_2^=}$ is a partition of $\mathcal{I}_1 \setminus \mathcal{I}_1^=$ and the equalities $J_{i_2}^2 = \{\phi(i_2)\}$ for $i_2 \in \mathcal{I}_2 \setminus \mathcal{I}_2^=$. In particular $\sum_{i_1 \in \mathcal{I}_1 \setminus \mathcal{I}_1^=} \eta_{i_1}^1 = \sum_{i_2 \in \mathcal{I}_2 \setminus \mathcal{I}_2^=} \eta_{\phi(i_2)}^2 = \sum_{i_2 \in \mathcal{I}_2 \setminus \mathcal{I}_2^=} \eta_{i_2}^2$. From Lemma C.5, it follows $\sum_{i_1 \in \mathcal{I}_1^=} \eta_{i_1}^1 = \sum_{i_2 \in \mathcal{I}_2^=} \eta_{i_2}^2$. As $\mathcal{D}_1 \preceq_{\mathcal{A}^*} \mathcal{D}_2$ and $\mathcal{D}_2 \preceq_{\mathcal{A}^*} \mathcal{D}_1$, there exists a disjoint covering of $(J_{i_1}^1)_{i_1 \in \mathcal{I}_1^=}$ of $\mathcal{I}_2^=$ and a disjoint covering $(J_{i_2}^2)_{i_2 \in \mathcal{I}_2^=}$ of $\mathcal{I}_1 \setminus \mathcal{I}_1^=$ such that,

$$\begin{cases} \forall i_1 \in \mathcal{I}_1^=, \eta_{i_1}^1 \geq \sum_{j_2 \in J_{i_1}^1} \eta_{j_2}^2 \\ \forall i_2 \in \mathcal{I}_1^=, \eta_{i_2}^2 \geq \sum_{j_1 \in J_{i_2}^2} \eta_{j_1}^1 \end{cases}.$$

Thus

$$\sum_{i_2 \in \mathcal{I}_2} \left(\eta_{i_2}^2 - \sum_{j_1 \in J_{i_2}^2} \eta_{i_1}^1 \right) = \sum_{i_2 \in \mathcal{I}_2^=} \eta_{i_2}^2 - \sum_{i_2 \in \mathcal{I}_2^=} \sum_{j_1 \in J_{i_2}^2} \eta_{i_1}^1 = \sum_{i_2 \in \mathcal{I}_2^=} \eta_{i_2}^2 - \sum_{i_1 \in \mathcal{I}_1^=} \eta_{i_1}^1 = 0, \tag{52}$$

which implies

$$\eta_{i_2}^2 = \sum_{j_1 \in J_{i_2}^2} \eta_{i_1}^1 \tag{53}$$

for every $i_2 \in \mathcal{I}_2^=$. Similarly, for all $i_1 \in \mathcal{I}_1^=$, $\eta_{i_1}^1 = \sum_{i_2 \in J_{i_1}^1} \eta_{i_2}^2$. From the previous equalities, the sets $J_{i_1}^1$ and $J_{i_2}^2$ are non-empty. Now, define $\tilde{J}_{i_2}^2 = \biguplus_{i_1 \in J_{i_2}^2} J_{i_1}^1$. For all $i_2 \in \mathcal{I}_2^=$, $\eta_{i_2}^2 = \sum_{j_2 \in \tilde{J}_{i_2}^2} \eta_{j_2}^2$. Since $(J_{i_1})_{i_1 \in \mathcal{I}_1^=}$ is a partition of $\mathcal{I}_2^=$ and $(J_{i_2}^2)_{i_2 \in \mathcal{I}_2^=}$ is a partition of $\mathcal{I}_1^=$, it directly follows that $(\tilde{J}_{i_2}^2)_{i_2 \in \mathcal{I}_2^=}$ is a partition of $\mathcal{I}_2^=$. According to Lemma C.14, there exists bijection $\psi : \mathcal{I}_2^= \to \mathcal{I}_2^=$ such that $\tilde{J}_{i_2}^2 = \{\psi(i_2)\}$ for all $i_2 \in \mathcal{I}_2^=$. This implies that, for all $i_2 \in \mathcal{I}_2^=$, there exists a unique $\phi(i_2) \in \mathcal{I}_2^=$ such that $J_{i_2}^2 = \{\phi(i_2)\}$. From Equation

