Irreducible Convex Paving for Decomposition of Multi-dimensional Martingale Transport Plans *

Hadrien De March† Nizar Touzi‡

February 28, 2017

Abstract

Martingale transport plans on the line are known from Beiglböck & Juillet [3] to have an irreducible decomposition on a (at most) countable union of intervals. We provide an extension of this decomposition for martingale transport plans in $\mathbb{R}^d$, $d \geqslant 1$. Our decomposition is a partition of $\mathbb{R}^d$ consisting of a possibly uncountable family of relatively open convex components, with the required measurability so that the disintegration is well-defined. We justify the relevance of our decomposition by proving the existence of a martingale transport plan filling these components. We also deduce from this decomposition a characterization of the structure of polar sets with respect to all martingale transport plans.

Key words. Martingale optimal transport, irreducible decomposition, polar sets.

1 Introduction

The problem of martingale optimal transport was introduced by P. Henry-Labordère as the dual of the problem of robust (model-free) superhedging of exotic derivatives in financial mathematics, see Beiglböck, Henry-Labordère & Penkner [2] in discrete time, and Galichon, Henry-Labordère & Touzi [9] in continuous-time. This robust superhedging problem was introduced by Hobson [14], and was addressing specific examples of exotic derivatives by means of corresponding solutions of the Skorohod embedding problem, see [6, 15, 16], and the survey [14].

Given two probability measures μ, ν on $\mathbb{R}^d$, with finite first order moment, martingale optimal transport differs from standard optimal transport in that the set of all interpolating

*The authors gratefully acknowledge the financial support of the ERC 321111 Rofirm, and the Chairs Financial Risks (Risk Foundation, sponsored by Société Générale) and Finance and Sustainable Development (IEF sponsored by EDF and CA).

†CMAP, École Polytechnique, hadrien.de-march@polytechnique.org.

‡CMAP, École Polytechnique, nizar.touzi@polytechnique.edu.

probability measures $\mathcal{P}(\mu,\nu)$ on the product space is reduced to the subset $\mathcal{M}(\mu,\nu)$ restricted by the martingale condition. We recall from Strassen [20] that $\mathcal{M}(\mu,\nu) \neq \emptyset$ if and only if $\mu \preceq \nu$ in the convex order, i.e. $\mu(f) \leqslant \nu(f)$ for all convex functions f. Notice that the inequality $\mu(f) \leqslant \nu(f)$ is a direct consequence of the Jensen inequality, the reverse implication follows from the Hahn-Banach theorem.

This paper focuses on the critical observation by Beiglböck & Juillet [3] that, in the one-dimensional setting $d = 1$, any such martingale interpolating probability measure $\mathbb{P}$ has a canonical decomposition $\mathbb{P} = \sum_{k\geqslant 0} \mathbb{P}_k$, where $\mathbb{P}_k \in \mathcal{M}(\mu_k,\nu_k)$ and μ_k,ν_k are the restrictions of μ,μ to the so-called irreducible components I_k, J_k, respectively, $k \geqslant 0$. Here, $(I_k)_{k\geqslant 1}$ are open intervals, $I_0 := \mathbb{R}\backslash(\cup_{k\geqslant 1} I_k)$, and J_k is an augmentation of I_k by the inclusion of either one of the endpoints of I_k, depending on whether they are charged by the distribution $\mathbb{P}_k$. Remarkably, the irreducible components $(I_k, J_k)_{k\geqslant 0}$ is independent of the choice of $\mathbb{P} \in \mathcal{M}(\mu,\nu)$. To understand this decomposition, notice that convex functions in one dimension are generated by the family $f_c(x) := |x - c|$, $x \in \mathbb{R}$, $c \in \mathbb{R}$. Then, in terms of the potential functions $U^\mu(c) := \mu(f_c)$, and $U^\nu(c) := \nu(f_c)$, $c \in \mathbb{R}$, we have $\mu \preceq \nu$ if and only if $U^\mu \leqslant U^\nu$ and μ,ν have same mean. Then, at any contact points c, of the potential functions, $U^\mu = U^\nu$, we have equality in the underlying Jensen equality, which means that the singularity c of the underlying function f_c is not seen by the measure. In other words, the point c acts as a barrier for the mass transfer in the sense that martingale transport maps do not cross the barrier c. Such contact points are precisely the endpoints of the intervals I_k, $k \geqslant 1$.

The decomposition in irreducible components plays a crucial role for the quasi-sure formulation introduced by Beiglböck, Nutz and Touzi [4], and represents an important difference between martingale transport and standard transport. Indeed, while the martingale transport problem is affected by the quasi-sure formulation, the standard optimal transport problem is not changed. We also refer to Ekren & Soner [8] for further functional analytic aspects of this duality.

Our objective in this paper is to extend the last decomposition to an arbitrary $d-$dimensional setting, $d \geqslant 1$. The main difficulty is that convex functions do not have anymore such a simple generating family. Therefore, all of our analysis is based on the set of convex functions. A first attempt to extend the last decomposition to the multi-dimensional case was achieved by Ghoussoub, Kim & Lim [10]. Motivated by the martingale monotonicity principle of Beiglböck & Juillet [3] (see also Zaev [22] for higher dimension and general linear constraints), their strategy is to find a monotone set $\Gamma \subset \mathbb{R}^d \times \mathbb{R}^d$, where the robust superhedging holds with equality, as a support of the optimal martingale transport in $\mathcal{M}(\mu,\nu)$. Denoting $\Gamma_x := \{y : (x,y) \in \Gamma\}$, this naturally induces the relation $x \,\mathrm{Rel}\, x'$ if $x \in \mathrm{ri}\,\mathrm{conv}(\Gamma_{x'})$, which is then completed to an equivalence relation $\sim$. The equivalence classes of the equivalence relation $\sim$ define their notion of irreducible components.

Our subsequent results differ from [10] from two perspectives. First, unlike [10], our decomposition is universal in the sense that it is not relative to any particular martingale measure in $\mathcal{M}(\mu,\nu)$ (see example 3.6). Second, our construction of the irreducible convex paving allows to prove the required measurability property, thus justifying completely the existence of a disintegration of martingale plans.

Finally, during the final stage of writing the present paper, we learned about the parallel work by Jan Obłój and Pietro Siorpaes [18]. Although the results are close, our approach is different from theirs. We are grateful to them for pointing to us to the notions of "convex face" and "Wijsmann topology" and the relative references, which allowed us to streamline our presentation. In an earlier version of this work we used instead a topology that we called the compacted Hausdorff distance, defined as the topology generated by the countable restrictions of the space to the closed balls centered in the origin with integer radia; the two are in our case the same topologies, as the Wijsman topology is locally equivalent to the Hausdorff topology in a locally compact set. We also owe Jan and Pietro special thanks for their useful remarks and comments on a first draft of this paper privately exchanged with them.

The paper is organized as follows. Section 2 collects the main technical ingredients needed for the statement of our main results. In particular, we introduce the new notions of relative face and tangent convex functions, together with the required topology on the set of such functions. Section 3 contains the main results of the paper, namely our decomposition in irreducible convex paving and the structure of polar sets, and shows the identity with the Beiglböck & Juillet [3] notion in the one-dimensional setting. The remaining sections contains the proofs of these results. In particular, the measurability of our irreducible convex paving is proved in Section 7.

Notations. We denote by $\bar{\mathbb{R}}$ the completed real line $\mathbb{R} \cup \{-\infty, \infty\}$, and similarly denote $\overline{\mathbb{R}}_+ := \mathbb{R}_+ \cup \{\infty\}$. We fix an integer $d \geqslant 1$. For $x \in \mathbb{R}^d$ and $r \geqslant 0$, we denote $B_r(x)$ the closed ball for the Euclidean distance, centered in x with radius r. We denote for simplicity $B_r := B_r(0)$. If $x \in \mathcal{X}$, and $A \subset \mathcal{X}$, where $(\mathcal{X}, \mathrm{d})$ is a metric space, $\mathrm{dist}(x, A) := \inf_{a \in A} \mathrm{d}(x, a)$. In all this paper, $\mathbb{R}^d$ is endowed with the Euclidean distance.

If V is a topological affine space and $A \subset V$ is a subset of V, $\mathrm{int} A$ is the interior of A, $\mathrm{cl}\, A$ is the closure of A, $\mathrm{Aff} A$ is the smallest affine subspace of V containing A, $\mathrm{conv} A$ is the convex hull of A, $\dim(A) := \dim(\mathrm{Aff} A)$, and $\mathrm{ri} A$ is the relative interior of A, which is the interior of A in the topology of $\mathrm{Aff} A$ induced by the topology of V. We also denote by $\partial A := \mathrm{cl}\, A \backslash \mathrm{ri} A$ the relative boundary of A, and by λ_A the Lebesgue measure of $\mathrm{Aff} A$.

The set $\mathcal{K}$ of all closed subsets of $\mathbb{R}^d$ is a Polish space when endowed with the Wijsman topology[1] (see Beer [1]). As $\mathbb{R}^d$ is separable, it follows from a theorem of Hess [11] that a function $F : \mathbb{R}^d \longrightarrow \mathcal{K}$ is Borel measurable with respect to the Wijsman topology if and only if its associated multifunction is Borel measurable, i.e.

$$F^-(V) := \{x \in \mathbb{R}^d : F(x) \cap V \neq \varnothing\} \quad \text{is Borel for each open subset } V \subset \mathbb{R}^d.$$

The subset $\widehat{\mathcal{K}} \subset \mathcal{K}$ of all the convex closed subsets of $\mathbb{R}^d$ is closed in $\mathcal{K}$ for the Wijsman topology,

[1] The Wijsman topology on the collection of all closed subsets of a metric space $(\mathcal{X}, \mathrm{d})$ is the weak topology generated by $\{\mathrm{dist}(x, \cdot) : x \in \mathcal{X}\}$.

and therefore inherits its Polish structure. Clearly, $\widehat{\mathcal{K}}$ is isomorphic to $\mathrm{ri}\,\widehat{\mathcal{K}}$ (with reciprocal isomorphism cl). We shall identify these two isomorphic sets in the rest of this text, when there is no possible confusion.

We denote $\Omega := \mathbb{R}^d \times \mathbb{R}^d$ and define the two canonical maps

$$X : (x, y) \in \Omega \longmapsto x \in \mathbb{R}^d \quad \text{and} \quad Y : (x, y) \in \Omega \longmapsto y \in \mathbb{R}^d.$$

For $\varphi, \psi : \mathbb{R}^d \longrightarrow \bar{\mathbb{R}}$, and $h : \mathbb{R}^d \longrightarrow \mathbb{R}^d$, we denote

$$\varphi \oplus \psi := \varphi(X) + \psi(Y), \quad \text{and} \quad h^{\otimes} := h(X) \cdot (Y - X),$$

with the convention $\infty - \infty = \infty$.

For a Polish space $\mathcal{X}$, we denote by $\mathcal{B}(\mathcal{X})$ the collection of Borel subsets of $\mathcal{X}$, and $\mathcal{P}(\mathcal{X})$ the set of all probability measures on $\big(\mathcal{X}, \mathcal{B}(\mathcal{X})\big)$. For $\mathbb{P} \in \mathcal{P}(\mathcal{X})$, we denote by $\mathcal{N}_{\mathbb{P}}$ the collection of all $\mathbb{P}-$null sets, $\mathrm{supp}\mathbb{P}$ the smallest closed support of $\mathbb{P}$, and $\widehat{\mathrm{supp}}\mathbb{P} := \mathrm{cl\,conv\,supp}\mathbb{P}$ the smallest convex closed support of $\mathbb{P}$. For a measurable function $f : \mathcal{X} \to \mathbb{R}$, we use again the convention $\infty - \infty = \infty$ to define its integral, and we denote

$$\mathbb{P}[f] := \mathbb{E}^{\mathbb{P}}[f] = \int_{\mathcal{X}} f d\mathbb{P} = \int_{\mathcal{X}} f(x)\mathbb{P}(dx) \quad \text{for all} \quad \mathbb{P} \in \mathcal{P}(\mathcal{X}).$$

Let $\mathcal{Y}$ be another Polish space, and $\mathbb{P} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$. The corresponding conditional kernel $\mathbb{P}_x$ is defined by:

$$\mathbb{P}(dx, dy) = \mu(dx) \otimes \mathbb{P}_x(dy), \text{ where } \mu := \mathbb{P} \circ X^{-1}.$$

We denote by $\mathbb{L}^0(\mathcal{X}, \mathcal{Y})$ the set of Borel measurable maps from $\mathcal{X}$ to $\mathcal{Y}$. We denote for simplicity $\mathbb{L}^0(\mathcal{X}) := \mathbb{L}^0(\mathcal{X}, \bar{\mathbb{R}})$ and $\mathbb{L}^0_+(\mathcal{X}) := \mathbb{L}^0(\mathcal{X}, \bar{\mathbb{R}}_+)$. For a measure m on $\mathcal{X}$, we denote $\mathbb{L}^1(\mathcal{X}, m) := \{f \in \mathbb{L}^0(\mathcal{X}) : m[|f|] < \infty\}$. We also denote simply $\mathbb{L}^1(m) := \mathbb{L}^1(\bar{\mathbb{R}}, m)$ and $\mathbb{L}^1_+(m) := \mathbb{L}^1_+(\bar{\mathbb{R}}_+, m)$.

We denote by $\mathfrak{C}$ the collection of all finite convex functions $f : \mathbb{R}^d \longrightarrow \mathbb{R}$. We denote by $\partial f(x)$ the corresponding subgradient at any point $x \in \mathbb{R}^d$. We also introduce the collection of all measurable selections in the subgradient, which is nonempty by Lemma 9.2,

$$\underline{\partial} f := \big\{p \in \mathbb{L}^0(\mathbb{R}^d, \mathbb{R}^d) : p(x) \in \partial f(x) \text{ for all } x \in \mathbb{R}^d\big\}.$$

We finally denote $\underline{f}_\infty := \liminf_{n\to\infty} f_n$, for any sequence $(f_n)_{n\geqslant 1}$ of real number, or of real-valued functions.

2 Preliminaries

Throughout this paper, we consider two probability measures μ and ν on $\mathbb{R}^d$ with finite first order moment, and $\mu \preceq \nu$ in the convex order, i.e. $\nu(f) \geqslant \mu(f)$ for all $f \in \mathfrak{C}$. Using the convention $\infty - \infty = \infty$, we may then define $(\nu - \mu)(f) \in [0, \infty]$ for all $f \in \mathfrak{C}$.

We denote by $\mathcal{M}(\mu, \nu)$ the collection of all probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals $\mathbb{P} \circ X^{-1} = \mu$ and $\mathbb{P} \circ Y^{-1} = \nu$. Notice that $\mathcal{M}(\mu, \nu) \neq \varnothing$ by Strassen [20].

An $\mathcal{M}(\mu, \nu)-$polar set is an element of $\cap_{\mathbb{P}\in\mathcal{M}(\mu,\nu)}\mathcal{N}_{\mathbb{P}}$. A property is said to hold $\mathcal{M}(\mu, \nu)-$quasi surely (abbreviated as q.s.) if it holds on the complement of an $\mathcal{M}(\mu, \nu)-$polar set.

2.1 Relative face of a set

For a subset $A \subset \mathbb{R}^d$ and $a \in \mathbb{R}^d$, we introduce the face of A relative to a (also denoted $a-$relative face of A):

$$\mathrm{rf}_a A := \big\{ y \in A : (a - \varepsilon(y-a), y + \varepsilon(y-a)) \subset A, \text{ for some } \varepsilon > 0 \big\} \quad . \tag{2.1}$$

Figure 1 illustrates examples of relative faces of a square S, relative to some points. For later

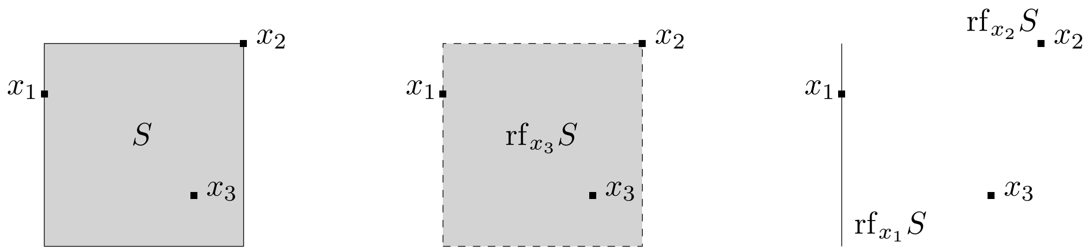

Figure 1: Examples of relative faces.

use, we list some properties whose proofs are reported in Section 9. [2]

Proposition 2.1. (i) *For $A, A' \subset \mathbb{R}^d$, we have $\mathrm{rf}_a(A \cap A') = \mathrm{rf}_a(A) \cap \mathrm{rf}_a(A')$, and $\mathrm{rf}_a A \subset \mathrm{rf}_a A'$ whenever $A \subset A'$. Moreover, $\mathrm{rf}_a A \neq \emptyset$ iff $a \in \mathrm{rf}_a A$ iff $a \in A$.*
(ii) *For a convex A, $\mathrm{rf}_a A = \mathrm{ri} A \neq \emptyset$ iff $a \in \mathrm{ri} A$. Moreover, $\mathrm{rf}_a A$ is convex relatively open, $A \setminus \mathrm{cl}\, \mathrm{rf}_a A$ is convex, and if $x_0 \in A \setminus \mathrm{cl}\, \mathrm{rf}_a A$ and $y_0 \in A$, then $[x_0, y_0) \subset A \setminus \mathrm{cl}\, \mathrm{rf}_a A$. Furthermore, if $a \in A$, then $\dim(\mathrm{rf}_a \mathrm{cl}\, A) = \dim(A)$ if and only if $a \in \mathrm{ri} A$. In this case, we have $\mathrm{cl}\, \mathrm{rf}_a \mathrm{cl}\, A = \mathrm{cl}\, \mathrm{ri}\, \mathrm{cl}\, A = \mathrm{cl}\, A = \mathrm{cl}\, \mathrm{rf}_a A$.*

2.2 Tangent Convex functions

Recall the notation (2.1), and denote for all $\theta : \Omega \to \bar{\mathbb{R}}$:

$$\mathrm{dom}_x \theta \quad := \quad \mathrm{rf}_x \mathrm{conv}\, \mathrm{dom}\, \theta(x, \cdot).$$

For $\theta_1, \theta_2 : \Omega \longrightarrow \mathbb{R}$, we say that $\theta_1 = \theta_2$, $\mu \otimes \mathrm{pw}$, if

$$\mathrm{dom}_X \theta_1 = \mathrm{dom}_X \theta_2, \quad \text{and} \quad \theta_1(X, \cdot) = \theta_2(X, \cdot) \text{ on } \mathrm{dom}_X \theta_1, \; \mu - \text{a.s.}$$

The main ingredient for our extension is the following.

Definition 2.2. *A measurable function $\theta : \Omega \to \overline{\mathbb{R}}_+$ is a tangent convex function if*

$$\theta(x, \cdot) \textit{ is convex, and } 0 = \theta(x, x) = \min_{\mathbb{R}^d} \theta(x, \cdot), \textit{ for all } x \in \mathbb{R}^d.$$

[2] $\mathrm{rf}_a A$ is equal to the only relative interior of face of A containing a, where we extend the notion of face to non-convex sets. A face F of A is a nonempty subset of A such that for all $[a, b] \subset A$, with $(a, b) \cap F \neq \emptyset$, we have $[a, b] \subset F$. It is proved in Hiriart-Urruty-Lemaréchal [13] that when A is convex, the relative interior of the faces of A form a partition of A, see also Rockafellar [19].

We denote by Θ the set of tangent convex functions, and we define

$$\Theta_\mu \quad := \quad \big\{\theta \in \mathbb{L}^0(\Omega, \overline{\mathbb{R}}_+) : \theta = \theta', \ \mu\otimes\text{pw}, \textit{and } \theta \geqslant \theta', \ \textit{for some } \theta' \in \Theta\big\}.$$

In order to introduce our main example of such functions, let

$$\mathbf{T}_p f(x,y) := f(y) - f(x) - p^{\otimes}(x,y) \geqslant 0, \quad \text{for all} \quad f \in \mathfrak{C}, \ \text{and } p \in \partial f.$$

Then, $\mathbf{T}(\mathfrak{C}) := \{\mathbf{T}_p f : f \in \mathfrak{C}, p \in \partial f\} \subset \Theta \subset \Theta_\mu$.

Example 2.3. *The second inclusion is strict. Indeed, let $d = 1$, and consider the convex function $f := \infty \mathbf{1}_{(-\infty,0)}$. Then $\theta := f(Y-X) \in \Theta$. Now let $\theta' = \theta + \sqrt{|Y-X|}$. Notice that since $\text{dom}_X\theta = \text{dom}_X\theta' = \{X\}$, we have $\theta = \theta'$, $\mu\otimes\text{pw}$ for any measure μ, and $\theta' \geqslant \theta$. Therefore $\theta' \in \Theta_\mu$. However, for all $x \in \mathbb{R}^d$, $\theta'(x,\cdot)$ is not convex, and therefore $\theta' \notin \Theta$.*

In higher dimension we may even have $X \in \text{ri}\,\text{dom}\theta'(X,\cdot)$, and $\theta'(X,\cdot)$ is not convex. Indeed, for $d = 2$, let $f : (y_1,y_2) \longmapsto \infty(\mathbf{1}_{\{|y_1|>1\}} + \mathbf{1}_{\{|y_2|>1\}})$, so that $\theta := f(Y-X) \in \Theta$. Let $x_0 := (1,0)$ and $\theta' := \theta + \mathbf{1}_{\{Y=X+x_0\}}$. Then, $\theta = \theta'$, $\mu\otimes\text{pw}$ for any measure μ, and $\theta' \geqslant \theta$. Therefore $\theta' \in \Theta_\mu$. However, $\theta' \notin \Theta$ as $\theta'(x,\cdot)$ is not convex for all $x \in \mathbb{R}^d$.

Proposition 2.4. (i) *Let $\theta \in \Theta_\mu$, $\text{dom}_X\theta = \text{rf}_X\text{dom}\theta(X,\cdot) \subset \text{dom}\theta(X,\cdot)$, $\mu-$a.s.*
(ii) *Let $\theta_1, \theta_2 \in \Theta_\mu$, $\text{dom}_X(\theta_1+\theta_2) = \text{dom}_X\theta_1 \cap \text{dom}_X\theta_2$, $\mu-$a.s.*
(iii) *Θ_μ is a convex cone.*

Proof. (i) It follows immediately from the fact that $\theta(X,\cdot)$ is convex and finite on $\text{dom}_X\theta$, $\mu-$a.s. by definition of Θ_μ. Then $\text{dom}_X\theta \subset \text{rf}_X\text{dom}\theta(X,\cdot)$. On the other side, as $\text{dom}\theta(X,\cdot) \subset \text{conv}\,\text{dom}\theta(X,\cdot)$, the monotony of rf_x gives the other inclusion: $\text{rf}_X\text{dom}\theta(X,\cdot) \subset \text{dom}_X\theta$
(ii) As $\theta_1, \theta_2 \geqslant 0$, $\text{dom}(\theta_1+\theta_2) = \text{dom}\theta_1 \cap \text{dom}\theta_2$. Then, for $x \in \mathbb{R}^d$, $\text{conv}\,\text{dom}(\theta_1(x,\cdot)+\theta_2(x,\cdot)) \subset \text{conv}\,\text{dom}\theta_1(x,\cdot) \cap \text{conv}\,\text{dom}\theta_2(x,\cdot)$. By Proposition 2.1 (i),

$$\text{dom}_x(\theta_1+\theta_2) \subset \text{dom}_x\theta_1 \cap \text{dom}_x\theta_2, \quad \text{for all} \quad x \in \mathbb{R}^d.$$

As for the reverse inclusion, notice that (i) implies that $\text{dom}_X\theta_1 \cap \text{dom}_X\theta_2 \subset \text{dom}\theta_1(X,\cdot) \cap \text{dom}\theta_2(X,\cdot) = \text{dom}\big(\theta_1(X,\cdot)+\theta_2(X,\cdot)\big) \subset \text{conv}\,\text{dom}\big(\theta_1(X,\cdot)+\theta_2(X,\cdot)\big)$, $\mu-$a.s. Observe that $\text{dom}_x\theta_1 \cap \text{dom}_x\theta_2$ is convex, relatively open, and contains x. Then,

$$\begin{aligned}\text{dom}_X\theta_1 \cap \text{dom}_X\theta_2 \;=\; \text{rf}_X\big(\text{dom}_X\theta_1 \cap \text{dom}_X\theta_2\big) \;&\subset\; \text{rf}_X\Big(\text{conv}\,\text{dom}\big(\theta_1(X,\cdot)+\theta_2(X,\cdot)\big)\Big)\\ &=\; \text{dom}_X(\theta_1+\theta_2) \ \ \mu-\text{a.s.}\end{aligned}$$

(iii) Given (ii), this follows from direct verification. □

Definition 2.5. *A sequence $(\theta_n)_{n\geqslant 1} \subset \mathbb{L}^0(\Omega)$ converges $\mu\otimes\text{pw}$ to some $\theta \in \mathbb{L}^0(\Omega)$ if*

$$\text{dom}_X\left(\underline{\theta}_\infty\right) = \text{dom}_X\theta \quad \textit{and} \quad \theta_n(X,\cdot) \longrightarrow \theta(X,\cdot), \textit{ pointwise on } \text{dom}_X\theta, \ \mu-a.s.$$

Notice that the $\mu\otimes$pw-limit is $\mu\otimes$pw unique. In particular, if θ_n converges to θ, $\mu\otimes$pw, it converges as well to $\underline{\theta}_\infty$.

Proposition 2.6. *Let* $(\theta_n)_{n\geqslant 1} \subset \Theta_\mu$, *and* $\theta : \Omega \longrightarrow \bar{\mathbb{R}}_+$, *such that* $\theta_n \underset{n\to\infty}{\longrightarrow} \theta$, $\mu\otimes\mathrm{pw}$,
(i) $\mathrm{dom}_X\theta \subset \liminf_{n\to\infty} \mathrm{dom}_X\theta_n$, *$\mu-$a.s.*
(ii) *If* $\theta'_n = \theta_n$, $\mu\otimes\mathrm{pw}$, *and* $\theta'_n \geqslant \theta_n$, *then* $\theta'_n \underset{n\to\infty}{\longrightarrow} \theta$, $\mu\otimes\mathrm{pw}$*;*
(iii) $\underline{\theta}_\infty \in \Theta_\mu$.

Proof. (i) Let $x \in \mathbb{R}^d$, such that $\theta_n(x,\cdot)$ converges on $\mathrm{dom}_x\theta$ to $\theta(x,\cdot)$. Let $y \in \mathrm{dom}_x\theta$, let $y' \in \mathrm{dom}_x\theta$ such that $y' = x - \epsilon(y-x)$, for some $\epsilon > 0$. As $\theta_n(x,y) \underset{n\to\infty}{\longrightarrow} \theta(x,y)$, and $\theta_n(x,y') \underset{n\to\infty}{\longrightarrow} \theta(x,y')$, then for n large enough, both are finite, and $y \in \mathrm{dom}_x\theta_n$. $y \in \liminf_{n\to\infty}\mathrm{dom}_x\theta_n$, and $\mathrm{dom}_x\theta \subset \liminf_{n\to\infty}\mathrm{dom}_x\theta_n$. The inclusion is true for $\mu-$a.e. $x \in \mathbb{R}^d$, which gives the result.
(ii) By (i), we have $dom_X\theta \subset \liminf_{n\to\infty}\mathrm{dom}_X\theta_n = \liminf_{n\to\infty}\mathrm{dom}_X\theta'_n$, $\mu-$a.s. As $\theta_n \leqslant \theta'_n$, $dom_X\underline{\theta'}_\infty \subset \mathrm{dom}_X\underline{\theta}_\infty \subset \liminf_{n\to\infty}\mathrm{dom}_X\theta_n$, $\mu-$a.s. We denote $N_\mu \in \mathcal{N}_\mu$, the set on which $\theta_n(X,\cdot)$ does not converge to $\theta(X,\cdot)$ on $\mathrm{dom}_X\theta(X,\cdot)$. For $x \notin N_\mu$, for $y \in \mathrm{dom}_x\theta$, $\theta_n(x,y) = \theta'_n(x,y)$, for n large enough, and $\theta'_n(x,y) \underset{n\to\infty}{\longrightarrow} \theta(x,y) < \infty$. Then $\mathrm{dom}_X\theta = \mathrm{dom}_X\underline{\theta'}_\infty$, and $\theta'_n(X,\cdot)$ converges to $\theta(X,\cdot)$, on $\mathrm{dom}_X\theta$, $\mu-$a.s. We proved that $\theta'_n \underset{n\to\infty}{\longrightarrow} \theta$, $\mu\otimes\mathrm{pw}$.
(iii) is proved in Subsection 8.2. □

The next result shows the relevance of this notion of convergence for our setting.

Proposition 2.7. *Let* $(\theta_n)_{n\geqslant 1} \subset \Theta_\mu$. *Then, we may find a sequence* $\widehat{\theta}_n \in \mathrm{conv}(\theta_k, k \geqslant n)$, *and* $\widehat{\theta}_\infty \in \Theta_\mu$ *such that* $\widehat{\theta}_n \longrightarrow \widehat{\theta}_\infty$, $\mu\otimes\mathrm{pw}$ *as* $n \to \infty$.

The proof is reported in Subsection 8.2.

Definition 2.8. (i) *A subset* $\mathcal{T} \subset \Theta_\mu$ *is* $\mu\otimes\mathrm{pw}$*-Fatou closed if* $\underline{\theta}_\infty \in \mathcal{T}$ *for all* $(\theta_n)_{n\geqslant 1} \subset \mathcal{T}$ *converging* $\mu\otimes\mathrm{pw}$ *(in particular,* Θ_μ *is* $\mu\otimes\mathrm{pw}-$*Fatou closed by Proposition 2.6* (iii)*).*
(ii) *The* $\mu\otimes\mathrm{pw}-$*Fatou closure of a subset* $A \subset \Theta_\mu$ *is the smallest* $\mu\otimes\mathrm{pw}-$*Fatou closed set containing A:*

$$\widehat{A} := \bigcap \big\{\mathcal{T} \subset \Theta_\mu : \ A \subset \mathcal{T}, \ \text{and } \mathcal{T} \ \mu\otimes\mathrm{pw}\text{-Fatou closed}\big\}.$$

We next introduce for $a \geqslant 0$ the set $\mathfrak{C}_a := \big\{f \in \mathfrak{C} : (\nu-\mu)(f) \leqslant a\big\}$, and

$$\widehat{\mathcal{T}}(\mu,\nu) := \bigcup_{a\geqslant 0} \widehat{\mathcal{T}}_a, \ \text{where } \widehat{\mathcal{T}}_a := \widehat{\mathbf{T}(\mathfrak{C}_a)}, \ \text{and } \mathbf{T}\big(\mathfrak{C}_a\big) := \big\{\mathbf{T}_p f : f \in \mathfrak{C}_a, p \in \partial f\big\}.$$

Proposition 2.9. $\widehat{\mathcal{T}}(\mu,\nu)$ *is a convex cone.*

Proof. We first prove that $\widehat{\mathcal{T}}(\mu,\nu)$ is a cone. We consider $\lambda, a > 0$, as we have $\lambda\mathfrak{C}_a = \mathfrak{C}_{\lambda a}$, and as convex combinations and inferior limit commute with the multiplication by λ, we have $\lambda\widehat{\mathcal{T}}_a = \widehat{\mathcal{T}}_{\lambda a}$. Then $\widehat{\mathcal{T}}(\mu,\nu) = \mathrm{cone}(\widehat{\mathcal{T}}_1)$, and therefore it is a cone.

We next prove that $\widehat{\mathcal{T}}_a$ is convex for all $a \geqslant 0$, which induces the required convexity of $\widehat{\mathcal{T}}(\mu,\nu)$ by the non-decrease of the family $\{\widehat{\mathcal{T}}_a, a \geqslant 0\}$. Fix $0 \leqslant \lambda \leqslant 1$, $a \geqslant 0$, $\theta_0 \in \widehat{\mathcal{T}}_a$, and denote $\mathcal{T}(\theta_0) := \big\{\theta \in \widehat{\mathcal{T}}_a : \lambda\theta_0 + (1-\lambda)\theta \in \widehat{\mathcal{T}}_a\big\}$. In order to complete the proof, we now verify that $\mathcal{T}(\theta_0) \supset \mathbf{T}\big(\mathfrak{C}_a\big)$ and is $\mu\otimes\mathrm{pw}-$Fatou closed, so that $\mathcal{T}(\theta_0) = \widehat{\mathcal{T}}_a$.

To see that $\mathcal{T}(\theta_0)$ is Fatou-closed, let $(\theta_n)_{n\geqslant 1} \subset \mathcal{T}(\theta_0)$, converging $\mu\otimes$pw. By definition of $\mathcal{T}(\theta_0)$, we have $\lambda\theta_0+(1-\lambda)\theta_n \in \widehat{\mathcal{T}}_a$ for all n. Then, $\lambda\theta_0+(1-\lambda)\theta_n \longrightarrow \liminf_{n\to\infty} \lambda\theta_0+(1-\lambda)\underline{\theta}_n$, $\mu\otimes$pw, and therefore $\lambda\theta_0 + (1-\lambda)\underline{\theta}_\infty \in \widehat{\mathcal{T}}_a$, which shows that $\underline{\theta}_\infty \in \mathcal{T}(\theta_0)$.

We finally verify that $\mathcal{T}(\theta_0) \supset \mathbf{T}\big(\mathfrak{C}_a\big)$. First, for $\theta_0 \in \mathbf{T}\big(\mathfrak{C}_a\big)$, this inclusion follows directly from the convexity of $\mathbf{T}\big(\mathfrak{C}_a\big)$, implying that $\mathcal{T}(\theta_0) = \widehat{\mathcal{T}}_a$ in this case. For general $\theta_0 \in \widehat{\mathcal{T}}_a$, the last equality implies that $\mathbf{T}\big(\mathfrak{C}_a\big) \subset \mathcal{T}(\theta_0)$, thus completing the proof. $\square$

Notice that even though $\mathbf{T}(\mathfrak{C}_a) \subset \Theta$, the functions in $\widehat{\mathcal{T}}(\mu,\nu)$ may not be in Θ as they may not be convex in y on $(\mathrm{dom}_x\theta)^c$ for some $x \in \mathbb{R}^d$. The following result shows that some convexity is still preserved.

Proposition 2.10. *For all $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$, we may find $N_\mu \in \mathcal{N}_\mu$ such that for $x_1, x_2 \notin N_\mu$, $y_1, y_2 \in \mathbb{R}^d$, and $\lambda \in [0,1]$ with $\bar{y} := \lambda y_1 + (1-\lambda)y_2 \in \mathrm{dom}_{x_1}\theta \cap \mathrm{dom}_{x_2}\theta$, we have:*

$$\lambda\theta(x_1,y_1) + (1-\lambda)\theta(x_1,y_1) - \theta(x_1,\bar{y}) = \lambda\theta(x_2,y_1) + (1-\lambda)\theta(x_2,y_1) - \theta(x_2,\bar{y}) \geqslant 0.$$

The proof of this claim is reported in Subsection 8.1. We observe that the statement also holds true for a finite number of points $y_1, ..., y_k$.[3]

2.3 Extended integral

We now introduce the extended $(\nu-\mu)-$integral:

$$\nu\widehat{\ominus}\mu[\theta] := \inf\big\{a \geqslant 0 : \theta \in \widehat{\mathcal{T}}_a\big\} \quad \text{for} \quad \theta \in \widehat{\mathcal{T}}(\mu,\nu).$$

Proposition 2.11. (i) *$\mathbb{P}[\theta] \leqslant \nu\widehat{\ominus}\mu[\theta] < \infty$ for all $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$ and $\mathbb{P} \in \mathcal{M}(\mu,\nu)$.*
(ii) *$\nu\widehat{\ominus}\mu[\mathbf{T}_p f] = (\nu-\mu)[f]$ for $f \in \mathfrak{C} \cap \mathbb{L}^1(\nu)$ and $p \in \partial f$.*
(iii) *$\nu\widehat{\ominus}\mu$ is homogeneous and convex.*

Proof. (i) For $a > \nu\widehat{\ominus}\mu[\theta]$, set $S^a := \big\{F \in \Theta_\mu : \mathbb{P}[F] \leqslant a \text{ for all } \mathbb{P} \in \mathcal{M}(\mu,\nu)\big\}$. Notice that S^a is $\mu\otimes$pw$-$Fatou closed by Fatou's lemma, and contains $\mathbf{T}(\mathfrak{C}_a)$, as for $f \in \mathfrak{C} \cap \mathbb{L}^1(\nu)$ and $p \in \partial f$, $\mathbb{P}[T_p f] = (\nu-\mu)[f]$ for all $\mathbb{P} \in \mathcal{M}(\mu,\nu)$. Then S^a contains $\widehat{\mathcal{T}}_a$ as well, which contains θ. Hence, $\theta \in S^a$ and $\mathbb{P}[\theta] \leqslant a$ for all $\mathbb{P} \in \mathcal{M}(\mu,\nu)$. The required result follows from the arbitrariness of $a > \nu\widehat{\ominus}\mu[\theta]$.
(ii) Let $\mathbb{P} \in \mathcal{M}(\mu,\nu)$. For $p \in \partial f$, notice that $T_p f \in \mathbf{T}(\mathfrak{C}_a) \subset \widehat{\mathcal{T}_a}$ for some $a = (\nu-\mu)[f]$, and therefore $(\nu-\mu)[f] \geqslant \nu\widehat{\ominus}\mu[T_p f]$. Then, the result follows from the inequality $(\nu-\mu)[f] = \mathbb{P}[T_p f] \leqslant \nu\widehat{\ominus}\mu[T_p f]$.
(iii) Similarly to the proof of Proposition 2.9, we have $\lambda\widehat{\mathcal{T}}_a = \widehat{\mathcal{T}}_{\lambda a}$, for all $\lambda, a > 0$. Then with the definition of $\nu\widehat{\ominus}\mu$ we have easily the homogeneity.

To see that the convexity holds, let $0 < \lambda < 1$, and $\theta, \theta' \in \widehat{\mathcal{T}}(\mu,\nu)$ with $a > \nu\widehat{\ominus}\mu[\theta]$, $a' > \nu\widehat{\ominus}\mu[\theta']$, for some $a, a' > 0$. By homogeneity and convexity of $\widehat{\mathcal{T}}_1$, $\lambda\theta + (1-\lambda)\theta' \in \widehat{\mathcal{T}}_{\lambda a+(1-\lambda)a'}$, so that $\nu\widehat{\ominus}\mu[\lambda\theta + (1-\lambda)\theta'] \leqslant \lambda a + (1-\lambda)a'$. The required convexity property now follows

[3] This is not a direct consequence of Proposition 2.10, as the barycentre $\bar{y}$ has to be in $\mathrm{dom}_{x_1}\theta \cap \mathrm{dom}_{x_2}\theta$.

from arbitrariness of $a > \nu\widehat{\ominus}\mu[\theta]$ and $a' > \nu\widehat{\ominus}\mu[\theta']$. □

The following compacteness result plays a crucial role.

Lemma 2.12. *Let $(\theta_n)_{n\geqslant 1} \subset \widehat{\mathcal{T}}(\mu,\nu)$ be such that $\sup_{n\geqslant 1} \nu\widehat{\ominus}\mu(\theta_n) < \infty$. Then we can find a sequence $\widehat{\theta}_n \in \mathrm{conv}(\theta_k, k \geqslant n)$ such that*

$$\underline{\widehat{\theta}}_\infty \in \widehat{\mathcal{T}}(\mu,\nu), \quad \widehat{\theta}_n \longrightarrow \underline{\widehat{\theta}}_\infty, \ \mu\otimes\text{pw}, \quad \textit{and} \ \ \nu\widehat{\ominus}\mu(\underline{\widehat{\theta}}_\infty) \leqslant \liminf_{n\to\infty} \nu\widehat{\ominus}\mu(\theta_n).$$

Proof. By possibly passing to a subsequence, we may assume that $\lim_{n\to\infty}(\nu\widehat{\ominus}\mu)(\theta_n)$ exists. The boundedness of $\nu\widehat{\ominus}\mu(\theta_n)$ ensures that this limit is finite. We next introduce the sequence $\widehat{\theta}_n$ of Proposition 2.7. Then $\widehat{\theta}_n \longrightarrow \widehat{\theta}_\infty$, $\mu\otimes\text{pw}$, and therefore $\underline{\widehat{\theta}}_\infty \in \widehat{\mathcal{T}}(\mu,\nu)$, because of the convergence $\widehat{\theta}_n \longrightarrow \underline{\widehat{\theta}}_\infty$, $\mu\otimes$pw. As $(\nu\widehat{\ominus}\mu)(\widehat{\theta}_n) \leqslant \sup_{k\geqslant n}(\nu\widehat{\ominus}\mu)(\theta_k)$ by Proposition 2.11 (iii), we have $\infty > \lim_{n\to\infty}(\nu\widehat{\ominus}\mu)(\theta_n) = \lim_{n\to\infty}\sup_{k\geqslant n}(\nu\widehat{\ominus}\mu)(\theta_k) \geqslant \limsup_{n\to\infty}(\nu\widehat{\ominus}\mu)(\widehat{\theta}_n)$. Set $l := \limsup_{n\to\infty} \nu\widehat{\ominus}\mu(\widehat{\theta}_n)$. For $\epsilon > 0$, we consider $n_0 \in \mathbb{N}$ such that $\sup_{k\geqslant n_0} \nu\widehat{\ominus}\mu(\widehat{\theta}_k) \leqslant l+\epsilon$. Then for $k \geqslant n_0$, $\widehat{\theta}_k \in \widehat{\mathcal{T}}_{l+2\epsilon}(\mu,\nu)$, and therefore $\underline{\widehat{\theta}}_\infty = \liminf_{k\geqslant n_0} \widehat{\theta}_k \in \widehat{\mathcal{T}}_{l+2\epsilon}(\mu,\nu)$, implying $\nu\widehat{\ominus}\mu(\widehat{\theta}) \leqslant l + 2\epsilon \longrightarrow l$, as $\epsilon \to 0$. Finally, $\liminf_{n\to\infty}(\nu\widehat{\ominus}\mu)(\theta_n) \geqslant \nu\widehat{\ominus}\mu(\underline{\widehat{\theta}}_\infty)$. □

3 Main results

3.1 The irreducible convex paving

Our final ingredient is the following measurement of subsets $K \subset \mathbb{R}^d$:

$$G(K) := \dim(K) + g_K(K) \quad \text{where} \quad g_K(dx) := (2\pi)^{-\frac{1}{2}\dim K} e^{-\frac{1}{2}|x|^2}\lambda_K(dx),$$

Notice that $0 \leqslant G \leqslant d+1$ and, for any convex subsets $C_1 \subset C_2$ of $\mathbb{R}^d$, we have

$$G(C_1) = G(C_2) \quad \text{iff} \quad \mathrm{ri}C_1 = \mathrm{ri}C_2 \quad \text{iff} \quad \mathrm{cl}\,C_1 = \mathrm{cl}\,C_2. \tag{3.2}$$

For $\theta \in \mathbb{L}^0_+(\Omega)$, $A \in \mathcal{B}(\mathbb{R}^d)$, we introduce the following map from $\mathbb{R}^d$ to the set $\widehat{\mathcal{K}}$ of all relatively open convex subsets of $\mathbb{R}^d$:

$$K_{\theta,A}(x) := \mathrm{rf}_x\mathrm{conv}(\mathrm{dom}\theta(x,\cdot)\backslash A) = \mathrm{dom}_X(\theta + \infty 1_{\mathbb{R}^d\times A}), \quad \text{for all} \quad x \in \mathbb{R}^d. \tag{3.3}$$

We recall that a function is universally measurable if it is measurable with respect to every complete probability measure that measures all Borel subsets.

Lemma 3.1. *For $\theta \in \mathbb{L}^0_+(\Omega)$ and $A \in \mathcal{B}(\mathbb{R}^d)$, we have:*
(i) *$\mathrm{cl\,conv\,dom}\theta(X,\cdot)$, $\mathrm{dom}_X\theta$, and $K_{\theta,A}$ are universally measurable;*
(ii) *$G : \widehat{\mathcal{K}} \longrightarrow \mathbb{R}$ is Borel-measurable;*
(iii) *if $A \in \mathcal{N}_\nu$, and $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$, then up to a modification on a $\mu-$null set, $K_{\theta,A}(\mathbb{R}^d) \subset \mathring{\mathcal{K}}$ is a partition of $\mathbb{R}^d$ with $x \in K_{\theta,A}(x)$ for all $x \in \mathbb{R}^d$.*

The proof is reported in Subsections 4.2, 7.1 and 7.2. The following property is the key-ingredient for our decomposition in irreducible convex paving.