(53), it follows that $\eta^2_{i_2} = \eta^1_{\phi(i_2)}$, i.e., $(\mu^2_{i_2}, \nu^2_{i_2}) = (\mu^1_{\phi(i_2)}, \nu^1_{\phi(i_2)})$. Observe that ϕ is a bijection from $\mathcal{I}_2^=$ to $\mathcal{I}_1^=$: this follows directly from the fact that $(J^2_{i_2})_{i_2 \in \mathcal{I}_2^=}$ is a partition of $\mathcal{I}_1^=$ and the equalities $J_{i_2} = \{\phi(i_2)^2\}$ for $i_2 \in \mathcal{I}_2^=$. We have thus established the existence of a bijection $\phi : (\mathcal{I}_2 \setminus \mathcal{I}_2^=) \uplus \mathcal{I}_1^= \to (\mathcal{I}_1 \setminus \mathcal{I}_1^=) \uplus \mathcal{I}_1^=$ such that, for all $i_2 \in \mathcal{I}_2$, we have $(\mu^2_{i_2}, \nu^2_{i_2}) = (\mu^1_{\phi(i_1)}, \nu^1_{\phi(i_1)})$. Therefore $\mathcal{D}_1 \sim \mathcal{D}_2$, and antisymmetry of $\overline{\mathcal{A}^*}$ follows directly.

We now prove that $\overline{\mathcal{D}^*_{\mathcal{K}}}$ is a minimal element of $(\overline{\mathcal{A}^*}, \preceq_{\overline{\mathcal{A}^*}})$, that is, $\mathcal{D}^*_{\mathcal{K}} \leq \mathcal{D}$ for every $\mathcal{D} \in \mathcal{A}^*$. By Point 3 of Remark C.4, $\mathcal{D}_{\mathcal{K}} \in \mathcal{A}^*$. By Remark B.6, for all $k \in \mathcal{K}^+$ (resp. $k \in \mathcal{K}^-$), $\eta(\tilde{\mu}^+_k, \tilde{\nu}^+_k) = 0$ (resp. $\eta(\tilde{\mu}^-_k, \tilde{\nu}^-_k) = 0$). Consider $\mathcal{D} = ((\mu_i, \nu_i))_{i \in I} \in \mathcal{A}^*$. For every $i \in \mathcal{I} \setminus \mathcal{I}^=$, define $J_i = K^-_i \uplus K^+_i$. According to Equation (46), $(J_i)_{i \in \mathcal{I} \setminus \mathcal{I}^=}$ is a partition of $\mathcal{K}^- \uplus \mathcal{K}^+$. By Lemma C.12, for all $i \in \mathcal{I} \setminus \mathcal{I}^=$,

$$\begin{cases} \mu_i - \eta_i = \sum_{k \in K^+_i} (\tilde{\mu}^+_k - 0) + \sum_{k \in K^-_i} (\tilde{\mu}^-_k - 0) \\ \nu_i - \eta_i = \sum_{k \in K^+_i} (\tilde{\nu}^+_k - 0) + \sum_{k \in K^-_i} (\tilde{\nu}^-_k - 0) \end{cases},$$

and

$$\eta_i \geq \sum_{k \in K^+_i} 0 + \sum_{k \in K^-_i} 0.$$

This establishes Requirement (41). Condition (40)follows directly from Lemma C.5: indeed $\eta = \sum_{i \in \mathcal{I}} \eta_i \geq \sum_{i \in \mathcal{I}^=} \eta_i$. Therefore $\mathcal{D}_{\mathcal{K}} \preceq_{\mathcal{A}^*} \mathcal{D}$, which establishes our claim. ☐

We are now ready to prove Theorem 2.47.