Proposition 3.2. *For all* $(\theta, N_\nu) \in \widehat{\mathcal{T}}(\mu,\nu) \times \mathcal{N}_\nu$*, we have* $Y \in \mathrm{cl}\, K_{\theta,N_\nu}(X)$*,* $\mathcal{M}(\mu,\nu)-$*q.s.*

Proof. For an arbitrary $\mathbb{P} \in \mathcal{M}(\mu,\nu)$, we have by Proposition 2.11 that $\mathbb{P}[\theta] < \infty$. Then, $\mathbb{P}\left[\mathrm{dom}\theta \backslash (\mathbb{R}^d \times N_\nu)\right] = 1$ i.e. $\mathbb{P}[Y \in D_X] = 1$ where $D_x := \mathrm{conv}(\mathrm{dom}\theta(x,\cdot)\backslash N_\nu)$. By the martingale property of $\mathbb{P}$, we deduce that

$$X \;=\; \mathbb{E}^{\mathbb{P}}[Y \mathbf{1}_{Y \in D_X}|X] \;=\; (1-\Lambda)E_I + \Lambda E_D, \quad \mu-\text{a.s.}$$

Where $\Lambda := \mathbb{P}_X[Y \in D_X \backslash \mathrm{cl}\, I(X)]$, $E_D := \mathbb{E}^{P_X}[Y|Y \in D_X \backslash \mathrm{cl}\, I(X)]$, $E_I := \mathbb{E}^{\mathbb{P}_X}[Y|Y \in \mathrm{cl}\, I(X)]$, and $\mathbb{P}_X$ is the conditional kernel to X of $\mathbb{P}$. We have $E_I \in \mathrm{cl}\,\mathrm{rf}_X D_X \subset D_X$ and $E_D \in D_X \backslash \mathrm{cl}\,\mathrm{rf}_X D_X$ because of the convexity of $D_X \backslash \mathrm{cl}\,\mathrm{rf}_X D_X$ given by Proposition 2.1 (ii) (D_X is convex). The lemma also gives that if $\Lambda \neq 0$, then $\mathbb{E}^{\mathbb{P}}[Y|X] = \Lambda E_D + (1-\Lambda)E_I \in D_X \backslash \mathrm{cl}\, I(X)$. This implies that

$$\{\Lambda \neq 0\} \;\subset\; \{\mathbb{E}^{\mathbb{P}}[Y|X] \in D_X \backslash \mathrm{cl}\, I(X)\} \;\subset\; \{\mathbb{E}^{\mathbb{P}}[Y|X] \notin I(X)\} \;\subset\; \{\mathbb{E}^{\mathbb{P}}[Y|X] \neq X\}.$$

Then $\mathbb{P}[\Lambda \neq 0] = 0$, and therefore $\mathbb{P}\left[Y \in D_X \backslash \mathrm{cl}\, I(X)\right] = 0$. Since $\mathbb{P}[Y \in D_X] = 1$, this shows that $\mathbb{P}[Y \in \mathrm{cl}\, I(X)] = 1$. □

In view of Proposition 3.2 and Lemma 3.1 (iii), we introduce the following optimization problem which will generate our irreducible convex paving decomposition:

$$\inf_{(\theta,N_\nu) \in \widehat{\mathcal{T}}(\mu,\nu) \times \mathcal{N}_\nu} \mu[G(K_{\theta,N_\nu})]. \tag{3.4}$$

Our first main result is the following.

Theorem 3.3. (i) *There is a* μ*-a.s. unique universally measurable minimizer* $I := K_{\widehat{\theta},\widehat{N}_\nu} : \mathbb{R}^d \to \widehat{\mathcal{K}}$ *of* (3.4)*, for some* $(\widehat{\theta}, \widehat{N}_\nu) \in \widehat{\mathcal{T}}(\mu,\nu) \times \mathcal{N}_\nu$*.*
I is called irreducible convex paving map, and satisfies $Y \in \mathrm{cl}\, I(X)$*,* $\mathcal{M}(\mu,\nu)-$*q.s.*
(ii) *for all* $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$ *and* $N_\nu \in \mathcal{N}_\nu$*, we have* $I(X) \subset K_{\theta,N_\nu}(X)$*,* μ*-a.s.*
(iii) *up to a modification on a* $\mu-$*null set,* $I(x)$ *is a relatively open convex subset of* $\mathbb{R}^d$ *with* $x \in I(x)$*, for all* $x \in \mathbb{R}^d$*, and* $\left\{I(x), x \in \mathbb{R}^d\right\}$ *is a partition of* $\mathbb{R}^d$*.*

In item (i), the measurability of I is induced by Lemma 3.1 (i), while the fact that $Y \in I(X)$, $\mathcal{M}(\mu,\nu)-$q.s. is a consequence of Proposition 3.2. Existence and uniqueness, together with (ii), are proved in Subsection 4.1. (iii) is implied by Lemma 3.1 (ii).

The next result shows the existence of a maximum support martingale transport plan, i.e. a martingale interpolating measure $\widehat{\mathbb{P}}$ whose disintegration $\widehat{\mathbb{P}}_x$ has a maximum convex hull of supports among all measures in $\mathcal{M}(\mu,\nu)$. For a probability measure $\mathbb{P}$ on a topological space, and a Borel subset A, $\mathbb{P}|_A := \mathbb{P}[\cdot \cap A]$ denotes its restriction to A.

Proposition 3.4. *There exists $\widehat{\mathbb{P}} \in \mathcal{M}(\mu,\nu)$ such that for all $\mathbb{P} \in \mathcal{M}(\mu,\nu)$:*
(i) $\widehat{\text{supp}}\,\mathbb{P}_X \subset \widehat{\text{supp}}\,\widehat{\mathbb{P}}_X = \text{cl}\,I(X)$, *$\mu-$a.s.,*
(ii) $J(X) := I(X) \cup \widehat{\text{supp}}\,\widehat{\mathbb{P}}_X|_{\partial I(X)}$ *is convex,*

$$Y \in J(X),\ \mathcal{M}(\mu,\nu) - q.s. \quad and \quad \text{supp}\,\mathbb{P}_X|_{\partial I(X)} \subset J \backslash I(X),\ \mu - a.s.$$

(iii) *J is unique $\mu-$a.s. and, up to a modification on a $\mu-$null set, Theorem 3.3 (iii) is preserved, we have $I \subset J \subset \text{cl}\,I$, and J is constant on $I(x)$, for all $x \in \mathbb{R}^d$.*

The proof is reported in Subsection 6.3, and is a consequence of Theorem 3.7 below. Proposition 3.4 provides a characterization of the irreducible convex paving by means of an optimality criterion on $\mathcal{M}(\mu,\nu)$. In particular, Example 3.6 uses this characterization to determine the irreducible convex paving in simple concrete cases.

Remark 3.5. We may chose the measure $\widehat{\mathbb{P}} \in \mathcal{M}(\mu,\nu)$ of Proposition 3.4 so that

$$\mu\big[\mathbb{P}_X[\{y\}] > 0\big] \leqslant \mu\big[\widehat{\mathbb{P}}_X[\{y\}] > 0\big] \quad \text{for all} \quad \mathbb{P} \in \mathcal{M}(\mu,\nu) \quad \text{and} \quad y \in \mathbb{R}^d.$$

Let $\underline{J}(X) := I(X) \cup \{y \in \mathbb{R}^d : \widehat{\mathbb{P}}_X[\{y\}] > 0\}$, and $J_\theta(X) := \text{dom}\,\theta(X,\cdot) \cap J(X)$, for some $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$. Then, up to modification on a $\mu-$null set preserving Proposition 3.4 (iii),

$$\begin{aligned} I \subset \underline{J} \subset J_\theta \subset J \subset \text{cl}\,I, \quad Y \in J_\theta(X),\ \mathcal{M}(\mu,\nu) - \text{q.s.} \\ \text{and} \quad I, J, \underline{J}, J_\theta \text{ constant on } I(x), \quad \text{for all} \quad x \in \mathbb{R}^d. \end{aligned} \tag{3.5}$$

These claims are justified in Subsection 6.4.

Example 3.6. In $\mathbb{R}^2$, we introduce $x_0 := (0,0)$, $x_1 := (1,0)$, $y_0 := x_0$, $y_{-1} := (0,-1)$, $y_1 := (0,1)$, and $y_2 := (2,0)$. Then we set $\mu := \frac{1}{2}(\delta_{x_0} + \delta_{x_1})$ and $\nu := \frac{1}{8}(4\delta_{y_0} + \delta_{y_{-1}} + \delta_{y_1} + 2\delta_{y_2})$. We can show easily that $\mathcal{M}(\mu,\nu)$ is the nonempty convex hull of $\mathbb{P}_1$ and $\mathbb{P}_2$ where

$$\mathbb{P}_1 := \frac{1}{8}\big(4\delta_{x_0,y_0} + 2\delta_{x_1,y_2} + \delta_{x_1,y_1} + \delta_{x_1,y_{-1}}\big)$$

and

$$\mathbb{P}_2 := \frac{1}{8}\big(2\delta_{x_0,y_0} + \delta_{x_0,y_1} + \delta_{x_0,y_{-1}} + 2\delta_{x_1,y_0} + 2\delta_{x_1,y_2}\big)$$

(i) *The Ghoussoub-Kim-Lim [10] irreducible convex paving.* Let $c_1 = \mathbf{1}_{X=Y}$, $c_2 = 1 - c_1 = \mathbf{1}_{X\neq Y}$, and notice that $\mathbb{P}_i$ is the unique optimal martingale transport plan for c_i, $i = 1,2$. Then, it follows that the corresponding $\mathbb{P}_i-$irreducible convex paving according to the definition of [10] are given by

$$\begin{aligned} C_{\mathbb{P}_1}(x_0) = \{x_0\},\ C_{\mathbb{P}_1}(x_1) = \text{ri}\,\text{conv}\{y_1, y_{-1}, y_2\}, \\ \text{and } C_{\mathbb{P}_2}(x_0) = \text{ri}\,\text{conv}\{y_1, y_{-1}\},\ C_{\mathbb{P}_2}(x_1) = \text{ri}\,\text{conv}\{y_0, y_2\}. \end{aligned}$$

Figure 2 shows the extreme probabilities $\mathbb{P}_1$ and $\mathbb{P}_2$, and their associated irreducible convex pavings map $C_{\mathbb{P}_1}$ and $C_{\mathbb{P}_2}$.

(ii) *Our irreducible convex paving.* The irreducible components are given by

$$I(x_0) = \text{ri}\,\text{conv}(y_1, y_{-1}) \quad \text{and} \quad I(x_1) = \text{ri}\,\text{conv}(y_1, y_{-1}, y_2).$$

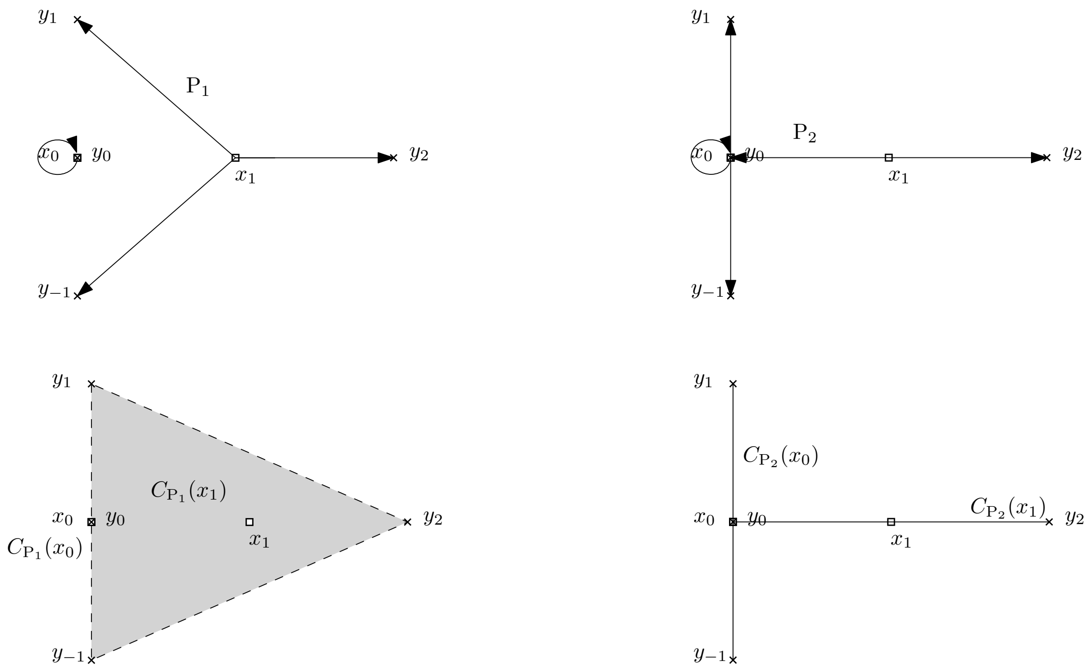

Figure 2: The extreme probabilities and associated irreducible paving.

To see this, we use the characterization of Proposition 3.4. Indeed, as $\mathcal{M}(\mu,\nu) = \mathrm{conv}(\mathbb{P}_1,\mathbb{P}_2)$, for any $\mathbb{P} \in \mathcal{M}(\mu,\nu)$, $\mathbb{P} \ll \widehat{\mathbb{P}} := \frac{\mathbb{P}_1+\mathbb{P}_2}{2}$, and $\mathrm{supp}\,\mathbb{P}_x \subset \mathrm{conv}\big(\mathrm{supp}\,\widehat{\mathbb{P}}_x\big)$ for $x = x_0, x_1$. Then $I(x) = \mathrm{riconv}\big(\mathrm{supp}\,\widehat{\mathbb{P}}_x\big)$ for $x = x_0, x_1$ (i.e. μ-a.s.) by Proposition 3.4.

Recall that the definition of the mapping I is given by a minimization problem on $\widehat{\mathcal{T}}(\mu,\nu) \times \mathcal{N}_\nu$. A $\widehat{\mathcal{T}}(\mu,\nu)$ function minimizing this problem (with $N_\nu := \varnothing \in \mathcal{N}_\nu$) is $\widehat{\theta} := \liminf_{n\to\infty} \mathbf{T}_{p_n} f_n$, where $f_n := nf$, $p_n := np$ for some $p \in \partial f$, and

$$f(x) := \mathrm{dist}\big(x, \mathrm{Aff}(y_1,y_{-1})\big) + \mathrm{dist}\big(x, \mathrm{Aff}(y_1,y_2)\big) + \mathrm{dist}\big(x, \mathrm{Aff}(y_2,y_{-1})\big).$$

One can easily check that $\mu[f] = \nu[f]$ for any $n \geqslant 1$: $f, f_n \in \mathfrak{C}_0$. These functions separate $I(x_0)$, $I(x_1)$ and $\big(I(x_0) \cup I(x_1)\big)^c$.

Notice that in this example, we may as well take $\theta := 0$, and $N_\nu := \{y_{-1}, y_0, y_1, y_2\}^c$, which minimizes the optimization problem as well.

3.2 Structure of polar sets

Recall the notations in Remark 3.5. Our second main result is:

Theorem 3.7. *A Borel set $N \in \mathcal{B}(\Omega)$ is $\mathcal{M}(\mu,\nu)-$polar if and only if*

$$N \subset \{X \in N_\mu\} \cup \{Y \in N_\nu\} \cup \{Y \notin J_\theta(X)\},\ \text{for some } (N_\mu, N_\nu) \in \mathcal{N}_\mu \times \mathcal{N}_\nu \text{ and } \theta \in \widehat{\mathcal{T}}(\mu,\nu).$$

The proof is reported in Section 6. We conclude this section by reporting a duality result which will be used for proof of Theorem 3.7. We emphasize that the primal objective of the accompanying paper De March [7] is to push further this duality result so as to be suitable for the robust superhedging problem in financial mathematics.

Let $c : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}_+$, and consider the martingale optimal transport problem:

$$\mathbf{S}_{\mu,\nu}(c) \;\; := \sup_{\mathbb{P}\in\mathcal{M}(\mu,\nu)} \mathbb{P}[c]. \tag{3.6}$$

Notice from Proposition 2.11 (i) that $\mathbf{S}_{\mu,\nu}(\theta) \leqslant \nu\widehat{\ominus}\mu(\theta)$ for all $\theta \in \widehat{\mathcal{T}}$. We denote by $\mathcal{D}^{mod}_{\mu,\nu}(c)$ the collection of all $(\varphi, \psi, h, \theta)$ in $\mathbb{L}^1_+(\mu) \times \mathbb{L}^1_+(\nu) \times \mathbb{L}^0(\mathbb{R}^d, \mathbb{R}^d) \times \widehat{\mathcal{T}}(\mu,\nu)$ such that

$$\mathbf{S}_{\mu,\nu}(\theta) = \nu\widehat{\ominus}\mu(\theta), \quad \text{and} \quad \varphi \oplus \psi + h^{\otimes} + \theta \;\geqslant\; c, \text{ on } \{Y \in \operatorname{Aff} K_{\theta,\{\psi=\infty\}}(X)\}.$$

The last inequality is an instance of the so-called robust superhedging property. The dual problem is defined by:

$$\mathbf{I}^{mod}_{\mu,\nu}(c) \;\; := \inf_{(\varphi,\psi,h,\theta)\in\mathcal{D}^{mod}_{\mu,\nu}(c)} \mu[\varphi] + \nu[\psi] + \nu\widehat{\ominus}\mu(\theta).$$

Notice that for any measurable function $c : \Omega \longrightarrow \mathbb{R}_+$, any $\mathbb{P} \in \mathcal{M}(\mu,\nu)$, and any $(\varphi, \psi, h, \theta) \in \mathcal{D}^{mod}_{\mu,\nu}(c)$, we have $\mathbb{P}[c] \leqslant \mu[\varphi] + \nu[\psi] + \mathbb{P}[\theta] \leqslant \mu[\varphi] + \nu[\psi] + \mathbf{S}_{\mu,\nu}(\theta)$, as a consequence of the above robust superhedging inequality, together with the fact that $Y \in \operatorname{Aff} K_{\theta,\{\psi=\infty\}}(X)$, $\mathcal{M}(\mu,\nu)$-q.s. This provides the weak duality:

$$\mathbf{S}_{\mu,\nu}(c) \;\;\leqslant\;\; \mathbf{I}^{mod}_{\mu,\nu}(c). \tag{3.7}$$

The following result states that the strong duality holds for upper semianalytic functions. We recall that a function $f : \mathbb{R}^d \to \mathbb{R}$ is upper semianalytic if $\{f \geqslant a\}$ is an analytic set for any $a \in \mathbb{R}$. In particular, a Borel function is upper semianalytic.

Theorem 3.8. *Let $c : \Omega \to \overline{\mathbb{R}}_+$ be upper semianalytic. Then we have*
(i) $\mathbf{S}_{\mu,\nu}(c) = \mathbf{I}^{mod}_{\mu,\nu}(c)$*;*
(ii) *If in addition $\mathbf{S}_{\mu,\nu}(c) < \infty$, then existence holds for the dual problem $\mathbf{I}^{mod}_{\mu,\nu}(c)$.*

Remark 3.9. *By allowing h to be infinite in some directions, orthogonal to $\operatorname{Aff} K_{\theta,\{\psi=\infty\}}(X)$, together with the convention $\infty - \infty = \infty$, we may reformulate the robust superhedging inequality in the dual set as $\varphi \oplus \psi + h^{\otimes} + \theta \geqslant c$ pointwise.*

3.3 The one-dimensional setting

In the one-dimensional case, the decomposition in irreducible components and the structure of $\mathcal{M}(\mu,\nu)-$polar sets were introduced in Beiglböck & Juillet [3] and Beiglböck, Nutz & Touzi [4], respectively.

Let us see how the results of this paper reduce to the known concepts in the one dimensional case. First, in the one-dimensional setting, $I(x)$ consists of open intervals (at most countable number) or single points. Following [3] Proposition 2.3, we denote the full dimension components $(I_k)_{k\geqslant 1}$.

We also have uniqueness of the $J_\theta(x)$ from Proposition 3.7, as $\underline{J} = J$ (see Proposition 3.10 below), and similarly to the map I, we introduce the corresponding sequence $(J_k)_{k\geqslant 1}$, defined this time in [4]. This Proposition is equivalent in dimension 1 to Theorem 3.2. Similar to [3], we denote by μ_k and ν_k the restrictions of μ and ν to I_k and J_k, respectively. We define the Beiglböck & Juillet (BJ)-irreducible components

$$(I^{BJ}, J^{BJ}) : x \mapsto \begin{cases} (I_k, J_k) & \text{if } x \in I_k^{BJ}, \text{ for some } k \geqslant 1, \\ \big(\{x\}, \{x\}\big) & \text{if } x \notin \cup_k I_k^{BJ}. \end{cases}$$

Proposition 3.10. *Let $d = 1$. Then $I = I^{BJ}$, and $J = \underline{J} = J^{BJ}$, $\mu - a.s.$*

Proof. By Proposition 3.4 (i)-(ii), we may find $\widehat{\mathbb{P}} \in \mathcal{M}(\mu,\nu)$ such that $\widehat{\operatorname{supp}}\,\widehat{\mathbb{P}}_X = \operatorname{cl} I(X)$, and $\widehat{\operatorname{supp}}\,\widehat{\mathbb{P}}_X|_{\partial I(X)} = J\backslash I(X)$, $\mu-$a.s. Notice that as $J\backslash I(\mathbb{R}^d)$ only consists in a countable set of points, we have $\underline{J} = J$. By Theorem 3.2 in [4], we have $Y \in J^{BJ}(X)$, $\mathcal{M}(\mu,\nu)-$q.s. Therefore, $Y \in J^{BJ}(X)$, $\widehat{\mathbb{P}}-$a.s. and we have $J(X) \subset J^{BJ}(X)$, $\mu-$a.s.

On the other hand, let $k \geqslant 1$. By the fact that $u_\nu - u_\mu > 0$ on I_k, together with the fact that $J_k\backslash I_k$ is constituted with atoms of ν, for any $N_\nu \in \mathcal{N}_\nu$, $J_k \subset \operatorname{conv}(J_k\backslash N_\nu)$. As $\mu = \nu$ out of the components,

$$J^{BJ}(X) \subset \operatorname{conv}(J^{BJ}(X)\backslash N_\nu), \quad \mu - \text{a.s.} \tag{3.8}$$

Then by Theorem 3.2 in [4], as $\{Y \notin J(X)\}$ is polar, we may find $N_\nu \in \mathcal{N}_\nu$ such that $J^{BJ}(X)\backslash N_\nu \subset J(X)$, $\mu-$a.s. The convex hull of this inclusion, together with (3.8) gives the remaining inclusion $J^{BJ}(X) \subset J(X)$, $\mu-$a.s.

The equality $I(X) = I^{BJ}(X)$, $\mu-$a.s. follows from the relative interior taken on the previous equality. □

Proposition 3.11. *Let $d = 1$, and $\mathfrak{C}(J)$ the set of convex functions on J. Then,*

$$\widehat{\mathcal{T}}(\mu,\nu) = \Big\{ \textstyle\sum_k \mathbf{1}_{\{X\in I_k\}} \mathbf{T}_{p_k} f_k : f_k \in \mathfrak{C}(J_k),\ p_k \in \partial f_k,\ \sum_k (\nu_k - \mu_k)(f_k) < \infty \Big\},\ \mathcal{M}(\mu,\nu)-q.s.$$

Proof. As all functions we consider are null on the diagonal, equality on $\cup_k I_k \times J_k$ implies $\mathcal{M}(\mu,\nu)-$q.s. equality by Theorem 3.2 in [4]. Let $\mathcal{L}$ be the set on the right hand side.

1. We first show $\subset$, for $a \geqslant 0$, we denote $\mathcal{L}_a := \{\theta \in \mathcal{L} : \sum_k (\nu_k - \mu_k)(f_k) \leqslant a\}$. Notice that $\mathcal{L}_a$ contains $\mathbf{T}(\mathcal{C}_a)$ modulo $\mathcal{M}(\mu,\nu)-$q.s. equality. We intend to prove that $\mathcal{L}_a$ is $\mu\otimes$pw$-$Fatou closed, so as to conclude that $\widehat{\mathcal{T}}_a \subset \mathcal{L}_a$, and therefore $\widehat{\mathcal{T}}(\mu,\nu) \subset \mathcal{L}$ by the arbitrariness of $a \geqslant 0$.

Let $\theta_n = \sum_k \mathbf{1}_{\{X\in I_k\}} \mathbf{T}_{p_k^n} f_k^n \in \mathcal{L}_a$ converging $\mu\otimes$pw. By Proposition 2.6, $\theta_n \longrightarrow \theta := \underline{\theta}_\infty$, $\mu\otimes$pw. For $k \geqslant 1$, let $x_k \in I_k$ be such that $\theta_n(x_k,\cdot) \longrightarrow \theta(x_k,\cdot)$ on $\operatorname{dom}_{x_k}\theta$, and set $f_k := \theta(x_k,\cdot)$. By Proposition 5.5 in [4], f_k is convex on I_k, finite on J_k, and we may find $p_k \in \partial f_k$ such that for $x \in I_k$, $\theta(x,\cdot) = \mathbf{T}_{p_k} f_k(x,\cdot)$. Hence, $\theta = \sum_k \mathbf{1}_{\{X\in I_k\}} \mathbf{T}_{p_k} f_k$, and $\sum_k (\nu_k - \mu_k)(f_k) \leqslant a$ by Fatou's Lemma, implying that $\theta \in \mathcal{L}_a$, as required.

2. To prove the reverse inclusion $\supset$, let $\theta = \sum_k \mathbf{1}_{\{X\in I_k\}} \mathbf{T}_{p_k} f_k \in \mathcal{L}$. Let f_k^ϵ be a convex function defined by $f_k^\epsilon := f_k$ on $J_k^\epsilon = J_k \cap \{x \in J_k : \operatorname{dist}(x, J_k^c) \geqslant \epsilon\}$, and f_k^ϵ afine on $\mathbb{R}\backslash J_k^\epsilon$. Set $\epsilon_n := n^{-1}$, $\bar{f}_n = \sum_{k=1}^n f_k^{\epsilon_n}$, and define the corresponding subgradient in $\partial \bar{f}_n$:

$$\bar{p}_n := p_k + \nabla(\bar{f}_n - f_k^{\varepsilon_n}) \ \text{ on } \ J_k^{\varepsilon_n},\ k \geqslant 1, \quad \text{and} \quad \bar{p}_n := \nabla \bar{f}_n \ \text{ on } \ \mathbb{R}\backslash\big(\cup_k J_k^{\varepsilon_n}\big).$$

We have $(\nu-\mu)(\bar f_n)=\sum_{k=1}^n(\nu_k-\mu_k)(f_k^{\epsilon_n})\leqslant\sum_k(\nu_k-\mu_k)(f_k)<\infty$. By definition, we see that $\mathbf{T}_{\bar p_n}\bar f_n$ converges to θ pointwise on $\cup_k(I_k)^2$ and to $\theta_*(x,y):=\liminf_{\bar y\to y}\theta(x,\bar y)$ on $\cup_k I_k\times \mathrm{cl}\, I_k$ where, using the convention $\infty-\infty=\infty$, $\theta':=\theta-\theta_*\geqslant 0$, and $\theta'=0$ on $\cup_k(I_k)^2$. For $k\geqslant 1$, set $\Delta_k^l:=\theta'(x_k,l_k)$, and $\Delta_k^r:=\theta'(x_k,l_k)$ where $I_k=(l_k,r_k)$. For positiove $\epsilon<\frac{r_k-l_k}{2}$, and $M\geqslant 0$, consider the piecewise affine function $g_k^{\epsilon,M}$ with break points $l_k+\epsilon$ and $r_k-\epsilon$, and:

$$g_k^{\epsilon,M}(l_k)=M\wedge\Delta_k^l,\quad g_k^{\epsilon,M}(r_k)=M\wedge\Delta_k^r,\quad g_k^{\epsilon,M}(l_k+\epsilon)=0,\ \text{and}\ g_k^{\epsilon,M}(r_k-\epsilon)=0.$$

Notice that $g_k^{\epsilon,M}$ is convex, and converges pointwise to $g_k^M:=M\wedge\theta'(\frac{l_k+r_k}{2},\cdot)$ on J_k, as $\epsilon\to 0$, with

$$\begin{aligned}(\nu_k-\mu_k)(g_k^M) &= \nu_k[l_k](M\wedge\Delta_k^l)+\nu_k[r_k](M\wedge\Delta_k^r)\\ &\leqslant (\nu_k-\mu_k)[f_k]-(\nu_k-\mu_k)[(f_k)_*]\ \leqslant\ (\nu_k-\mu_k)[f_k],\end{aligned}$$

where $(f_k)_*$ is the lower semi-continuous envelop of f_k, then by the dominated convergence theorem, we may find positive $\epsilon_k^{n,M}<\frac{r_k-l_k}{2n}$ such that

$$(\nu_k-\mu_k)(g_k^{\epsilon_k^{n,M},M})\leqslant(\nu_k-\mu_k)(f_k)+2^{-k}/n.$$

Now let $\bar g_n=\sum_{k=1}^n g_k^{\epsilon_k^{n,n},n}$, and $\bar p_n'\in\partial\bar g_n$. Notice that $\mathbf{T}_{\bar p_n'}g_n\longrightarrow\theta'$ pointwise on $\cup_k I_k\times J_k$, furthermore, $(\nu-\mu)(\bar g_n)\leqslant\sum_k(\nu_k-\mu_k)(f_k)+1/n\leqslant\sum_k(\nu_k-\mu_k)(f_k)+1<\infty$.

Then we have $\theta_n:=\mathbf{T}_{\bar p_n}\bar f_n+\mathbf{T}_{\bar p_n'}\bar g_n$ converges to θ' pointwise on $\cup_k I_k\times J_k$, and therefore $\mathcal{M}(\mu,\nu)-$q.s. by Theorem 3.2 in [4]. Since $(\nu-\mu)(\bar f_n+\bar g_n)$ is bounded, we see that $(\theta_n)_{n\geqslant 1}\subset\mathbf{T}(\mathfrak{C}_a)$ for some $a\geqslant 0$. Notice that θ_n may fail to converge $\mu\otimes$pw. However, we may use Proposition 2.7 to get a sequence $\widehat\theta_n\in\mathrm{conv}(\theta_k,k\geqslant n)$, and $\widehat\theta_\infty\in\Theta_\mu$ such that $\widehat\theta_n\longrightarrow\widehat\theta_\infty$, $\mu\otimes$pw as $n\to\infty$, and satisfies the same $\mathcal{M}(\mu,\nu)-$q.s. convergence properties than θ_n. Then $\underline{\widehat\theta}_\infty\in\widehat{\mathcal{T}}(\mu,\nu)$, and $\underline{\widehat\theta}_\infty=\theta$, $\mathcal{M}(\mu,\nu)-$q.s. □

Remark 3.12. In the present one-dimensional setting, we have

$$\nu\widehat\ominus\mu(\theta)=\sum_k(\nu_k-\mu_k)(f_k)\quad\text{for all}\quad\theta\in\widehat{\mathcal{T}}(\mu,\nu),$$

where the convex functions f_k are induced by the characterization of Proposition 3.11.

4 The irreducible convex paving

4.1 Existence and uniqueness

Proof of Theorem 3.3 (i) The measurability follows from Lemma 3.1. We first prove the existence of a minimizer for the problem (3.4). Let m denote the infimum in (3.4), and consider a minimizing sequence $(\theta_n,N_\nu^n)_{n\in\mathbb{N}}\subset\widehat{\mathcal{T}}(\mu,\nu)\times\mathcal{N}_\nu$ with $\mu[G(K_{\theta_n,N_\nu^n})]\leqslant m+1/n$. By possibly normalizing the functions θ_n, we may assume that $\nu\widehat\ominus\mu(\theta_n)\leqslant 1$. Set

$$\widehat\theta\ :=\ \sum_{n\geqslant 1}2^{-n}\theta_n\quad\text{and}\quad\widehat N_\nu:=\cup_{n\geqslant 1}N_\nu^n\in\mathcal{N}_\nu.$$

Notice that $\widehat{\theta}$ is well-defined as the pointwise limit of a sequence of the nonnegative functions $\widehat{\theta}_N := \sum_{n\leqslant N} 2^{-n}\theta_n$. Since $\nu\widehat{\ominus}\mu\big[\widehat{\theta}_N\big] \leqslant \sum_{n\geqslant 1} 2^{-n} < \infty$ by convexity of $\nu\widehat{\ominus}\mu$, $\widehat{\theta}_N \longrightarrow \widehat{\theta}$, pointwise, and $\widehat{\theta} \in \mathcal{T}(\mu,\nu)$ by Lemma 2.12, since any convex extraction of $(\theta_n)_{n\geqslant 1}$ still converges to $\widehat{\theta}$. Since $\theta_n^{-1}(\{\infty\}) \subset \widehat{\theta}^{-1}(\{\infty\})$, it follows from the definition of $\widehat{N}_\nu$ that $m + 1/n \geqslant \mu[G(K_{\theta_n,N_\nu^n})] \geqslant \mu[G(K_{\widehat{\theta},\widehat{N}_\nu})]$, hence $\mu[G(K_{\widehat{\theta},\widehat{N}_\nu})] = m$ as $\widehat{\theta} \in \widehat{\mathcal{T}}(\mu,\nu)$, $\widehat{N}_\nu \in \mathcal{N}_\nu$.
(ii) For an arbitrary $(\theta, N_\nu) \in \widehat{\mathcal{T}}(\mu,\nu)\times\mathcal{N}_\nu$, we define $\bar{\theta} := \theta + \widehat{\theta} \in \widehat{\mathcal{T}}(\mu,\nu)$ and $\bar{N}_\nu := \widehat{N}_\nu \cup N_\nu$, so that $K_{\bar{\theta},\bar{N}_\nu} \subset K_{\widehat{\theta},\widehat{N}_\nu}$. By the non-negativity of θ and $\widehat{\theta}$, we have $m \leqslant \mu[G\big(K_{\bar{\theta},\bar{N}_\nu}\big)] \leqslant \mu[G\big(K_{\widehat{\theta},\widehat{N}_\nu}\big)] = m$. Then $G\big(K_{\bar{\theta},\bar{N}_\nu}\big) = G\big(K_{\widehat{\theta},\widehat{N}_\nu}\big)$, μ-a.s. By (3.2), we see that, $\mu-$a.s. $K_{\bar{\theta},\bar{N}_\nu} = K_{\widehat{\theta},\widehat{N}_\nu}$ and $K_{\bar{\theta},\bar{N}_\nu} = K_{\widehat{\theta},\widehat{N}_\nu} = I$. This shows that $I \subset K_{\theta,N_\nu}$, μ-a.s. □

4.2 Partition of the space in convex components

This section is dedicated to the proof of Lemma 3.1 (iii), which is an immediate consequence of Proposition 4.1 (ii).

Proposition 4.1. *Let $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$, and $A \in \mathcal{B}(\mathbb{R}^d)$. We may find $N_\mu \in \mathcal{N}_\mu$ such that:*
(i) *for all $x_1, x_2 \notin N_\mu$ with $K_{\theta,A}(x_1) \cap K_{\theta,A}(x_2) \neq \varnothing$, we have $K_{\theta,A}(x_1) = K_{\theta,A}(x_2)$;*
(ii) *if $A \in \mathcal{N}_\nu$, then $x \in K_{\theta,A}(x)$ for $x \notin N_\mu$, and up to a modification of $K_{\theta,A}$ on N_μ, $K_{\theta,A}(\mathbb{R}^d)$ is a partition of $\mathbb{R}^d$ such that $x \in K_{\theta,A}(x)$ for all $x \in \mathbb{R}^d$.*

Proof. (i) Let N_μ be the $\mu-$null set given by Proposition 2.10 for θ. For $x_1, x_2 \notin N_\mu$, we suppose that we may find $\bar{y} \in K_{\theta,A}(x_1) \cap K_{\theta,A}(x_2)$. Consider $y \in \mathrm{cl}\, K_{\theta,A}(x_1)$, as $K_{\theta,A}(x_1)$ is open in its affine span, $y' := \bar{y} + \frac{\epsilon}{1-\epsilon}(\bar{y} - y) \in K_{\theta,A}(x_1)$ for $0 < \epsilon < 1$ small enough. Then $\bar{y} = \epsilon y + (1-\epsilon)y'$, and by Proposition 2.10, we get

$$\epsilon\theta(x_1, y) + (1-\epsilon)\theta(x_1, y') - \theta(x_1, \bar{y}) = \epsilon\theta(x_2, y) + (1-\epsilon)\theta(x_2, y') - \theta(x_2, \bar{y})$$

By convexity of $\mathrm{dom}_{x_i}\theta$, $K_{\theta,A}(x_i) \subset \mathrm{dom}_{x_i}\theta \subset \mathrm{dom}\theta(x_i,\cdot)$. Then $\theta(x_1, y')$, $\theta(x_1,\bar{y})$, $\theta(x_2, y')$, and $\theta(x_2,\bar{y})$ are finite and

$$\theta(x_1, y) < \infty \quad \text{if and only if} \quad \theta(x_2, y) < \infty.$$

Therefore $\mathrm{cl}\, K_{\theta,A}(x_1) \cap \mathrm{dom}\theta(x_1,\cdot) \subset \mathrm{dom}\theta(x_2,\cdot)$. We have obviously $\mathrm{cl}\, K_{\theta,A}(x_2) \cap \mathrm{dom}\theta(x_2,\cdot) \subset \mathrm{dom}\theta(x_2,\cdot)$ as well. Subtracting A, we get

$$\big(\mathrm{cl}\, K_{\theta,A}(x_1) \cap \mathrm{dom}\theta(x_1,\cdot)\backslash A\big) \cup \big(\mathrm{cl}\, K_{\theta,A}(x_2) \cap \mathrm{dom}\theta(x_2,\cdot)\backslash A\big) \subset \mathrm{dom}\theta(x_2,\cdot)\backslash A.$$

Taking the convex hull and using the fact that the relative face of a set is included in itself, we see that $\mathrm{conv}\big(K_{\theta,A}(x_1) \cup K_{\theta,A}(x_2)\big) \subset \mathrm{conv}\big(\mathrm{dom}\theta(x_2,\cdot)\backslash A\big)$. Finally, as $K_{\theta,A}(x_1)$ and $K_{\theta,A}(x_2)$ are convex sets intersecting in relative interior points and $x_2 \in \mathrm{ri}K_{\theta,A}(x_2)$, it follows from Lemma 9.1 that $x_2 \in \mathrm{riconv}\big(K_{\theta,A}(x_1) \cup K_{\theta,A}(x_2)\big)$. Then by Proposition 2.1 (ii),

$$\mathrm{rf}_{x_2}\mathrm{conv}\big(K_{\theta,A}(x_1) \cup K_{\theta,A}(x_2)\big) = \mathrm{ri}\,\mathrm{conv}\big(K_{\theta,A}(x_1) \cup K_{\theta,A}(x_2)\big) = \mathrm{conv}\big(K_{\theta,A}(x_1) \cup K_{\theta,A}(x_2)\big).$$

Then, we have $\operatorname{conv}\big(K_{\theta,A}(x_1) \cup K_{\theta,A}(x_2)\big) \subset \operatorname{rf}_{x_2}\operatorname{conv}\big(\operatorname{dom}\theta(x_2,\cdot)\backslash A\big) = K_{\theta,A}(x_2)$, as rf_{x_2} is increasing. Therefore $K_{\theta,A}(x_1) \subset K_{\theta,A}(x_2)$ and by symmetry between x_1 and x_2, $K_{\theta,A}(x_1) = K_{\theta,A}(x_2)$.

(ii) We suppose that $A \in \mathcal{N}_\nu$. First, notice that, as $K_{\theta,A}(X)$ is defined as the $X-$relative face of some set, either $x \in K_{\theta,A}(x)$ or $K_{\theta,A}(x) = \emptyset$ for $x \in \mathbb{R}^d$ by the properties of rf_x. Consider $\mathbb{P} \in \mathcal{M}(\mu,\nu)$. By Proposition 3.2, $\mathbb{P}[Y \in \operatorname{cl} K_{\theta,A}(X)] = 1$. As $\operatorname{supp}(\mathbb{P}_X) \subset \operatorname{cl} K_{\theta,A}(X)$, μ-a.s., $K_{\theta,A}(X)$ is non-empty, which implies that $x \in K_{\theta,A}(x)$. Hence, $\{X \in K_{\theta,A}(X)\}$ holds on the set $N_\mu^0 := \{\operatorname{supp}(\mathbb{P}_X) \not\subset \operatorname{cl} I(X)\} \in \mathcal{N}_\mu$. Then we just need to have this property to replace N_μ by $N_\mu \cup N_\mu^0 \in \mathcal{N}_\mu$.

Finally, to get a partition of $\mathbb{R}^d$, we just need to redefine $K_{\theta,A}$ on N_μ. If $x \in \bigcup_{x' \notin N_\mu} K_{\theta,A}(x')$ then $K_{\theta,A}(x')$ is independent of x' such that $x \in K_{\theta,A}(x')$ by definition of N_μ. We set $K_{\theta,A}(x) := K_{\theta,A}(x')$. Otherwise, if $x \notin \bigcup_{x' \notin N_\mu} K_{\theta,A}(x')$, we set $K_{\theta,A}(x) := \{x\}$ which is trivially convex and relatively open. With this definition, $K_{\theta,A}(\mathbb{R}^d)$ is a partition of $\mathbb{R}^d$. □

5 Proof of the duality

For simplicity, we denote $\operatorname{Val}(\xi) := \mu[\varphi] + \nu[\psi] + \nu\widehat{\ominus}\mu(\theta)$, for $\xi := (\varphi,\psi,h,\theta) \in \mathcal{D}^{mod}_{\mu,\nu}(c)$.

5.1 Existence of dual optimizer

Lemma 5.1. *Let $c, c_n : \Omega \longrightarrow \overline{\mathbb{R}_+}$, and $\xi_n \in \mathcal{D}^{mod}_{\mu,\nu}(c_n)$, $n \in \mathbb{N}$, be such that*

$$c_n \longrightarrow c, \ \textit{pointwise, and} \ \operatorname{Val}(\xi_n) \longrightarrow \mathbf{S}_{\mu,\nu}(c) < \infty \ \textit{as} \ n \to \infty.$$

Then there exists $\xi \in \mathcal{D}^{mod}_{\mu,\nu}(c)$ such that $\operatorname{Val}(\xi_n) \longrightarrow \operatorname{Val}(\xi)$ as $n \to \infty$.