Definition C.16. We define $\mathcal{A}^\triangle$ as the set of elements $\mathcal{D} = ((\mu_i, \nu_i))_{i \in \mathcal{I}} \in \mathcal{A}^*$ such that there exists a family $(J_i)_{i \in \mathcal{I} \setminus \mathcal{I}^=}$ and families $((a^j_i, b^j_i))_{j \in J_i} \in (\mathbb{R} \times \mathbb{R})^{J_i}$ with $a^j_i < b^j_i$ such that the following two conditions are satisfied:

1. For every $i \in \mathcal{I}^=$, μ_i is concentrated on $\bigcap_{l \in \mathcal{I} \setminus \mathcal{I}^=} \bigcap_{j \in J_l}]a^j_l, b^j_l[^c$.

2. For all $i \in \mathcal{I} \setminus \mathcal{I}^=$, μ_i and ν_i are concentrated on $\bigcup_{j \in J_i} [a^j_i, b^j_i]$, and:

$$\forall (i_1, i_2) \in (\mathcal{I} \setminus \mathcal{I}^=)^2, \forall (j_1, j_2) \in J_{i_1} \times J_{i_2} : (i_1, j_1) \neq (i_2, j_2) \implies (b^{j_2}_{i_2} \leq a^{j_1}_{i_1} \text{ or } b^{j_1}_{i_1} \leq a^{j_2}_{i_2}). \quad (54)$$

Theorem C.17. *The decomposition $\overline{\mathcal{D}_{\mathcal{K}}}$ is the minimum element of $\overline{\mathcal{A}^\triangle}$ with respect to the order $\preceq_{\mathcal{A}^*}$.*

Proof. Consider $\mathcal{D} = ((\mu_i, \nu_i))_{i \in \mathcal{I}} \in \mathcal{A}^\triangle$, and let the families $((a^j_i, b^j_i))_{j \in J_i}$ for $i \in \mathcal{I} \setminus \mathcal{I}^=$ be as in Definition C.16. For every $i \in \mathcal{I} \setminus \mathcal{I}^=$, define

$$B_i = \bigcup_{j \in J_i} \left([a^j_i, b^j_i]^2 \setminus \{(a^j_i, a^j_i)\,(b^j_i, b^j_i)\} \right).$$

For every $i \in \mathcal{I}$, the measures μ_i and ν_i are concentrated on $\bigcup_{j \in J_i} [a^j_i, b^j_i]$: it directly follows that $\mathcal{C}_i \preceq ([a^j_i, b^j_i]^2)_{j \in J_i}$. Therefore, for every $k \in K^+_i$, we have $A^+_k \subset A^*_i \subset B_i$: as $(\mathrm{id}, \mathrm{id})_\# \eta^+_k$ is concentrated on

A_k^+, we obtain $(\mathrm{id},\mathrm{id})_\# \eta_k^+ \llcorner_{B_i} = (\mathrm{id},\mathrm{id})_\# \eta_k^+$. Similarly, for all $k \in \mathcal{K}_i^-$, $(\mathrm{id},\mathrm{id})_\# \eta_k^- \llcorner_{B_i} = (\mathrm{id},\mathrm{id})_\# \eta_k^-$. Therefore,

$$\begin{aligned} H(\mu,\nu)\llcorner_{B_i} &= \left(\sum_{k\in\mathcal{K}^+} (\mathrm{id},\mathrm{id})_\# \eta_k^+ + \sum_{k\in\mathcal{K}^+} (\mathrm{id},\mathrm{id})_\# \eta_k^+ + (\mathrm{id},\mathrm{id})_\# \mu^= \right)\llcorner_{B_i} \\ &\geq \sum_{k\in K_i^+} (\mathrm{id},\mathrm{id})_\# \eta_k^+ + \sum_{k\in K_i^-} (\mathrm{id},\mathrm{id})_\# \eta_k^-. \end{aligned} \tag{55}$$

Moreover, for every $l \in \mathcal{I} \setminus \mathcal{I}^=$ with $l \neq i$, H_l is concentrated on $\bigcup_{j\in B_l} [a_l^j, b_l^j] \subset B_i^c$. For every $l \in \mathcal{I}^=$, H_l is concentrated on $\bigcap_{k\notin\mathcal{I}^=} \bigcap_{j\in J_k} \left( ]a_k^j, b_k^j[^2 \right)^c \subset \bigcap_{j\in J_i} \left( ]a_i^j, b_i^j[^2 \right)^c \subset B_i^c$. Therefore, by applying Lemma C.5, we obtain

$$H_i \geq H_i\llcorner_{B_i} = \sum_{l\in\mathcal{I}\setminus\{i\}} H_l\llcorner_{B_i} + H_i\llcorner_{B_i} = H(\mu,\nu)\llcorner_{B_i}. \tag{56}$$