Proof. Denote $\xi_n := (\varphi_n, \psi_n, h_n, \theta_n)$, and observe that the convergence of $\operatorname{Val}(\xi_n)$ implies that the sequence $\big(\mu(\varphi_n), \nu(\psi_n), \nu\widehat{\ominus}\mu(\theta_n)\big)_n$ is bounded, by the non-negativity of φ_n, ψ_n and $\nu\widehat{\ominus}\mu(\theta_n)$. We also recall the robust superhedging inequality

$$\varphi_n \oplus \psi_n + h_n^{\otimes} + \theta_n \geqslant c_n, \quad \text{on } \{Y \in \operatorname{Aff} K_{\theta_n,\{\psi_n=\infty\}}(X)\}, \quad n \geqslant 1. \tag{5.9}$$

<u>Step 1.</u> By Komlòs Lemma together with Lemma 2.12, we may find a sequence $(\widehat{\varphi}_n, \widehat{\psi}_n, \widehat{\theta}_n) \in \operatorname{conv}\{(\varphi_k, \psi_k, \theta_k), k \geqslant n\}$ such that

$$\widehat{\varphi}_n \longrightarrow \varphi := \underline{\widehat{\varphi}}_\infty, \ \mu-\text{a.s.}, \ \widehat{\psi}_n \longrightarrow \psi := \underline{\widehat{\psi}}_\infty, \ \nu-\text{a.s., and}$$
$$\widehat{\theta}_n \longrightarrow \widetilde{\theta} := \underline{\widehat{\theta}}_\infty \in \widehat{\mathcal{T}}(\mu,\nu), \ \mu \otimes \text{pw}.$$

Set $\varphi := \infty$ and $\psi := \infty$ on the corresponding non-convergence sets, and observe that $\mu[\varphi] + \nu[\psi] < \infty$, by the Fatou Lemma, and therefore $N_\mu := \{\varphi = \infty\} \in \mathcal{N}_\mu$ and $N_\nu := \{\psi = \infty\} \in \mathcal{N}_\nu$. We denote by $(\widehat{h}_n, \widehat{c}_n)$ the same convex extractions from $\{(h_k, c_k), k \geqslant n\}$, so that the sequence $\widehat{\xi}_n := (\widehat{\varphi}_n, \widehat{\psi}_n, \widehat{h}_n, \widehat{\theta}_n)$ inherits from (5.9) the robust superhedging property, as for

$\theta_1, \theta_1 \in \widehat{\mathcal{T}}(\mu,\nu)$, $\psi_1, \psi_2 \in \mathbb{L}^1_+(\mathbb{R}^d)$, and $0 < \lambda < 1$, we have $\mathrm{Aff} K_{\lambda\theta_1+(1-\lambda)\theta_2, \{\lambda\psi_1+(1-\lambda)\psi_2=\infty\}} \subset \mathrm{Aff} K_{\theta_1,\{\psi_1=\infty\}} \cap \mathrm{Aff} K_{\theta_2,\{\psi_2=\infty\}}$:

$$\widehat{\varphi}_n \oplus \widehat{\psi}_n + \widehat{\theta}_n + \widehat{h}^{\otimes}_n \geqslant \widehat{c}_n \geqslant 0, \quad \text{pointwise on } \mathrm{Aff} K_{\widehat{\theta}_n,\{\widehat{\psi}_n=\infty\}}(X). \tag{5.10}$$

Step 2. Next, notice that $l_n := (\widehat{h}^{\otimes}_n)^- \in \Theta$ for all $n \in \mathbb{N}$. By the convergence Proposition 2.7, we may find convex combinations $\widehat{l}_n := \sum_{k \geqslant n} \lambda^n_k l_k \longrightarrow l := \underline{\widehat{l}}_\infty$, $\mu\otimes$pw. Updating the definition of φ by setting $\varphi := \infty$ on the zero $\mu-$measure set on which the last convergence does not hold on $(\partial^x \mathrm{dom} l)^c$, it follows from (5.10), and the fact that $\mathrm{Aff} K_{\bar{\theta},\{\psi=\infty\}} \subset \liminf_{n\to\infty} \mathrm{Aff} K_{\widehat{\theta}_n,\{\widehat{\psi}_n=\infty\}}$, that

$$l = \underline{\widehat{l}}_\infty \leqslant \liminf_n \sum_{k\geqslant n} \lambda^n_k \big(\widehat{\varphi}_k \oplus \widehat{\psi}_k + \widehat{\theta}_k\big) \leqslant \varphi \oplus \psi + \bar{\theta}, \text{ pointwise on } \{Y \in \mathrm{Aff} K_{\bar{\theta},\{\psi=\infty\}}(X)\}.$$

where $\bar{\theta} := \liminf_n \sum_{k\geqslant n} \lambda^n_k \widehat{\theta}_k \in \widehat{\mathcal{T}}(\mu,\nu)$. Consequently, updating also $N_\mu := N_\mu \cup \{\varphi = \infty\}$ which is still in $\mathcal{N}_\mu$, we see that $\mathrm{dom}\, l \supset (N^c_\mu \times N^c_\nu) \cap \mathrm{dom}\bar{\theta} \cap \{Y \in \mathrm{Aff} K_{\bar{\theta},\{\psi=\infty\}}(X)\}$, and therefore

$$K_{\bar{\theta},\{\psi=\infty\}}(X) \subset \mathrm{dom}_X l' \subset \mathrm{dom}\, l'(X,\cdot), \quad \mu\text{-a.s.} \tag{5.11}$$

Step 3. Let $\widehat{\widehat{h}}_n := \sum_{k\geqslant n} \lambda^n_k \widehat{h}_k$. Then $b_n := \widehat{\widehat{h}}^{\otimes}_n + \widehat{l}_n = \sum_{k\geqslant n} \lambda^n_k (\widehat{h}^{\otimes}_k)^+$ defines a non-negative sequence in Θ. By Proposition 2.7, we may find a sequence $\widehat{b}_n =: \widetilde{h}^{\otimes}_n + \widetilde{l}_n \in \mathrm{conv}(b_k, k \geqslant n)$ such that $\widehat{b}_n \longrightarrow b := \underline{\widehat{b}}_\infty$, $\mu\otimes$pw, where b takes values in $[0,\infty]$. $\widehat{b}_n(X,\cdot) \longrightarrow b(X,\cdot)$ pointwise on $\mathrm{dom}_X b$, $\mu-$a.s. Combining with (5.11), this shows that

$$\widetilde{h}^{\otimes}_n(X,\cdot) \longrightarrow (b-l)(X,\cdot) \quad \text{pointwise on} \quad \mathrm{dom}_X b \cap K_{\bar{\theta},\{\psi=\infty\}}(X), \quad \mu-\text{a.s.}$$

$(b-l)(X,\cdot) = \liminf_n \widetilde{h}^{\otimes}_n(X,\cdot)$, pointwise on $K_{\bar{\theta},\{\psi=\infty\}}(X)$ (where l is a limit of l_n), $\mu-$a.s. Clearly, on the last convergence set, $(b-l)(X,\cdot) > -\infty$ on $K_{\bar{\theta},\{\psi=\infty\}}(X)$, and we now argue that $(b-l)(X,\cdot) < \infty$ on $K_{\bar{\theta},\{\psi=\infty\}}(X)$, therefore $K_{\bar{\theta},\{\psi=\infty\}}(X) \subset \mathrm{dom}_X b$, so that we deduce from the structure of $\widetilde{h}^{\otimes}_n$ that the last convergence holds also on $\mathrm{Aff} K_{\bar{\theta},\{\psi=\infty\}}(X)$:

$$\widetilde{h}^{\otimes}_n(X,\cdot) \longrightarrow (b-l)(X,\cdot) =: h^{\otimes}(X,\cdot) \quad \text{pointwise on} \quad K_{\bar{\theta},\{\psi=\infty\}}(X), \quad \mu-\text{a.s.} \tag{5.12}$$

Indeed, let x be an arbitrary point of the last convergence set, and consider an arbitrary $y \in K_{\bar{\theta},\{\psi=\infty\}}(x)$. By the definition of $K_{\bar{\theta},\{\psi=\infty\}}$, we have $x \in \mathrm{ri} K_{\bar{\theta},\{\psi=\infty\}}(x)$, and we may therefore find $y' \in K_{\bar{\theta},\{\psi=\infty\}}(x)$ with $x = py + (1-p)y'$ for some $p \in (0,1)$. Then, $p\,\widetilde{h}^{\otimes}_n(x,y) + (1-p)\widetilde{h}^{\otimes}_n(x,y') = 0$. Sending $n \to \infty$, by concavity of the $\liminf$, this provides $p(b-l)(x,y) + (1-p)(b-l)(x,y') \leqslant 0$, so that $(b-l)(x,y') > -\infty$ implies that $(b-l)(x,y) < \infty$.

Step 4. Notice that by dual reflexivity of finite dimensional vector spaces, (5.12) defines a unique $h(X)$ in the vector space $\mathrm{Aff} K_{\bar{\theta},\{\psi=\infty\}}(X) - X$, such that $(b-l)(X,\cdot) = h^{\otimes}(X,\cdot)$ on $\mathrm{Aff} K_{\bar{\theta},\{\psi=\infty\}}(X)$. At this point, we have proceeded to a finite number of convex combinations which induce a final convex combination with coefficients $(\bar{\lambda}^k_n)_{k\geqslant n\geqslant 1}$. denote $\bar{\xi}_n := \sum_{k\geqslant n} \bar{\lambda}^k_n \xi_k$, and set $\theta := \underline{\bar{\theta}}_\infty$. Then, applying this convex combination to the robust superhedging inequality (5.9), we obtain by sending $n \to \infty$ that $(\varphi \oplus \psi + h^{\otimes} + \theta)(X,\cdot) \geqslant c(X,\cdot)$ on $\mathrm{Aff} K_{\bar{\theta},\{\psi=\infty\}}(X)$,

$\mu-$a.s. and $\varphi \oplus \psi + h^{\otimes} + \theta = \infty$ on the complement μ null-set. As θ is the liminf of a convex extraction of $(\widehat{\theta}_n)$, we have $\theta \geqslant \underline{\widehat{\theta}}_\infty = \bar{\theta}$, and therefore $\mathrm{Aff}\, K_{\theta,\{\psi=\infty\}} \subset \mathrm{Aff}\, K_{\bar{\theta},\{\psi=\infty\}}$. This shows that the limit point $\xi := (\varphi, \psi, h, \theta)$ satisfies the pointwise robust superhedging inequality

$$\varphi \oplus \psi + \theta + h^{\otimes} \geqslant c, \quad \text{on } \{Y \in \mathrm{Aff}\, K_{\theta,\{\psi=\infty\}}(X)\}. \tag{5.13}$$

Step 5. By Fatou's Lemma and Lemma 2.12, we have

$$\mu[\varphi] + \nu[\psi] + \nu\widehat{\ominus}\mu[\theta] \;\leqslant\; \liminf_n \mu[\varphi_n] + \nu[\psi_n] + \nu\widehat{\ominus}\mu[\theta_n] \;=\; \mathbf{S}_{\mu,\nu}(c). \tag{5.14}$$

By (5.13), we have $\mu[\varphi] + \nu[\psi] + \mathbb{P}[\theta] \geqslant \mathbb{P}[c]$ for all $\mathbb{P} \in \mathcal{M}(\mu,\nu)$. Then, $\mu[\varphi] + \nu[\psi] + \mathbf{S}_{\mu,\nu}[\theta] \geqslant \mathbf{S}_{\mu,\nu}[c]$. By Proposition 2.11 (i), we have $\mathbf{S}_{\mu,\nu}[\theta] \leqslant \nu\widehat{\ominus}\mu[\theta]$, and therefore

$$\mathbf{S}_{\mu,\nu}[c] \;\leqslant\; \mu[\varphi] + \nu[\psi] + \mathbf{S}_{\mu,\nu}[\theta] \;\leqslant\; \mu[\varphi] + \nu[\psi] + \nu\widehat{\ominus}\mu[\theta] \;\leqslant\; \mathbf{S}_{\mu,\nu}(c),$$

by (5.14). Then we have $\mathrm{Val}(\xi) = \mu[\varphi] + \nu[\psi] + \nu\widehat{\ominus}\mu[\theta] = \mathbf{S}_{\mu,\nu}(c)$ and $\mathbf{S}_{\mu,\nu}[\theta] = \nu\widehat{\ominus}\mu[\theta]$, so that $\xi \in \mathcal{D}^{mod}_{\mu,\nu}(c)$. □

5.2 Duality result

We first prove the duality in the set USC_b of all bounded upper semicontinuous functions $\Omega \to \overline{\mathbb{R}_+}$. This is a classical result using the Hahn-Banach Theorem, the proof is reported for completeness.

Lemma 5.2. *Let $f \in \mathrm{USC}_b$, then $\mathbf{S}_{\mu,\nu}(f) = \mathbf{I}^{mod}_{\mu,\nu}(f)$*

Proof. We have $\mathbf{S}_{\mu,\nu}(f) \leqslant \mathbf{I}^{mod}_{\mu,\nu}(f)$ by weak duality (3.7), let us now show the converse inequality $\mathbf{S}_{\mu,\nu}(f) \geqslant \mathbf{I}^{mod}_{\mu,\nu}(f)$. By standard approximation technique, it suffices to prove the result for bounded continuous f. We denote by $\mathrm{C}_l(\mathbb{R}^d)$ the set of continuous mappings $\mathbb{R}^d \to \mathbb{R}$ with linear growth at infinity, and by $\mathrm{C}_b(\mathbb{R}^d, \mathbb{R}^d)$ the set of continuous bounded mappings $\mathbb{R}^d \longrightarrow \mathbb{R}^d$. Define

$$\mathcal{D}(f) := \Big\{(\bar{\varphi}, \bar{\psi}, \bar{h}) \in \mathrm{C}_l(\mathbb{R}^d) \times \mathrm{C}_l(\mathbb{R}^d) \times \mathrm{C}_b(\mathbb{R}^d, \mathbb{R}^d) : \; \bar{\varphi} \oplus \bar{\psi} + \bar{h}^{\otimes} \geqslant f\Big\},$$

and the associated $\mathbf{I}_{\mu,\nu}(f) := \inf_{(\bar{\varphi},\bar{\psi},\bar{h}) \in \mathcal{D}(f)} \mu(\bar{\varphi}) + \nu(\bar{\psi})$. By Theorem 2.1 in Zaev [22], and Lemma 5.3 below, we have

$$\mathbf{S}_{\mu,\nu}(f) \;\;=\;\; \mathbf{I}_{\mu,\nu}(f) \;\;=\;\; \inf_{(\bar{\varphi},\bar{\psi},\bar{h}) \in \mathcal{D}(f)} \mu(\bar{\varphi}) + \nu(\bar{\psi}) \;\geqslant\; \mathbf{I}^{mod}_{\mu,\nu}(f),$$

which provides the required result. □

Proof of Theorem 3.8 The existence of a dual optimizer follows from a direct application of the compactness Lemma 5.1 to a minimizing sequence of robust superhedging strategies.

As for the extension of the duality result to non-negative upper semi-analytic functions, we shall use the capacitability theorem of Choquet, similar to [17] and [4]. Let $[0,\infty]^\Omega$ denote the set of all nonnegative functions $\Omega \to [0,\infty]$, USA_+ the sublattice of upper semianalytic functions, and USC_b the sublattice of bounded upper semicontinuous fonctions. Note that USC_b is stable by infimum.

Recall that a USC_b-capacity is a monotone map $\mathbf{C} : [0,\infty]^\Omega \longrightarrow [0,\infty]$, sequentially continuous upwards on $[0,\infty]^\Omega$, and sequentially continuous downwards on USC_b. The Choquet capacitability theorem states that a USC_b−capacity $\mathbf{C}$ extends to USA_+ by:

$$\mathbf{C}(f) \;=\; \sup\big\{\mathbf{C}(g) : g \in \mathrm{USC}_b \text{ and } g \leqslant f\big\} \quad \text{for all} \quad f \in \mathrm{USA}_+.$$

In order to prove the required result, it suffices to verify that $\mathbf{S}_{\mu,\nu}$ and $\mathbf{I}^{mod}_{\mu,\nu}$ are USC_b-capacities. As $\mathcal{M}(\mu,\nu)$ is weakly compact, it follows from similar argument as in Prosposition 1.21, and Proposition 1.26 in Kellerer [17] that $\mathbf{S}_{\mu,\nu}$ is a USC_b-capacity. We next verify that $\mathbf{I}^{mod}_{\mu,\nu}$ is a USC_b-capacity. Indeed, the upwards continuity is inherited from $\mathbf{S}_{\mu,\nu}$ together with the compactness lemma 5.1, and the downwards continuity follows from the downwards continuity of $\mathbf{S}_{\mu,\nu}$ together with the duality result on USC_b of Lemma 5.2. □

Lemma 5.3. *Let $c : \Omega \to \overline{\mathbb{R}}_+$, and $(\bar\varphi,\bar\psi,\bar h) \in \mathcal{D}(c)$. Then, we may find $\xi \in \mathcal{D}^{mod}_{\mu,\nu}(c)$ such that* $\mathrm{Val}(\xi) = \mu[\bar\varphi] + \nu[\bar\psi]$.

Proof. Let us consider $(\bar\varphi,\bar\psi,\bar h) \in \mathcal{D}(c)$. Then $\bar\varphi \oplus \bar\psi + \bar h^{\otimes} \geqslant c \geqslant 0$, and therefore

$$\bar\psi(y) \geqslant f(y) := \sup_{x\in\mathbb{R}^d} -\bar\varphi(x) - \bar h(x)\cdot(y-x).$$

Clearly, f is convex, and $f(x) \geqslant -\bar\varphi(x)$ by taking value $x = y$ in the supremum. Hence $\bar\psi - f \geqslant 0$ and $\bar\varphi + f \geqslant 0$, implying in particular that f is finite on $\mathbb{R}^d$. As $\bar\varphi$ and $\bar\psi$ have linear growth at infinity, f is in $\mathbb{L}^1(\nu) \cap \mathbb{L}^1(\mu)$. We have $f \in \mathfrak{C}_a$ for $a = \nu[f] - \mu[f] \geqslant 0$. Then we consider $p \in \partial f$ and denote $\theta := \mathbf{T}_p f$. $\theta \in \mathbf{T}\big(\mathfrak{C}_a\big) \subset \widehat{\mathcal{T}}(\mu,\nu)$. Then denoting $\varphi := \bar\varphi + f$, $\psi := \bar\psi - f$, and $h := \bar h + p$, we have $\xi := (\varphi,\psi,h,\theta) \in \mathcal{D}^{mod}_{\mu,\nu}(c)$ and

$$\mu[\bar\varphi] + \nu[\bar\psi] = \mu[\varphi] + \nu[\psi] + (\nu-\mu)[f] = \mu[\varphi] + \nu[\psi] + \nu\widehat{\ominus}\mu[\theta] = \mathrm{Val}(\xi).$$

□

6 Polar sets and maximum support martingale plan

6.1 The mapping J

Consider the optimization problems:

$$\inf_{(\theta,N_\nu)\in\widehat{\mathcal{T}}(\mu,\nu)\times\mathcal{N}_\nu} \mu[G(R_{\theta,N_\nu})], \text{ with } R_{\theta,N_\nu} := \mathrm{cl\,conv}\big(\mathrm{dom}\theta(X,\cdot) \cap \partial I(X) \cap N_\nu^c\big), \tag{6.15}$$

and

$$\inf_{(\theta,N_\nu)\in\widehat{\mathcal{T}}(\mu,\nu)\times\mathcal{N}_\nu} \mu\big[y \in \partial I(X) \cap \mathrm{dom}\theta(X,\cdot) \cap N_\nu^c\big] \quad \text{for all} \quad y \in \mathbb{R}^d. \tag{6.16}$$

These problems are well defined by the following measurability result, whose proof is reported in Subsection 7.2.

Lemma 6.1. *Let $F : \mathbb{R}^d \longrightarrow \mathcal{K}$, universally measurable. For all $\gamma \in \mathcal{P}(\mathbb{R}^d)$, we may find $N_\gamma \in \mathcal{N}_\gamma$ such that $\mathbf{1}_{Y \in F(X)} \mathbf{1}_{X \notin N_\gamma}$ is Borel measurable, and if $X \in \mathrm{ri} F(X)$ convex, $\gamma-$a.s., then $\mathbf{1}_{Y \in \partial F(X)} \mathbf{1}_{X \notin N_\gamma}$ is Borel measurable as well.*

By the same argument than that of the proof of existence and uniqueness in Theorem 3.3, we see that the problem (6.15), (resp. (6.16) for $y \in \mathbb{R}^d$) has an optimizer $(\theta^*, N_\nu^*) \in \mathcal{T}(\mu,\nu) \times \mathcal{N}_\nu$, (resp. $(\theta_y^*, N_{\nu,y}^*) \in \mathcal{T}(\mu,\nu) \times \mathcal{N}_\nu$). Furthermore, $D := R_{\theta^*, N_\nu^*}$, (resp $D_y(x) := \{y\}$ if $y \in \partial I(x) \cap \mathrm{dom}\theta_y^*(x,\cdot) \cap N_{\nu,y}^*$, and $\varnothing$ otherwise, for $x \in \mathbb{R}^d$) does not depend on the choice of (θ^*, N_ν^*), (resp. θ_y^*) up to a $\mu-$negligible modification.

We define $J := D \cup I$, and recall that for $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$, we denote $J_\theta(X) := \mathrm{dom}\theta(X,\cdot) \cap J(X)$. Notice that if $y \in \mathbb{R}^d$ is not an atom of ν, we may chose $N_{\nu,y}$ containing y, which means that Problem (6.16) is non-trivial only if y is an atom of ν. We denote $\mathrm{atom}(\nu)$, the (at most countable) atoms of ν, and define the mapping $\underline{J} := (\cup_{y \in \mathrm{atom}(\nu)} D_y) \cup I$,

Proposition 6.2. *Let $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$, up to a modification on a $\mu-$null set preserving Theorem 3.3 (iii), we have,*
(i) *J is convex valued, moreover $Y \in J(X)$, and $Y \in J_\theta(X)$, $\mathcal{M}(\mu,\nu)-$q.s.*
(ii) *$I \subset \underline{J} \subset J_\theta \subset J \subset \mathrm{cl}\, I$,*
(iii) *$\underline{J}$, J_θ, and J are constant on $I(x)$, for all $x \in \mathbb{R}^d$.*

Proof. We suppose that I is changed such that I satisfies Theorem 3.3 (iii). (i) For $x \in \mathbb{R}^d$, $J(x) = D(x) \cup I(x)$. Let $y_1, y_2 \in J(x)$, $\lambda \in (0,1)$, and set $y := \lambda y_1 + (1-\lambda) y_2$. If $y_1, y_2 \in I(x)$, or $y_1, y_2 \in D(x)$, we get $y \in J(x)$ by convexity of $I(x)$, or $D(x)$. Now, up to switching the indices, we may assume that $y_1 \in I(x)$, and $y_2 \in D(x) \backslash I(x)$. As $D(x) \backslash I(x) \subset \partial I(x)$, $y \in I(x)$, as $\lambda > 0$. Then $y \in J(x)$. Hence, J is convex valued.

Since $\mathrm{dom}\theta^*(X,\cdot) \backslash N_\nu^* \cap \mathrm{cl}\, I \backslash I \subset R_{\theta^*, N_\nu^*}$, we have the inclusion $\mathrm{dom}\theta^*(X,\cdot) \backslash N_\nu^* \cap \mathrm{cl}\, I \subset R_{\theta^*, N_\nu^*} \cup I = J$. Then, as $Y \in \mathrm{dom}\theta^*(X,\cdot) \backslash N_\nu^*$, and $Y \in \mathrm{cl}\, I(X)$, $Y \in J(X)$, $\mathcal{M}(\mu,\nu)-$q.s.

Let $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$, $Y \in \mathrm{dom}(X,\theta)$, $\mathcal{M}(\mu,\nu)-$q.s. Finally we get $Y \in \mathrm{dom}(X,\theta) \cap J(X) = J_\theta(X)$, $\mathcal{M}(\mu,\nu)-$q.s.
(ii) As $R_{\theta, N_\nu}(X) \subset \mathrm{cl}\,\mathrm{conv}\partial I(X) = \mathrm{cl}\, I(X)$, $J \subset \mathrm{cl}\, I$. By definition, $J_\theta \subset J$, and $I \subset \underline{J}$. For $y \in \mathrm{atom}(\nu)$, and $\theta_0 \in \widehat{\mathcal{T}}(\mu,\nu)$, by minimality, $D_y(X) \subset \mathrm{dom}\theta_0(X,\cdot)\theta \cap \partial I(X)$, $\mu-$a.s. Applying it for $\theta_0 = \theta$, we get $D_y \subset \mathrm{dom}\theta(X,\cdot)$, and for $\theta_0 = \theta^*$, $D_y(X) \subset J(X)$, $\mu-$a.s. Taking the countable union: $\underline{J} \subset J_\theta$, $\mu-$a.s. (This is the only inclusion that is not pointwise). The we change $\underline{J}$ to I on this set to get this inclusion pointwise.
(iii) For $\theta_0 \in \widehat{\mathcal{T}}(\mu,\nu)$, we may apply Proposition 2.10. For some $N_\mu \in \mathcal{N}_\mu$, for $x \in N_\mu^c$, for $y \in \partial I(X)$, y':= $\frac{x+y}{2} \in I(x)$. Then for any other $x' \in I(x) \cap N_\mu^c$, $\frac{1}{2}\theta_0(x,y) - \theta_0(x,y') = \frac{1}{2}\theta_0(x',y) - \theta_0(x',y')$, in particular, $y \in \mathrm{dom}\theta(x,\cdot)$ if and only if $y \in \mathrm{dom}\theta(x',\cdot)$. Applying this result to θ, θ^*, and θ_y^* for all $y \in \mathrm{atom}(\nu)$, we get N_μ such that for any $x \in \mathbb{R}^d$, J, J_θ, and $\underline{J}$ are constant on $I(x) \cap N_\mu^c$. To get it pointwise, we redefine these mappings to this constant value on $I(x) \cap N_\mu$, or to $I(x)$, if $I(x) \cap N_\mu^c = \varnothing$. The previous properties are preserved. □

6.2 Structure of polar sets

Proof of Proposition 3.7 One implication is trivial as $Y \in J_\theta(X)$, $\mathcal{M}(\mu,\nu)-$q.s. for all $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$, by Proposition 6.2. We only focus on the non-trivial implication. For an $\mathcal{M}(\mu,\nu)$-polar set N, we have $\mathbf{S}_{\mu,\nu}(\infty\mathbf{1}_B) = 0$, and it follows from the dual formulation of Theorem 3.8 that $0 = \mathrm{Val}(\xi)$ for some $\xi = (\varphi,\psi,h,\theta) \in \mathcal{D}^{mod}_{\mu,\nu}(\infty\mathbf{1}_B)$. Then,

$$\varphi < \infty,\ \mu-\text{a.s.},\quad \psi < \infty,\ \nu-\text{a.s.}\quad \text{and } \theta \in \widehat{\mathcal{T}}(\mu,\nu),$$

As h is finite valued, and φ,ψ are non-negative functions, the superhedging inequality $\varphi\oplus\psi+\theta+h^\otimes \geqslant \infty\mathbf{1}_B$ on $\{Y \in \mathrm{Aff}K_{\theta,\{\psi=\infty\}}(X)\}$ implies that

$$\mathbf{1}_{\{\varphi=\infty\}}\oplus\mathbf{1}_{\{\psi=\infty\}}+\mathbf{1}_{\{(\mathrm{dom}\theta)^c\}} \geqslant \mathbf{1}_B \quad \text{on} \quad \{Y \in \mathrm{Aff}K_{\theta,\{\psi=\infty\}}(X)\} \tag{6.17}$$

By Theorem 3.3 (ii), we have $I(X) \subset K_{\theta,\{\psi=\infty\}}(X)$, $\mu-$a.s. Then $J(X) \subset \mathrm{Aff}I(X) \subset K_{\theta,\{\psi=\infty\}}(X)$, which implies that

$$J_\theta(X) := \mathrm{dom}\theta(X,\cdot)\cap J(X) \subset \mathrm{dom}\theta(X,\cdot)\cap \mathrm{Aff}K_{\theta,\{\psi=\infty\}}(X),\quad \mu-\text{a.s.} \tag{6.18}$$

We denote $N_\mu := \{\phi=\infty\}\cup\{J_\theta(X)\not\subset \mathrm{dom}\theta(X,\cdot)\cap\mathrm{Aff}K_{\theta,\{\psi=\infty\}}(X)\}\in\mathcal{N}_\mu$, and $N_\nu := \{\psi=\infty\}\in\mathcal{N}_\nu$. Then by (6.17), $\mathbf{1}_B = 0$ on $(\{\phi=\infty\}^c\times\{\psi=\infty\}^c)\cap\{Y\in\mathrm{dom}\theta(X,\cdot)\cap\mathrm{Aff}K_{\theta,\{\psi=\infty\}}(X)\}$, and therefore by (6.18), $B\subset\{X\in N_\mu\}\cup\{Y\in N_\nu\}\cup\{Y\notin J_\theta(X)\}$. □

6.3 The maximal support probability

In order to prove the existence of a maximum support martingale transport plan, we introduce the minimization problem.

$$M := \sup_{\mathbb{P}\in\mathcal{M}(\mu,\nu)} \mu[G(\widehat{\mathrm{supp}}\mathbb{P}_X)]. \tag{6.19}$$

where we rely on the following measurability result whose proof is reported in Subsection 7.2.

Lemma 6.3. *For $\mathbb{P}\in\mathcal{P}(\Omega)$, the map $\widehat{\mathrm{supp}}\mathbb{P}_X$ is Borel-measurable, regardless of the choice of the kernel $\mathbb{P}_X$. Furthermore, the map $\widehat{\mathrm{supp}}(\mathbb{P}_X|_{\partial I(X)})$ is $\mu-$measurable.*

Proof of Proposition 3.4 (i) We procede in three steps.
<u>*Step 1:*</u> We first prove existence for the problem 6.19. Let $(\mathbb{P}^n)_{n\geqslant 1}\subset\mathcal{M}(\mu,\nu)$ be a maximizing sequence. Then the measure $\widehat{\mathbb{P}} := \sum_{n\geqslant 1}2^{-n}\,\mathbb{P}^n\in\mathcal{M}(\mu,\nu)$, and satisfies $\widehat{\mathrm{supp}}\mathbb{P}^n_X\subset\widehat{\mathrm{supp}}\widehat{\mathbb{P}}_X$ for all $n\geqslant 1$. Consequently $\mu[G(\widehat{\mathrm{supp}}_X\mathbb{P}^n_X)]\leqslant\mu[G(\widehat{\mathrm{supp}}\widehat{\mathbb{P}}_X)]$, and therefore $M = \mu[G(\widehat{\mathrm{supp}}\widehat{\mathbb{P}}_X)]$.
<u>*Step 2:*</u> We next prove that $\widehat{\mathrm{supp}}\mathbb{P}_X\subset\widehat{\mathrm{supp}}\widehat{\mathbb{P}}_X$, μ-a.s. for all $\mathbb{P}\in\mathcal{M}(\mu,\nu)$. Indeed, the measure $\overline{\mathbb{P}} := \frac{\widehat{\mathbb{P}}+\mathbb{P}}{2}\in\mathcal{M}(\mu,\nu)$ satisfies $M\geqslant\mu[G(\widehat{\mathrm{supp}}\overline{\mathbb{P}}_X)]\geqslant\mu[G(\widehat{\mathrm{supp}}\widehat{\mathbb{P}}_X)] = M$, implying that $G(\widehat{\mathrm{supp}}\overline{\mathbb{P}}_X) = G(\widehat{\mathrm{supp}}\widehat{\mathbb{P}}_X)$, $\mu-$a.s. The required result now follows from the inclusion $\widehat{\mathrm{supp}}\widehat{\mathbb{P}}_X\subset\widehat{\mathrm{supp}}\overline{\mathbb{P}}_X$.
<u>*Step 3:*</u> We finally prove that $S(X) := \widehat{\mathrm{supp}}\widehat{\mathbb{P}}_X = \mathrm{cl}\,I(X)$, $\mu-$a.s. By the previous step, we

have $\text{supp}(\mathbb{P}_X) \subset S(X)$, μ−a.s. Then $\{Y \notin S(X)\}$ is $\mathcal{M}(\mu,\nu)$−polar. By Theorem 3.7, we see that $\{Y \notin S(X)\} \subset \{X \in N_\mu\} \cup \{Y \in N_\nu\} \cup \{Y \notin J_\theta(X)\}$, or equivalently,

$$\{Y \in S(X)\} \quad \supset \quad \{X \notin N_\mu\} \cap \{Y \in J_\theta(X)\backslash N_\nu\}, \tag{6.20}$$

for some $N_\mu \in \mathcal{N}_\mu$, $N_\nu \in \mathcal{N}_\nu$, and $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$. Observe that (6.20) says that $J_\theta(X)\backslash N_\nu \subset S(X)$, μ-a.s. However, by 3.3 (ii), $I(X) \subset \text{conv}\big(\text{dom}\theta(X,\cdot)\backslash N_\nu\big)$, μ−a.s. Then, since $S(X)$ is closed and convex, we see that $\text{cl}\, I(X) \subset S(X)$.

To obtain the reverse inclusion, we recall from Theorem 3.3 (i) that $\{Y \in \text{cl}\, I(X)\}$, $\mathcal{M}(\mu,\nu)$−q.s. In particular $\widehat{\mathbb{P}}[Y \in \text{cl}\, I(X)] = 1$, implying that $S(X) \subset \text{cl}\, I(X)$, μ-a.s. as $\text{cl}\, I(X)$ is closed convex.

(ii) By the same argument as in (i), we may find $\widehat{\mathbb{P}}' \in \mathcal{M}(\mu,\nu)$ such that

$$M' := \sup_{\mathbb{P}\in\mathcal{M}(\mu,\nu)} \mu\Big[G\Big(\widehat{\text{supp}}(\mathbb{P}_X|_{\partial I(X)})\Big)\Big] = \mu\Big[G\Big(\widehat{\text{supp}}(\widehat{\mathbb{P}}'_X|_{\partial I(X)})\Big)\Big]. \tag{6.21}$$

We also have similarly that $\widehat{\text{supp}}(\mathbb{P}_X|_{\partial I(X)}) \subset \widehat{\text{supp}}(\widehat{\mathbb{P}}'_X|_{\partial I(X)})$, μ-a.s. for all $\mathbb{P} \in \mathcal{M}(\mu,\nu)$. Then we prove similarly that $S'(X) := \widehat{\text{supp}}(\widehat{\mathbb{P}}'_X|_{\partial I(X)}) = D(X)$, μ−a.s., where recall that D is the optimizer for (6.15). Indeed, by the previous step, we have $\widehat{\text{supp}}(\mathbb{P}_X|_{\partial I(X)}) \subset S'(X)$, μ−a.s. Then $\{Y \notin S'(X) \cup I(X)\}$ is $\mathcal{M}(\mu,\nu)$−polar. By Theorem 3.7, we see that $\{Y \notin S'(X) \cup I(X)\} \subset \{X \in N_\mu\} \cup \{Y \in N_\nu\} \cup \{Y \notin J_\theta(X) \cup I(X)\}$, or equivalently,

$$\{Y \in S'(X) \cup I(X)\} \quad \supset \quad \{X \notin N_\mu\} \cap \{Y \in J_\theta(X)\backslash N_\nu\}, \tag{6.22}$$

for some $N_\mu \in \mathcal{N}_\mu$, $N_\nu \in \mathcal{N}_\nu$, and $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$. Same than previously, we have $J_\theta(X)\backslash N_\nu\backslash I(X) \subset S'(X)$, μ-a.s. Then, since $S'(X)$ is closed and convex, we see that $D(X) \subset S'(X)$.

To obtain the reverse inclusion, we recall from Proposition 6.2 that $\{Y \in J(X)\}$, $\mathcal{M}(\mu,\nu)$−q.s. In particular $\widehat{\mathbb{P}}^*[Y \in I(X) \cup D(X)] = 1$, implying that $S'(X) \subset D(X)$, μ-a.s.

Finally, $\frac{\widehat{\mathbb{P}}+\widehat{\mathbb{P}}'}{2}$ is optimal for both problems (6.19), and (6.21). The remaining properties follow from Proposition 6.2. □

6.4 Properties of the J-mappings

Proof of Remark 3.5 Let $y \in \text{atom}(\nu)$, by the same argument as in the proof of Proposition 3.4, we may find $\widehat{\mathbb{P}}''_y \in \mathcal{M}(\mu,\nu)$ such that

$$M'' := \sup_{\mathbb{P}\in\mathcal{M}(\mu,\nu)} \mu\Big[\mathbb{P}_X\big[\{y\} \cap \text{cl}\, I(X)\big] > 0\Big] = \mu\Big[\widehat{\mathbb{P}}''_X\big[\{y\} \cap \text{cl}\, I(X)\big] > 0\Big]. \tag{6.23}$$

We denote $S''(X) := \text{supp}\widehat{\mathbb{P}}''_X|_{\text{Aff}\, I(X)\cap\{y\}}$. Recall that D_y is the notation for the optimizer of problem (6.16). We consider the set $N := \big\{Y \notin (\text{cl}\, I(X)\backslash\{y\}) \cup S''(X)\big\}$. N is polar as $Y \in \text{cl}\, I(X)$, q.s., and by definition of S''. Then $N \subset \{X \in N_\mu\} \cup \{Y \in N_\nu\} \cup \{Y \notin J_\theta(X)\}$, or equivalently,

$$\big\{Y \notin (\text{cl}\, I(X)\backslash\{y\}) \cup S''(X)\big\} \quad \supset \quad \{X \notin N_\mu\} \cap \{Y \in J_\theta(X)\backslash N_\nu\}, \tag{6.24}$$

for some $N_\mu \in \mathcal{N}_\mu$, $N_\nu \in \mathcal{N}_\nu$, and $\theta \in \widehat{\mathcal{T}}(\mu,\nu)$. Then $D_y(X) \subset J_\theta(X)\backslash N_\nu \subset \mathrm{cl}\, I(X)\backslash\{y\} \cup S''(X)$, $\mu-$a.s. Finally $D_y(X) \subset S''(X)$, $\mu-$a.s.

On the other hand, $S'' \subset D_y$, $\mu-$a.s., as if $\widehat{\mathbb{P}}''_X[\{y\}] > 0$, we have $\theta(X,y) < \infty$, $\mu-$a.s. at the corresponding points. Hence, $D_y(X) = S''(X)$, $\mu-$a.s. Now if we sum up the countable optimizers for $y \in \mathrm{atom}(\nu)$, with the previous optimizers, then the probability $\widehat{\mathbb{P}}$ we get is an optimizer for (6.19), (6.21), and (6.23), for all $y \in \mathbb{R}^d$ (the optimum is 0 if it is not an atom of ν). The remaining properties follow from Proposition 6.2. $\square$

7 Measurability of the irreducible components

7.1 Measurability of G

Proof of Lemma 3.1 (ii) As $\mathbb{R}^d$ is locally totally bounded, the Wijsman topology is locally equivalent to the Hausdorff topology[4], i.e. as $n \to \infty$, $K_n \longrightarrow K$ for the Wijsman topology if and only if $K_n \cap B_M \longrightarrow K \cap B_M$ for the Hausdorff topology, for all $M \geqslant 0$.

We first prove that $K \longmapsto \dim \mathrm{Aff} K$ is a lower semi-continuous map $\mathcal{K} \to \mathbb{R}$. Let $(K_n)_{n\geqslant 1} \subset \mathcal{K}$ with dimension $d_n \leqslant d' \leqslant d$ converging to K. We consider $A_n := \mathrm{Aff}\, K_n$. As A_n is a sequence of affine spaces, it is homeomorphic to a $d+1$-uplet. Observe that the convergence of K_n allow us to chose this $d+1$-uplet to be bounded. Then up to taking a subsequence, we may suppose that A_n converges to an affine subspace A of dimension less than d'. By continuity of the inclusion under the Wijsman topology, $K \subset A$ and $\dim K \leqslant \dim A \leqslant d'$.

We next prove that the mapping $K \mapsto g_K(K)$ is continuous on $\{\dim K = d'\}$ for $0 \leqslant d' \leqslant d$, which implies the required measurability. Let $(K_n)_{n\geqslant 1} \subset \mathcal{K}$ be a sequence with constant dimension d', converging to a $d'-$dimensional subset, K in $\mathcal{K}$. Define $A_n := \mathrm{Aff} K_n$ and $A := \mathrm{Aff} K$, A_n converges to A as for any accumulation set A' of A_n, $K \subset A'$ and $\dim A' = \dim A$, implying that $A' = A$. Now we consider the map $\phi_n : A_n \to A$, $x \mapsto proj_A(x)$. For all $M > 0$, it follows from the compactness of the closed ball B_M that ϕ_n converges uniformly to identity as $n \to \infty$ on B_M. Then, $\phi_n(K_n) \cap B_M \longrightarrow K \cap B_M$ as $n \to \infty$, and therefore $\lambda_A[\phi_n(K_n \cap B_M)\backslash K] + \lambda_A[K\backslash\phi_n(K_n) \cap B_M] \longrightarrow 0$. As the Gaussian density is bounded, we also have

$$g_A[\phi_n(K_n \cap B_M)] \longrightarrow g_A[K \cap B_M].$$

We next compare $g_A[\phi_n(K_n \cap B_M)]$ to $g_{K_n}(K_n \cap B_M)$. As (ϕ_n) is a sequence of linear functions that converges uniformly to identity, we may assume that ϕ_n is a C^1-diffeomorphism. Furthermore, its constant Jacobian J_n converges to 1 as $n \to \infty$. Then,

$$\int_{K_n \cap B_N} \frac{e^{-|\phi_n(x)|^2/2}}{(2\pi)^{d'/2}} \lambda_{K_n}(dx) = \int_{\phi_n(K_n \cap B_M)} \frac{e^{-|y|^2/2} J_n^{-1}}{(2\pi)^{d'/2}} \lambda_A(dy) = J_n^{-1} g_A[\phi_n(K_n \cap B_M)].$$

[4]The Haussdorff distance on the collection of all compact subsets of a compact metric space $(\mathcal{X}, \mathrm{d})$ is defined by $d_H(K_1,K_2) = \sup_{x\in\mathcal{X}} |\mathrm{dist}(x,K_1) - \mathrm{dist}(x,K_2)|$, for $K_1, K_2 \subset \mathcal{X}$, compact subsets.

As the Gaussian distribution function is 1-Lipschitz, we have

$$\Big|\int_{K_n\cap B_M}\frac{e^{-|\phi_n(x)|^2/2}}{(2\pi)^{d'/2}}\lambda_{K_n}(dx)-g_{K_n}(K_n\cap B_M)\Big| \;\leqslant\; \lambda_{K_n}[K_n\cap B_M]|\phi_n-Id_A|_\infty,$$

where $|\cdot|_\infty$ is taken on $K_n\cap B_M$. Now for arbitrary $\epsilon>0$, by choosing M sufficiently large so that $g_V[V\backslash B_M]\leqslant\epsilon$ for any d'−dimensional subspace V, we have

$$\begin{aligned}|g_{K_n}[K_n]-g_K[K]| &\leqslant |g_{K_n}[K_n\cap B_M]-g_A[K\cap B_M]|+2\epsilon\\ &\leqslant \Big|g_{K_n}[K_n\cap B_M]-\int_{K_n\cap B_M}C\exp(\frac{-|\phi_n(x)|^2}{2})\lambda_{K_n}(dx)\Big|\\ &\qquad\qquad+\big|J_n^{-1}g_A[\phi_n(K_n\cap B_M)]-g_A[K\cap B_M]\big|+2\epsilon\;\leqslant\;4\epsilon,\end{aligned}$$

for n sufficiently large, by the previously proved convergence. Hence $G_{d'}:=G\big|_{\dim^{-1}\{d'\}}$ is continuous, implying that $G:K\longmapsto\sum_{d'=0}^d\mathbf{1}_{\dim^{-1}\{d'\}}(K)G_{d'}(K)$ is Borel-measurable. □

7.2 Further measurability of set-valued maps

This subsection is dedicated to the proof of Lemmas 3.1 (i), 6.1, and 6.3. In preparation for the proofs, we start by giving some lemmas on measurability of set-valued maps.

Lemma 7.1. *Let* $(F_n)_{n\geqslant 1}\subset\mathbb{L}^0(\mathbb{R}^d,\mathcal{K})$*. Then* $\mathrm{cl}\cup_{n\geqslant1}F_n$ *and* $\cap_{n\geqslant1}F_n$ *are measurable.*

Proof. The measurability of the union is a consequence of Propositions 2.3 and 2.6 in Himmelberg [12]. The measurability of the intersection follows from the fact that $\mathbb{R}^d$ is σ-compact, together with Corollary 4.2 in [12]. □

Lemma 7.2. *Let* $F\in\mathbb{L}^0(\mathbb{R}^d,\mathcal{K})$*. Then,* $\mathrm{cl\,conv}F$*,* $\mathrm{Aff}F$*, and* $\mathrm{cl\,rf}_X\mathrm{cl\,conv}F$ *are measurable.*

Proof. The measurability of $\mathrm{cl\,conv}F$ is a direct application of Theorem 9.1 in [12].

We next verify that $\mathrm{Aff}F$ is measurable. Since the values of F are closed, we deduce from Theorem 4.1 in Wagner [21], that we may find a measurable $x\longmapsto y(x)$, such that $y(x)\in F(x)$ if $F(x)\neq\varnothing$, for all $x\in\mathbb{R}^d$. Then we may write $\mathrm{Aff}F(x)=\mathrm{cl\,conv\,cl}\,\cup_{q\in\mathbb{Q}}\big(y(x)+q\left(F(x)-y(x)\right)\big)$ for all $x\in\mathbb{R}^d$. The measurability follows from Lemmas 7.1, together with Step (i) of the present proof.

We finally justify that $\mathrm{cl\,rf}_X\mathrm{cl\,conv}F$ is measurable. We may assume that F takes convex values. By convexity, we may reduce the definition of rf_x to a sequential form:

$$\begin{aligned}\mathrm{cl\,rf}_xF(x) &= \mathrm{cl}\cup_{n\geqslant1}\left\{y\in\mathbb{R}^d, y+\frac1n(y-x)\in F(x)\text{ and }x-\frac1n(y-x)\in F(x)\right\}\\ &= \mathrm{cl}\cup_{n\geqslant1}\left[\left\{y\in\mathbb{R}^d, y+\frac1n(y-x)\in F(x)\right\}\cap\left\{y\in\mathbb{R}^d, x-\frac1n(y-x)\in F(x)\right\}\right]\\ &= \mathrm{cl}\cup_{n\geqslant1}\left[\left(\frac{1}{n+1}x+\frac{n}{n+1}F(x)\right)\cap\left(-(n+1)x-nF(x)\right)\right],\end{aligned}$$

so that the required measurability follows from Lemma 7.1. □

We denote $\mathcal{S}$ the set of finite sequences of positive integers, and Σ the set of infinite sequences of positive integers. Let $s \in \mathcal{S}$, and $\sigma \in \Sigma$. We shall denote $s < \sigma$ whenever s is a prefix of σ.