From Equation (55), it follows that $H_i \geq \sum_{k\in K_i^+} (\mathrm{id},\mathrm{id})_\# \eta_k^+ + \sum_{k\in K_i^-} (\mathrm{id},\mathrm{id})_\# \eta_k^-$, i.e.,

$$\eta_i \geq \sum_{k\in K_i^+} \eta_k^+ + \sum_{k\in K_i^-} \eta_k^-. \tag{57}$$

Moreover, according to Lemma C.12, we have

$$\begin{cases} \mu_i - \eta_i = \sum_{k\in K_i^+} (\mu_k^+ - \eta_k^+) + \sum_{k\in K_i^-} (\mu_k^- - \eta_k^-) \\ \nu_i - \eta_i = \sum_{k\in K_i^+} (\nu_k^+ - \eta_k^+) + \sum_{k\in K_i^-} (\nu_k^- - \eta_k^-) \end{cases}.$$

Combined with Equation (57), this establishes Condition (40). Moreover,

$$\sum_{i\in\mathcal{I}^=} \eta_i = \eta - \sum_{i\in\mathcal{I}\setminus\mathcal{I}^=} \eta_i \leq \eta - \sum_{i\in\mathcal{I}\setminus\mathcal{I}^=} \left(\sum_{k\in K_i^+} \eta_k^+ + \sum_{k\in K_i^-} \eta_k^- \right) = \eta - \sum_{k\in\mathcal{K}^+} \eta_k^+ - \sum_{k\in\mathcal{K}^-} \eta_k^- = \mu^=,$$

which establishes Condition (40). Thus $\mathcal{D}_\mathcal{K} \preceq_{\mathcal{A}^*} \mathcal{D}$, thereby establishing the claim. □

Remark C.18. Observe that $\overline{\mathcal{A}} \subset \overline{\mathcal{A}^\triangle} \subset \overline{\mathcal{A}^*}$. Since $\overline{\mathcal{D}_\mathcal{K}} \in \overline{\mathcal{A}}$ is the minimum of $\overline{\mathcal{A}^\triangle}$ with respect to $\preceq_{\mathcal{A}^*}$, the proof of Theorem 2.47 boils down to the following condition:

$$\forall (\mathcal{D}_1, \mathcal{D}_2) \in \mathcal{A}^2 : \mathcal{D}_1 \preceq_{\mathcal{A}^*} \mathcal{D}_2 \implies \mathcal{D}_1 \preceq_\mathcal{A} \mathcal{D}_2.$$

This condition follows directly from Proposition C.8.

References

[1] Luigi Ambrosio, *Lecture Notes on Optimal Transport Problems*, Mathematical Aspects of Evolving Interfaces: Lectures given at the C.I.M.-C.I.M.E. joint Euro-Summer School held in Madeira, Funchal, Portugal, July 3-9, 2000, Springer, Berlin, Heidelberg, 2003, pp. 1–52.

[2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient Flows*, Birkhäuser, Basel, 2008 (en).

[3] Shrey Aryan and Promit Ghosal, *Entropic Selection Principle for Monge's Optimal Transport*, February 2025, arXiv:2502.16370.

[4] Mathias Beiglböck and Nicolas Juillet, *On a problem of optimal transport under marginal martingale constraints*, The Annals of Probability **44** (2016), no. 1, 42–106, Publisher: Institute of Mathematical Statistics.

[5] Espen Bernton, Promit Ghosal, and Marcel Nutz, *Entropic optimal transport: Geometry and large deviations*, Duke Mathematical Journal **171** (2022), no. 16, 3363–3400.

[6] Charles Boubel and Nicolas Juillet, *The Markov-quantile process attached to a family of marginals*, J. Éc. Polytech., Math. **9** (2022), 1–62 (English).