Lemma 7.3. *Let $(F_s)_{s\in\mathcal{S}}$ be a family of universally measurable functions $\mathbb{R}^d \longrightarrow \mathcal{K}$ with convex image. Then the mapping* $\mathrm{cl\,conv}\big(\cup_{\sigma\in\Sigma}\cap_{s<\sigma}F_s\big)$ *is universally measurable.*

Proof. Let $\mathcal{U}$ the collection of universally measurable maps from $\mathbb{R}^d$ to $\mathcal{K}$ with convex image. For an arbitrary $\gamma \in \mathcal{P}(\mathbb{R}^d)$, and $F : \mathbb{R}^d \longrightarrow \mathcal{K}$, we introduce the map

$$\gamma G^*[F] := \inf_{F\subset F'\in\mathcal{U}} \gamma G[F'], \quad \text{where} \quad \gamma G[F'] := \gamma\big[G\big(F'(X)\big)\big] \text{ for all } F' \in \mathcal{U}.$$

Clearly, γG and γG^* are non-decreasing, and it follows from the dominated convergence theorem that γG, and thus γG^*, are upward continuous.

<u>*Step 1:*</u> In this step we follow closely the line of argument in the proof of Proposition 7.42 of Bertsekas and Shreve [5]. Set $F := \mathrm{cl\,conv}\big(\cup_{\sigma\in\Sigma}\cap_{s<\sigma}F_s\big)$, and let $(\bar{F}_n)_n$ a minimizing sequence for $\gamma G^*[F]$. Notice that $F \subset \bar{F} := \cap_{n\geqslant 1}\bar{F}_n \in \mathcal{U}$, by Lemma 7.1. Then $\bar{F}$ is a minimizer of $\gamma G^*[F]$.

For $s, s' \in S$, we denote $s \leqslant s'$ if they have the same length $|s| = |s'|$, and $s_i \leqslant s'_i$ for $1 \leqslant i \leqslant |s|$. For $s \in S$, let

$$R(s) := \text{cl conv} \cup_{s'\leqslant s} \cup_{\sigma>s'} \cap_{s''<\sigma} F_{s''} \quad \text{and} \quad K(s) := \text{cl conv} \cup_{s'\leqslant s} \cap_{j=1}^{|s'|} F_{s'_1,\ldots,s'_j}.$$

Notice that $K(s)$ is universally measurable, by Lemmas 7.1 and 7.2, and

$$R(s) \subset K(s), \quad \text{cl} \cup_{s_1\geqslant 1} R(s_1) = F, \quad \text{and} \quad \text{cl} \cup_{s_k\geqslant 1} R(s_1, ..., s_{k-1}, s_k) = R(s_1, ..., s_{k-1}).$$

By the upwards continuity of γG^*, we may find for all $\epsilon > 0$ a sequence $\sigma^\epsilon \in \Sigma$ s.t.

$$\gamma G^*[F] \leqslant \gamma G^*[R(\sigma_1^\epsilon)] + 2^{-1}\epsilon, \quad \text{and} \quad \gamma G^*[R(\underline{\sigma}_{k-1})] \leqslant \gamma G^*[R(\underline{\sigma}_k)] + 2^{-k}\epsilon, \quad k \geqslant 1,$$

with the notation $\underline{\sigma}_k^\varepsilon := (\sigma_1^\epsilon, \ldots, \sigma_k^\varepsilon)$. Recall that the minimizer $\overline{F}$ and $K(s)$ are in $\mathcal{U}$ for all $s \in \mathcal{S}$. We then define the sequence $K_k^\epsilon := \overline{F} \cap K(\underline{\sigma}_k^\epsilon) \in \mathcal{U}$, $k \geqslant 1$, and we observe that

$$(K_k^\epsilon)_{k\geqslant 1} \text{ decreasing}, \ \underline{F}^\epsilon := \cap_{k\geqslant 1} K_k^\epsilon \subset F, \text{ and } \gamma G[K_k^\epsilon] \geqslant \gamma G^*[F] - \epsilon = \gamma G[\overline{F}] - \epsilon, \quad (7.25)$$

by the fact that $R(\underline{\sigma}_k^\epsilon) \subset K_k^\epsilon$. We shall prove in Step 2 that, for an arbitrary $\alpha > 0$, we may find $\varepsilon = \varepsilon(\alpha) \leqslant \alpha$ such that (7.25) implies that

$$\gamma G[\underline{F}^\epsilon] \geqslant \inf_{k\geqslant 1} \gamma G[K_k^\epsilon] - \alpha \geqslant \gamma G[\overline{F}] - \epsilon - \alpha. \quad (7.26)$$

Now let $\alpha = \alpha_n := n^{-1}$, $\varepsilon_n := \epsilon(\alpha_n)$, and notice that $\underline{F} := \mathrm{cl\,conv} \cup_{n\geqslant 1} \underline{F}^{\epsilon_n} \in \mathcal{U}$, with $\underline{F}^{\epsilon_n} \subset \underline{F} \subset F \subset \overline{F}$, for all $n \geqslant 1$. Then, it follows from (7.26) that $\gamma G[\underline{F}] = \gamma G[\overline{F}]$, and therefore $\underline{F} = F = \overline{F}$, $\gamma-$a.s. In particular, F is $\gamma-$measurable, and we conclude that $F \in \mathcal{U}$ by the arbirariness of $\gamma \in \mathcal{P}(\mathbb{R}^d)$.

Step 2: It remains to prove that, for an arbitrary $\alpha > 0$, we may find $\varepsilon = \varepsilon(\alpha) \leqslant \alpha$ such that (7.25) implies (7.26). Notice that this is the point where we have to deviate from the argument of [5] because γG is not downwards continuous, as the dimension can jump down.

Set $A_n := \{G\big(\overline{F}(X)\big) - \dim \overline{F}(X) \leqslant 1/n\}$, and notice that $\cap_{n\geqslant 1} A_n = \varnothing$. Let $n_0 \geqslant 1$ such that $\gamma[A_{n_0}] \leqslant \frac{1}{2}\frac{\alpha}{d+1}$, and set $\epsilon := \frac{1}{2}\frac{1}{n_0}\frac{\alpha}{d+1} > 0$. Then, it follows from (7.25) that

$$\begin{aligned}
\gamma\big[\inf_n G(K_n^\epsilon) - \dim \overline{F} \leqslant 0\big] &\leqslant \gamma\big[\inf_n G(K_n^\epsilon) - G(\overline{F}) \leqslant n_0^{-1}\big] + \gamma\big[G(\overline{F}) - \dim \overline{F} \leqslant -n_0^{-1}\big] \\
&\leqslant n_0\big(\gamma\big[G(\overline{F})\big] - \gamma\big[\inf_n G(K_n^\epsilon)\big]\big) + \gamma\big[A_{n_0}\big] \\
&= n_0\big(\gamma\big[G(\overline{F})\big] - \inf_n \gamma\big[G(K_n^\epsilon)\big]\big) + \gamma\big[A_{n_0}\big] \\
&\leqslant n_0\epsilon + \frac{1}{2}\frac{\alpha}{d+1} \;=\; \frac{\alpha}{d+1},
\end{aligned} \tag{7.27}$$

where we used the Markov inequality and the monotone convergence theorem. Then:

$$\begin{aligned}
\gamma\big[\inf_n G(K_n^\epsilon) - G\big(\underline{F}^\epsilon\big)\big] &\leqslant \gamma\Big[\mathbf{1}_{\{\inf_n G(K_n^\epsilon) - \dim \overline{F} \leqslant 0\}}\big(\inf_n G(K_n^\epsilon) - G\big(\underline{F}^\epsilon\big)\big) \\
&\qquad + \mathbf{1}_{\{\inf_n G(K_n^\epsilon) - \dim \overline{F} > 0\}}\big(\inf_n G(K_n^\epsilon) - G\big(\underline{F}^\epsilon\big)\big)\Big] \\
&\leqslant \gamma\Big[(d+1)\mathbf{1}_{\{\inf_n G(K_n^\epsilon) - \dim \overline{F} \leqslant 0\}} \\
&\qquad + \mathbf{1}_{\{\inf_n G(K_n^\epsilon) - \dim \overline{F} > 0\}}\big(\inf_n G(K_n^\epsilon) - G\big(\underline{F}^\epsilon\big)\big)\Big].
\end{aligned}$$

We finally note that $\inf_n G(K_n^\epsilon) - G\big(\underline{F}^\epsilon\big) = 0$ on $\{\inf_n G(K_n^\epsilon) - \dim \overline{F} > 0\}$. Then (7.26) follows by substituting the estimate in (7.27). □

Proof of Lemma 3.1 (i) We consider the mappings $\theta : \Omega \to \bar{\mathbb{R}}_+$ such that $\theta = \sum_{k=1}^n \lambda_k \mathbf{1}_{C_k^1 \times C_k^2}$ where $n \in \mathbb{N}$, the λ_k are non-negative numbers, and the C_k^1, C_k^2 are closed convex subsets of $\mathbb{R}^d$. We denote the collection of all these mappings $\mathcal{F}$. Notice that $\mathrm{cl}\,\mathcal{F}$ for the pointwise limit topology contains all $\mathbb{L}^0_+(\Omega)$. Then for any $\theta \in \mathbb{L}^0_+(\Omega)$, we may find a family $(\theta_s)_{s\in\Sigma} \subset \mathcal{F}$, such that $\theta = \inf_{\sigma\in\Sigma}\sup_{s<\sigma}\theta_s$. For $\theta \in \mathbb{L}^0_+(\Omega)$, and $M \geqslant 0$, we denote $F_\theta : x \longmapsto \mathrm{cl}\,\mathrm{conv}\,\mathrm{dom}\theta(x,\cdot)$, and $F_{\theta,M} : x \longmapsto \mathrm{cl}\,\mathrm{conv}\,\theta(x,\cdot)^{-1}([0,M])$. Notice that $F_\theta = \mathrm{cl}\,\cup_{n\geqslant 1} F_{\theta,n}$. Notice as well that $F_{\theta,M}$ is Borel measurable for $\theta \in \mathcal{F}$, and $M \geqslant 0$, as it takes values in a finite set, from a finite number of measurable sets. Let $\theta \in \mathbb{L}^0_+(\Omega)$, we consider the associated family $(\theta_s)_{s\in\Sigma} \subset \mathcal{F}$, such that $\theta = \inf_{\sigma\in\Sigma}\sup_{s<\sigma}\theta_s$. Notice that $F_{\theta,M} = \mathrm{cl}\,\mathrm{conv}\big(\cup_{\sigma\in\Sigma}\cap_{s<\sigma}F_{\theta_s,M}\big)$ is universally measurable by Lemma 7.3, thus implying the universal measurability of $F_\theta = \mathrm{cl}\,\mathrm{dom}\theta(X,\cdot)$ by Lemma 7.1.

In order to justify the measurability of $\mathrm{dom}_X\theta$, we now define

$$F_\theta^0 := F_\theta \quad \text{and} \quad F_\theta^k := \mathrm{cl}\,\mathrm{conv}(\mathrm{dom}\theta(X,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_X F_\theta^{k-1}),\; k \geqslant 1.$$

Note that $F_\theta^k = \mathrm{cl}\,\cup_{n\geqslant 1}\big(\mathrm{cl}\,\mathrm{conv}\,\cup_{\sigma\in\Sigma}\cap_{s<\sigma}F_{\theta_s,n} \cap \mathrm{Aff}\,\mathrm{rf}_x F_\theta^{k-1}\big)$. Then, as F_θ^0 is universally measurable, we deduce that $\big(F_\theta^k\big)_{k\geqslant 1}$ are universally measurable, by Lemmas 7.2 and 7.3.

As $\mathrm{dom}_X\theta$ is convex and relatively open, the required measurability follows from the claim:

$$F_\theta^d \quad = \quad \mathrm{cl}\,\mathrm{dom}_X\theta.$$

To prove this identity, we start by observing that $F_\theta^k(x) \supset \mathrm{cl}\,\mathrm{dom}_x\theta$. Since the dimension cannot decrease more than d times, we have $\mathrm{Aff}\,\mathrm{rf}_x F_\theta^d(x) = \mathrm{Aff} F_\theta^d(x)$ and

$$F_\theta^{d+1}(x) = \mathrm{cl}\,\mathrm{conv}\big(\mathrm{dom}\theta(x,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_x F_\theta^d(x)\big) = \mathrm{cl}\,\mathrm{conv}\big(\mathrm{dom}\theta(x,\cdot) \cap \mathrm{Aff}\mathrm{rf}_x F_\theta^{d-1}(x)\big) = F_\theta^d(x).$$

i.e. $(F_\theta^{d+1})_k$ is constant for $k \geqslant d$. Consequently,

$$\begin{aligned}\dim \mathrm{rf}_x\mathrm{conv}(\mathrm{dom}\theta(x,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_x F_\theta^d(x)) \quad &= \quad \dim F_\theta^d(x) \\ &\geqslant \quad \dim \mathrm{conv}(\mathrm{dom}\theta(x,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_x F_\theta^d(x)).\end{aligned}$$

As $\dim \mathrm{conv}\big(\mathrm{dom}\theta(x,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_x F_\theta^d(x)\big) \geqslant \dim \mathrm{rf}_x\mathrm{conv}\big(\mathrm{dom}\theta(x,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_x F_\theta^d(x)\big)$, we have equality of the dimension of $\mathrm{conv}\big(\mathrm{dom}\theta(x,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_x F_\theta^d(x)\big)$ with its rf_x. Then it follows from Proposition 2.1 (ii) that $x \in \mathrm{ri}\,\mathrm{conv}\big(\mathrm{dom}\theta(x,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_x F_\theta^d(x)\big)$, and therefore:

$$\begin{aligned}F_\theta^d(x) \quad &= \quad \mathrm{cl}\,\mathrm{conv}\big(\mathrm{dom}\theta(x,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_x F_\theta^d(x)\big) = \mathrm{cl}\,\mathrm{ri}\,\mathrm{conv}\big(\mathrm{dom}\theta(x,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_x F_\theta^d(x)\big) \\ &= \quad \mathrm{cl}\,\mathrm{rf}_x\mathrm{conv}\big(\mathrm{dom}\theta(x,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_x F_\theta^d(x)\big) \ \subset \ \mathrm{cl}\,\mathrm{dom}_x\theta.\end{aligned}$$

Hence $F_\theta^d(x) = \mathrm{cl}\,\mathrm{dom}_x\theta$.

Finally, $K_{\theta,A} = \mathrm{dom}_X(\theta + \infty 1_{\mathbb{R}^d\times A})$ is universally measurable by the universal measurability of dom_X. □

Proof of Lemma 6.1 We may find $(F_n)_{n\geqslant 1}$, Borel-measurable with finite image, converging $\gamma-$a.s. to F. We denote $N_\gamma \in \mathcal{N}_\gamma$, the set on which this convergence does not hold. If for $\epsilon > 0$, we denote $F_k^\epsilon(X) := \{y \in \mathbb{R}^d : dist(y, F_k(X) \leqslant \epsilon)\}$, we have

$$F(x) = \cap_{i\geqslant 1} \liminf_{n\to\infty} F_n^{1/i}(x), \quad \text{for all} \quad x \notin N_\gamma.$$

Then, as $\mathbf{1}_{Y\in F(X)}\mathbf{1}_{X\notin N_\gamma} = \inf_{i\geqslant 1}\liminf_{n\to\infty}\mathbf{1}_{Y\in F_n^{1/i}(X)}\mathbf{1}_{X\notin N_\gamma}$, the Borel-measurability of this function follows from the Borel-measurability of each $\mathbf{1}_{Y\in F_n^{1/i}(X)}$.

Now we suppose that $X \in \mathrm{ri}F(X)$ convex, $\gamma-$a.s. Up to redefining N_γ, we may suppose that this property holds on N_γ^c, then $\partial F(x) = \cap_{n\geqslant 1} F(x)\backslash\big(x + \frac{n}{n+1}(F(x) - x)\big)$, for $x \notin N_\gamma$. We denote $a := \mathbf{1}_{Y\in F(X)}\mathbf{1}_{X\notin N_\gamma}$. The result follows from the identity $\mathbf{1}_{Y\in\partial F(X)}\mathbf{1}_{X\notin N_\gamma} = a - \sup_{n\geqslant 1} a\big(X, X + \frac{n}{n+1}(Y - X)\big)$. □

Proof of Lemma 6.3 Let $\mathcal{K}_\mathbb{Q} := \{K = \mathrm{conv}(x_1, \ldots, x_n) : n \in \mathbb{N}, (x_i)_{i\leqslant n} \subset \mathbb{Q}^d\}$. Then

$$\widehat{\mathrm{supp}}\mathbb{P}_x \quad = \quad \mathrm{cl}\ \cup_{N\geqslant 1} \cap\{K \in \mathcal{K}_\mathbb{Q} : \widehat{\mathrm{supp}}\mathbb{P}_x \cap B_N \subset K\} \ = \ \mathrm{cl}\ \cup_{N\geqslant 1} \cap_{K\in\mathcal{K}_\mathbb{Q}} F_K^N(x),$$

where $F_K^N(x) := K$ if $\mathbb{P}_x(B_N \cap K) = \mathbb{P}_x(B_N)$, and $F_K^N(x) := \mathbb{R}^d$ otherwise. As for any $K \in \mathcal{K}_\mathbb{Q}$ and $N \geqslant 1$, the map $\mathbb{P}_X(B_N \cap K) - \mathbb{P}_X(B_N)$ is measurable, and therefore F_K^N is measurable. The required measurability result follows from lemma 7.1.

Now, in order to get the measurability of $\widehat{\text{supp}}(\mathbb{P}_X|_{\partial I(X)})$, we have in the same way

$$\widehat{\text{supp}}(\mathbb{P}_X|_{\partial I(X)}) \quad = \quad \text{cl} \cup_{n\geqslant 1} \cap_{K\in\mathcal{K}_\mathbb{Q}} F'^N_K(x),$$

where $F'^N_K(x) := K$ if $\mathbb{P}_x(\partial I(x) \cap B_N \cap K) = \mathbb{P}_x(\partial I(x) \cap B_N)$, and $F'^N_K(x) := \mathbb{R}^d$ otherwise. As $\mathbb{P}_X(\partial I(X) \cap B_N \cap K) = \mathbb{P}_X[\mathbf{1}_{Y\in\partial I(X)}\mathbf{1}_{X\notin N_\mu}\mathbf{1}_{Y\notin B_N\cap K}]$, $\mu-$a.s., where $N_\mu \in \mathcal{N}_\mu$ is taken from Lemma 6.1, $\mathbb{P}_X(\partial I(X) \cap B_N \cap K)$ is $\mu-$measurable, as equal $\mu-$a.s. to a Borel function. Then similarly, $\mathbb{P}_X(\partial I(X) \cap B_N \cap K) - \mathbb{P}_X(\partial I(X) \cap B_N)$ is $\mu-$measurable, and therefore $\widehat{\text{supp}}(\mathbb{P}_X|_{\partial I(X)})$ is $\mu-$measurable. □

8 Properties of tangent convex functions

8.1 x-invariance of the y-convexity for tangent convex functions

We first report a convex analysis lemma.

Lemma 8.1. *Let $f : \mathbb{R}^d \to \bar{\mathbb{R}}$ be convex finite on some convex open subset $U \subset \mathbb{R}^d$. We denote $f_* : \mathbb{R}^d \to \bar{\mathbb{R}}$ the lower-semicontinuous envelop of f on U, then*

$$f_*(y) = \lim_{\epsilon\searrow 0} f\big(\epsilon x + (1-\epsilon)y\big), \quad \text{for all} \quad (x, y) \in U \times \text{cl}\, U.$$

Proof. f_* is the lower semi-continuous envelop of f on U, i.e. the lower semi-continuous envelop of $f' := f + \infty\mathbf{1}_{U^c}$. Notice that f' is convex $\mathbb{R}^d \longrightarrow \mathbb{R} \cup \{\infty\}$. Then by Proposition 1.2.5 in Chapter IV of [13], we get the result as $f = f'$ on U. □

Proof of Proposition 2.10 The result is obvious in $\mathbf{T}(\mathfrak{C}_1)$, as the affine part depending on x vanishes. We may use $N_\nu = \varnothing$. Now we denote $\mathcal{T}$ the set on mappings in Θ_μ such that this identity is verified. Then we have $\mathbf{T}(\mathfrak{C}_1) \subset \mathcal{T}$.