[7] Andrea Braides, *[Gamma]-convergence for beginners*, Oxford lecture series in mathematics and its applications, no. 22, Oxford University Press, Oxford, 2002 (eng).

[8] Guillaume Carlier, Vincent Duval, Gabriel Peyré, and Bernhard Schmitzer, *Convergence of entropic schemes for optimal transport and gradient flows*, SIAM J. Math. Anal. **49** (2017), no. 2, 1385–1418 (English).

[9] Marco Cuturi, *Sinkhorn Distances: Lightspeed Computation of Optimal Transport*, Advances in neural information processing systems **26** (2013) (en).

[10] Simone Di Marino and Jean Louet, *The entropic regularization of the Monge problem on the real line*, SIAM J. Math. Anal. **50** (2018), no. 4, 3451–3477 (English).

[11] Wilfrid Gangbo and Robert J. McCann, *The geometry of optimal transportation*, Acta Math. **177** (1996), no. 2, 113–161 (en).

[12] Promit Ghosal, Marcel Nutz, and Espen Bernton, *Stability of entropic optimal transport and Schrödinger bridges*, Journal of Functional Analysis **283** (2022), no. 9, 109622 (en).

[13] Nassif Ghoussoub, Young-Heon Kim, and Tongseok Lim, *Structure of optimal martingale transport plans in general dimensions*, The Annals of Probability **47** (2019), no. 1, 109–164.

[14] Nicolas Juillet, *On a Solution to the Monge Transport Problem on the Real Line Arising from the Strictly Concave Case*, SIAM J. Math. Anal. **52** (2020), no. 5, 4783–4805.

[15] Hans G. Kellerer, *Order conditioned independence of real random variables*, Math. Ann. **273** (1986), no. 3, 507–528 (en).

[16] ———, *Markov property of point processes*, Probab. Th. Rel. Fields **76** (1987), no. 1, 71–80 (en).

[17] Christian Léonard, *From the Schrödinger problem to the Monge–Kantorovich problem*, Journal of Functional Analysis **262** (2012), no. 4, 1879–1920 (en).

[18] Francesco Maggi, *Optimal Mass Transport on Euclidean Spaces*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2023.

[19] Hadrien De March and Nizar Touzi, *Irreducible convex paving for decomposition of multidimensional martingale transport plans*, The Annals of Probability **47** (2019), no. 3, 1726–1774.

[20] Marcel Nutz, *Introduction to Entropic Optimal Transport* (en).

[21] Jan Obłój and Pietro Siorpaes, *Structure of martingale transports in finite dimensions*, February 2017, arXiv:1702.08433.

[22] L. De Pascale, A. Kausamo, and K. Wyczesany, *60 years of cyclic monotonicity: a survey*, February 2024, arXiv:2308.07682.

[23] Gabriel Peyré and Marco Cuturi, *Computational Optimal Transport: With Applications to Data Science*, MAL **11** (2019), no. 5-6, 355–607 (English).

[24] S. T. Rachev and Ludger Rüschendorf, *Mass transportation problems*, Probability and its applications, Springer, New York, 1998.

[25] Filippo Santambrogio, *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*, Progress in Nonlinear Differential Equations and Their Applications, vol. 87, Springer International Publishing, Cham, 2015 (en).

[26] Richard Sinkhorn, *A Relationship Between Arbitrary Positive Matrices and Doubly Stochastic Matrices*, The Annals of Mathematical Statistics **35** (1964), no. 2, 876–879.

[27] V. Strassen, *The existence of probability measures with given marginals*, Ann. Math. Stat. **36** (1965), 423–439 (English).

[28] V. N. Sudakov, *Geometric problems in the theory of infinite-dimensional probability distributions*, Proceedings of the Steklov Institute of Mathematics ; 1979, issue 2, American Mathematical Society, Providence, R. I, 1979 (eng).

[29] Cédric Villani, *Topics in optimal transportation*, vol. 58, American Mathematical Society (AMS), Providence, RI, 2003 (English).

[30] ———, *Optimal transport. Old and new*, Grundlehren Math. Wiss., vol. 338, Springer, Berlin, 2009 (English), ISSN: 0072-7830.