We prove that $\mathcal{T}$ is $\mu\otimes$pw$-$Fatou closed. Let $(\theta_n)_n$ be a sequence in $\mathcal{T}$ converging $\mu\otimes$pw to $\theta \in \Theta_\mu$. We denote N_μ, the set in $\mathcal{N}_\mu$ such that the identity does not hold for θ_n, for some n, and such that the $\mu\otimes$pw convergence holds. Let $x_1, x_2 \notin N_\mu$, and $\bar{y} \in \text{dom}_{x_1}\theta \cap \text{dom}_{x_2}\theta$. Let $y_1, y_2 \in \text{dom}_{x_1}\theta$, such that we have the convex combination $\bar{y} = \lambda y_1 + (1-\lambda)y_2$, and $0 \leqslant \lambda \leqslant 1$. Then for $i = 1, 2$, $\theta_n(x_1, y_i) \longrightarrow \theta(x_1, y_i)$, and $\theta_n(x_1, \bar{y}) \longrightarrow \theta(x_1, \bar{y})$, as $n \to \infty$. Using the fact that $\theta_n \in \mathcal{T}$, for all n, we have

$$\lambda\theta_n(x_1, y_1)+(1-\lambda)\theta_n(x_1, y_2)-\theta_n(x_1, \bar{y}) = \lambda\theta_n(x_2, y_1)+(1-\lambda)\theta_n(x_2, y_2)-\theta_n(x_2, \bar{y}) \geqslant 0. \quad (8.28)$$

Taking the limit $n \to \infty$ gives that $\underline{\theta}_\infty(x_2, y_i) < \infty$, and $y_i \in \text{dom}\underline{\theta}_\infty(x_2, \cdot)$. $\bar{y}$ is interior to $\text{dom}_{x_1}\theta$, then for any $y \in \text{dom}_{x_1}\theta$, $y' := \bar{y} + \frac{\epsilon}{1-\epsilon}(\bar{y} - y) \in \text{dom}_{x_1}\theta$ for $0 < \epsilon < 1$ small enough. Then $\bar{y} = \epsilon y + (1-\epsilon)y'$. As we may chose any $y \in \text{dom}_{x_1}\theta$, we have $\text{dom}_{x_1}\theta \subset \text{dom}\underline{\theta}_\infty(x_2, \cdot)$. Then, we have

$$\text{rf}_{x_2}\text{conv}(\text{dom}_{x_1}\theta \cup \text{dom}_{x_2}\theta) \subset \text{rf}_{x_2}\text{conv}\,\text{dom}\big(\underline{\theta}_\infty(x_2, \cdot)\big) = \text{dom}_{x_2}\theta. \quad (8.29)$$

By Lemma 9.1, as $\mathrm{dom}_{x_1}\theta \cap \mathrm{dom}_{x_2}\theta \neq \emptyset$, $\mathrm{conv}(\mathrm{dom}_{x_1}\theta \cup \mathrm{dom}_{x_2}\theta) = \mathrm{riconv}(\mathrm{dom}_{x_1}\theta \cup \mathrm{dom}_{x_2}\theta)$. In particular, $\mathrm{conv}(\mathrm{dom}_{x_1}\theta \cup \mathrm{dom}_{x_2}\theta)$ is relatively open and contains x_2, and therefore $\mathrm{rf}_{x_2}\mathrm{conv}(\mathrm{dom}_{x_1}\theta \cup \mathrm{dom}_{x_2}\theta) = \mathrm{conv}(\mathrm{dom}_{x_1}\theta \cup \mathrm{dom}_{x_2}\theta)$. Finally, by (8.29), $\mathrm{dom}_{x_1}\theta \subset \mathrm{dom}_{x_2}\theta$. As there is a symmetry between x_1, and x_2, we have $\mathrm{dom}_{x_1}\theta = \mathrm{dom}_{x_2}\theta$. Then we may go to the limit in equation (8.28):

$$\lambda\theta(x_1, y_1) + (1-\lambda)\theta(x_1, y_2) - \theta(x_1, \bar{y}) = \lambda\theta(x_2, y_1) + (1-\lambda)\theta(x_2, y_2) - \theta(x_2, \bar{y}) \geqslant 0. \quad (8.30)$$

Now, let $y_1, y_2 \in \mathbb{R}^d$, such that we have the convex combination $\bar{y} = \lambda y_1 + (1-\lambda)y_2$, and $0 \leqslant \lambda \leqslant 1$. we have three cases to study.
Case 1: $y_i \notin \mathrm{cl\,dom}_{x_1}\theta$ for some $i = 1, 2$. Then, as the average $\bar{y}$ of the y_i is in $\mathrm{dom}_{x_1}\theta$, by Proposition 2.1 (ii), me may find $i' = 1, 2$ such that $y_{i'} \notin \mathrm{conv\,dom}\theta(x_1, \cdot)$. Then $\lambda\theta(x_1, y_1) + (1-\lambda)\theta(x_1, y_2) - \theta(x_1, \bar{y}) = \infty \geqslant 0$. As $\mathrm{dom}_{x_1}\theta = \mathrm{dom}_{x_2}\theta$, we may apply the same reasoning to x_2, we get $\lambda\theta(x_1, y_1) + (1-\lambda)\theta(x_2, y_2) - \theta(x_2, \bar{y}) = \infty \geqslant 0$. We get the result.
Case 2: $y_1, y_2 \in \mathrm{dom}_{x_1}\theta$. This case is (8.30).
Case 3: $y_1, y_2 \in \mathrm{cl\,dom}_{x_1}\theta$. The problem arises here if some y_i is in the border $\partial\mathrm{dom}_{x_1}\theta$. Let $x \notin N_\mu$, we denote the lower semi-continuous envelop of $\theta(x, \cdot)$ in $\mathrm{cl\,dom}_x\theta$, $\theta_*(x, y) := \lim_{\epsilon \searrow 0} \theta(x, \epsilon x + (1-\epsilon)y')$, for $y \in \mathrm{cl\,dom}_x\theta$, where the latest equality follows from Lemma 8.1 together with that fact that $\theta(x, \cdot)$ is convex on $\mathrm{dom}_x\theta$. Let $y \in \mathrm{cl\,dom}_{x_1}\theta$, for $1 \geqslant \epsilon > 0$, $y^\epsilon := \epsilon x_1 + (1-\epsilon)y \in \mathrm{dom}_{x_1}\theta$. By (8.28), $(1-\epsilon)\theta_n(x_1, y) - \theta_n(x_1, y^\epsilon) = (1-\epsilon)\theta_n(x_2, y) - \theta_n(x_2, y^\epsilon)$. Taking the $\liminf$, we have $(1-\epsilon)\theta(x_1, y) - \theta(x_1, y^\epsilon) = (1-\epsilon)\theta(x_2, y) - \theta(x_2, y^\epsilon)$. Now making $\epsilon \searrow 0$, we have $\theta(x_1, y) - \theta_*(x_1, y) = \theta(x_2, y) - \theta_*(x_2, y)$. Then the jump of $\theta(x, \cdot)$ in y is independent of $x = x_1$ or x_2. Now for $1 \geqslant \epsilon > 0$, by (8.30)

$$\lambda\theta(x_1, y_1^\epsilon) + (1-\lambda)\theta(x_1, y_2^\epsilon) - \theta(x_1, \bar{y}^\epsilon) = \lambda\theta(x_2, y_1^\epsilon) + (1-\lambda)\theta(x_2, y_2^\epsilon) - \theta(x_2, \bar{y}^\epsilon) \geqslant 0.$$

By going to the limit $\epsilon \searrow 0$, we get

$$\lambda\theta_*(x_1, y_1) + (1-\lambda)\theta_*(x_1, y_2) - \theta_*(x_1, \bar{y}) = \lambda\theta_*(x_2, y_1) + (1-\lambda)\theta_*(x_2, y_2) - \theta_*(x_2, \bar{y}) \geqslant 0.$$

As the (nonnegative) jumps do not depend on $x = x_1$ or x_2, we finally get

$$\lambda\theta(x_1, y_1) + (1-\lambda)\theta(x_1, y_2) - \theta(x_1, \bar{y}) = \lambda\theta(x_2, y_1) + (1-\lambda)\theta(x_2, y_2) - \theta(x_2, \bar{y}) \geqslant 0.$$

Finally, $\mathcal{T}$ is $\mu\otimes$pw$-$Fatou closed, and convex. $\widehat{\mathcal{T}}_1 \subset \mathcal{T}$. As the result is clearly invariant when the function is multiplied by a scalar, the Result is proved on $\widehat{\mathcal{T}}(\mu, \nu)$. □

8.2 Compactness of tangent convex functions

Proof of Proposition 2.7 We first prove the result for $\theta = (\theta_n)_{n \geqslant 1} \subset \Theta$. Denote $\mathrm{conv}(\theta) := \{\theta' \in \Theta^{\mathbb{N}} : \theta'_n \in \mathrm{conv}(\theta_k, k \geqslant n), n \in \mathbb{N}\}$. Consider the minimization problem:

$$m \;:=\; \inf_{\theta' \in \mathrm{conv}(\theta)} \mu[G(\mathrm{dom}_X \underline{\theta'_\infty})], \quad (8.31)$$

where the measurability of $G(\mathrm{dom}_X\underline{\theta}'_\infty)$ follows from Lemma 3.1.
Step 1: We first prove the existence of a minimizer. Let $(\theta'^k)_{k\in\mathbb{N}} \in \mathrm{conv}(\theta)^{\mathbb{N}}$ be a minimizing sequence, and define the sequence $\widehat{\theta} \in \mathrm{conv}(\theta)$ by:

$$\widehat{\theta}_n := (1-2^{-n})^{-1}\sum_{k=1}^n 2^{-k}\theta'^k_n,\ n \geqslant 1.$$

Then, $\mathrm{dom}(\widehat{\underline{\theta}}_\infty) \subset \bigcap_{k\geqslant 1}\mathrm{dom}(\underline{\theta}'^k_\infty)$ by the non-negativity of θ', and we have the inclusion $\big\{\widehat{\theta}_n \underset{n\to\infty}{\longrightarrow} \infty\big\} \subset \big\{\theta'^k_n \underset{n\to\infty}{\longrightarrow} \infty \text{ for some } k \geqslant 1\big\}$. Consequently,

$$\mathrm{dom}_x\widehat{\underline{\theta}}_\infty \ \subset\ \mathrm{conv}\big(\textstyle\bigcap_{k\geqslant 1}\mathrm{dom}\underline{\theta}'^k_\infty(x,\cdot)\big) \ \subset\ \textstyle\bigcap_{k\geqslant 1}\mathrm{dom}_x\underline{\theta}'^k_\infty \quad \text{for all} \ \ x \in \mathbb{R}^d.$$

Since $(\theta'^k)_k$ is a minimizing sequence, and $\widehat{\theta} \in \mathrm{conv}(\theta)$, this implies that $\mu[G(\mathrm{dom}_X\widehat{\underline{\theta}}_\infty)] = m$.
Step 2: We next prove that we may find a sequence $(y_i)_{i\geqslant 1} \subset \mathbb{L}^0(\mathbb{R}^d,\mathbb{R}^d)$ such that

$$y_i(X) \in \mathrm{Aff}(\mathrm{dom}_X\widehat{\underline{\theta}}_\infty) \quad \text{and} \quad (y_i(X))_{i\geqslant 1} \text{ dense in } \mathrm{Aff}\,\mathrm{dom}_X\widehat{\underline{\theta}}_\infty,\ \ \mu-\text{a.s.} \tag{8.32}$$

Indeed, it follows from Lemmas 3.1, and 7.2 that the map $x \mapsto \mathrm{Aff}(\mathrm{dom}_x\widehat{\underline{\theta}}_\infty)$ is universally measurable, and therefore Borel-measurable up to a modification on a $\mu-$null set. Since its values are closed and nonempty, we deduce from the implication $(ii) \implies (ix)$ in Theorem 4.2 of the survey on measurable selection [21] the existence of a sequence $(y_i)_{i\geqslant 1}$ satisfying (8.32).
Step 3: Let $m(dx,dy) := \mu(dx)\otimes\sum_{i\geqslant 0}2^{-i}\delta_{\{y_i(x)\}}(dy)$. By the Komlòs lemma, we may find $\widetilde{\theta} \in \mathrm{conv}(\widehat{\theta})$ such that $\widetilde{\theta}_n \longrightarrow \widetilde{\theta}_\infty \in \mathbb{L}^0(\Omega)$, $m-$a.s. Clearly, $\mathrm{dom}_x\widetilde{\underline{\theta}}_\infty \subset \mathrm{dom}_x\widehat{\underline{\theta}}_\infty$, and therefore $\mu\big[G(\mathrm{dom}_X\widetilde{\underline{\theta}}_\infty)\big] \leqslant \mu\big[G(\mathrm{dom}_x\widehat{\underline{\theta}}_\infty)\big]$, for all $x \in \mathbb{R}^d$. This shows that

$$G(\mathrm{dom}_X\widetilde{\underline{\theta}}_\infty) \ = \ G(\mathrm{dom}_X\widehat{\underline{\theta}}_\infty),\quad \mu-\text{a.s.} \tag{8.33}$$

so that $\widetilde{\theta}$ is also a solution of the minimization problem (8.31). Moreover, it follows from (3.2) that

$$\mathrm{ri}\,\mathrm{dom}_X\widetilde{\underline{\theta}}_\infty = \mathrm{ri}\,\mathrm{dom}_X\widehat{\underline{\theta}}_\infty,\quad \text{and therefore} \quad \mathrm{Aff}\,\mathrm{dom}_X\widetilde{\underline{\theta}}_\infty = \mathrm{Aff}\,\mathrm{dom}_X\widehat{\underline{\theta}}_\infty,\quad \mu-a.s.$$

Step 4: Notice that the values taken by $\widetilde{\theta}_\infty$ are only fixed on an $m-$full measure set. By the convexity of elements of Θ in the $y-$variable, $\mathrm{dom}_X\widetilde{\theta}_n$ has a nonempty interior in $\mathrm{Aff}(\mathrm{dom}_X\widetilde{\underline{\theta}}_\infty)$. Then as $\mu-$a.s., $\widetilde{\theta}_n(X,\cdot)$ is convex, the following definition extends $\widetilde{\theta}_\infty$ to Ω:

$$\widetilde{\theta}_\infty(x,y) := \sup\big\{a\cdot y+b : (a,b) \in \mathbb{R}^d\times\mathbb{R},\ a\cdot y_n(x)+b \leqslant \widetilde{\theta}_\infty(x,y_n(x)) \text{ for all } n \geqslant 0\big\}.$$

This extension coincides with $\widetilde{\theta}_\infty$, in $(x,y_n(x))$ for $\mu-$a.e. $x \in \mathbb{R}^d$, and all $n \geqslant 1$ such that $y_n(x) \notin \partial\mathrm{dom}_X\widetilde{\theta}_k$ for some $k \geqslant 1$ such that $\mathrm{dom}_x\widetilde{\theta}_n$ has a nonempty interior in $\mathrm{Aff}(\mathrm{dom}_x\widetilde{\underline{\theta}}_\infty)$. As for k large enough, $\partial\mathrm{dom}_X\widetilde{\theta}_k$ is Lebesgue negligible in $\mathrm{Aff}(\mathrm{dom}_x\widetilde{\underline{\theta}}_\infty)$, the remaining $y_n(x)$ are still dense in $\mathrm{Aff}(\mathrm{dom}_x\widetilde{\underline{\theta}}_\infty)$. Then, for $\mu-$a.e. $x \in \mathbb{R}^d$, $\widetilde{\theta}_n(x,\cdot)$ converges to $\widetilde{\theta}_\infty(x,\cdot)$ on a dense subset of $\mathrm{Aff}(\mathrm{dom}_x\widetilde{\underline{\theta}}_\infty)$. We shall prove in Step 6 below that

$$\mathrm{dom}\,\widetilde{\theta}_\infty(X,\cdot) \text{ has nonempty interior in } \mathrm{Aff}(\mathrm{dom}_X\widetilde{\underline{\theta}}_\infty),\quad \mu-\text{a.s.} \tag{8.34}$$

Then, by Theorem 9.3, $\widetilde{\theta}_n(X,\cdot) \longrightarrow \widetilde{\theta}_\infty(X,\cdot)$ pointwise on $\mathrm{Aff}(\mathrm{dom}_X\underline{\widetilde{\theta}}_\infty)\backslash\partial\mathrm{dom}\widetilde{\theta}_\infty(X,\cdot)$, $\mu-$a.s. Since $\mathrm{dom}_X\theta_\infty = \mathrm{dom}_X\underline{\theta}_\infty$, and $\widetilde{\theta}$ converges to θ_∞ on $\mathrm{dom}_X\theta_\infty$, $\mu-$a.s., $\widetilde{\theta}$ converges to $\theta_\infty \in \Theta$, $\mu\otimes$pw.

Step 5: Finally for general $(\theta_n)_{n\geqslant 1} \subset \Theta_\mu$, we consider θ'_n, equal to θ_n, $\mu\otimes$pw, such that $\theta'_n \leqslant \theta_n$, for $n \geqslant 1$, from the definition of Θ_μ. Then we may find λ_n^k, coefficients such that $\widehat{\theta}'_n := \sum_{k\geqslant n}\lambda_n^k\theta'_k \in \mathrm{conv}(\theta')$ converges $\mu\otimes$pw to $\widehat{\theta}_\infty \in \Theta$. We denote $\widehat{\theta}_n := \sum_{k\geqslant n}\lambda_n^k\theta_k \in \mathrm{conv}(\theta)$, $\widehat{\theta}_n = \widehat{\theta}'_n$, $\mu\otimes$pw, and $\widehat{\theta}_n \geqslant \widehat{\theta}'_n$. By Proposition 2.6 (iii), $\widehat{\theta}$ converges to $\widehat{\theta}_\infty$, $\mu\otimes$pw. The Proposition is proved.

Step 6: In order to prove (8.34), suppose to the contrary that there is a set A such that $\mu[A] > 0$ and $\mathrm{dom}\widetilde{\theta}_\infty(x,\cdot)$ has an empty interior in $\mathrm{Aff}(\mathrm{dom}_x\underline{\widetilde{\theta}}_\infty)$ for all $x \in A$. Then, by the density of the sequence $(y_n(x))_{n\geqslant 1}$ stated in (8.32), we may find for all $x \in A$ an index $i(x) \geqslant 0$ such that

$$\widehat{y}(x) := y_{i(x)}(x) \in \mathrm{ri}\,\mathrm{dom}_x\underline{\widetilde{\theta}}_\infty, \quad \text{and} \quad \widetilde{\theta}_\infty(x,\widehat{y}(x)) = \infty. \tag{8.35}$$

Moreover, since $i(x)$ takes values in $\mathbb{N}$, we may reduce to the case where $i(x)$ is a constant integer, by possibly shrinking the set A, thus guaranteeing that $\widehat{y}$ is measurable. Define the measurable function on Ω:

$$\theta_n^0(x,y) := \mathrm{dist}(y, L_x^n) \quad \text{with} \quad L_x^n := \big\{y \in \mathbb{R}^d : \widetilde{\theta}_n(x,y) < \widetilde{\theta}_n(x,\widehat{y}(x))\big\}. \tag{8.36}$$

Since L_x^n is convex, and contains x for n sufficiently large by (8.35), we see that

$$\theta_n^0 \;\text{ is convex in } y \text{ and } \quad \theta_n^0(x,y) \leqslant |x-y|, \quad \text{for all } \;(x,y) \in \Omega. \tag{8.37}$$

In particular, this shows that $\theta_n^0 \in \Theta$. By Komlòs Lemma, we may find

$$\widehat{\theta}_n^0 := \textstyle\sum_{k\geqslant n}\lambda_k^n\theta_k^0 \in \mathrm{conv}(\theta^0) \quad \text{such that} \quad \widehat{\theta}_n^0 \longrightarrow \widehat{\theta}_\infty^0, \quad m-\text{a.s.}$$

for some non-negative coefficients $(\lambda_k^n, k \geqslant n)_{n\geqslant 1}$ with $\sum_{k\geqslant n}\lambda_k^n = 1$. By convenient extension of this limit, we may assume that $\widehat{\theta}_\infty^0 \in \Theta$. We claim that

$$\widehat{\theta}_\infty^0 > 0 \;\text{ on } \quad H_x := \{h(x)\cdot(y-\widehat{y}(x)) > 0\}, \quad \text{for some} \quad h(x) \in \mathbb{R}^d. \tag{8.38}$$

We defer the proof of this claim to Step 7 below and we continue in view of the required contradiction. By definition of θ_n^0 together with (8.37), we compute that

$$\begin{aligned}\theta_n^1(x,y) \;:=\; \sum_{k\geqslant n}\lambda_k^n\widetilde{\theta}_k(x,y) \;\geqslant\; \sum_{k\geqslant n}\lambda_k^n\widetilde{\theta}_k(x,\widehat{y}(x))\mathbf{1}_{\{\theta_n^0>0\}} \;&\geqslant\; \sum_{k\geqslant n}\lambda_k^n\widetilde{\theta}_k(x,\widehat{y}(x))\frac{\theta_k^0(x,y)}{|x-y|}\\ &\geqslant\; \frac{\widehat{\theta}_n^0(x,y)}{|x-y|}\inf_{k\geqslant n}\widetilde{\theta}_k(x,\widehat{y}(x)).\end{aligned}$$

By (8.35) and (8.38), this shows that the sequence $\theta^1 \in \mathrm{conv}(\theta)$ satisfies

$$\theta_n^1(x,\cdot) \longrightarrow \infty, \quad \text{on} \quad H_x, \quad \text{for all} \;\; x \in A.$$

We finally consider the sequence $\widetilde{\theta}^1 := \frac{1}{2}(\widetilde{\theta}+\theta^1) \in \mathrm{conv}(\theta)$. Clearly, $\mathrm{dom}\underline{\widetilde{\theta}}^1_\infty(X,\cdot) \subset \mathrm{dom}\underline{\widetilde{\theta}}_\infty(X,\cdot)$, and it follows from the last property of θ^1 that $\mathrm{dom}\underline{\widetilde{\theta}}^1_\infty(x,\cdot) \subset H_x^c \cap \mathrm{dom}\underline{\widetilde{\theta}}_\infty(x,\cdot)$ for all

$x \in A$. Notice that $\widehat{y}(x)$ lies on the boundary of the half space H_x and, by (8.35), $\widehat{y}(x) \in \text{ridom}_x \underline{\widetilde{\theta}}_\infty$. Then $G(\text{dom}_x \underline{\widetilde{\theta}}^1_\infty) < G(\text{dom}_x \underline{\widetilde{\theta}}_\infty)$ for all $x \in A$ and, since $\mu[A] > 0$, we deduce that $\mu\big[G(\text{dom}_X \underline{\widetilde{\theta}}^1_\infty)\big] < \mu\big[G(\text{dom}_X \underline{\widetilde{\theta}}_\infty)\big]$, contradicting the optimality of $\widetilde{\theta}$, by (8.33), for the minimization problem (8.31).

Step 7: It remains to justify (8.38). Since $\widetilde{\theta}_n(x,\cdot)$ is convex, it follows from the Hahn-Banach separation theorem that:

$$\widetilde{\theta}_n(x,\cdot) \geqslant \widetilde{\theta}_n(x,\widehat{y}(x)) \text{ on } H^n_x := \big\{y \in \mathbb{R}^d : h_n(x)\cdot(y-\widehat{y}(x)) > 0\big\}, \text{ for some } h^n(x) \in \mathbb{R}^d,$$

so that it follows from (8.36) that $L^n_x \subset (H^n_x)^c$, and

$$\theta^0_n(x,y) \geqslant \text{dist}\big(y,(H^n_x)^c\big) = \big[\big(y-\widehat{y}(x)\big)\cdot h^n(x)\big]^+.$$

Denote $g_x := g_{\text{dom}_x \underline{\widehat{\theta}}_\infty}$ the Gaussian kernel restricted to the affine span of $\text{dom}_x \underline{\widehat{\theta}}_\infty$, and $B_r(x_0)$ the corresponding ball with radius r, centered at some point x_0. By (8.35), we may find r^x so that $B^x_r := B_r(\widehat{y}(x)) \subset \text{ri}\,\text{dom}_x \underline{\widetilde{\theta}}_\infty$ for all $r \leqslant r^x$, and

$$\int_{B^x_r} \theta^0_n(x,y) g_x(y) dy \quad \geqslant \quad \int_{B^x_r} \big[\big(y-\widehat{y}(x)\big)\cdot h^n(x)\big]^+ g_x(y) dy \geqslant \min_{B^x_r} g_x \int_{B_r(0)} (y\cdot e_1)^+ dy =: b^r_x > 0,$$

where e_1 is an arbitrary unit vector of the affine span of $\text{dom}_x \underline{\widehat{\theta}}_\infty$. Then we have the inequality $\int_{B^x_r} \widehat{\theta}^0_n(x,y) g_x(y) dy \geqslant b^r_x$, and since $\widehat{\theta}^0_n$ has linear growth in y by (8.37), it follows from the dominated convergence theorem that $\int_{B^r_x} \widehat{\theta}^0_\infty(x,y) g(dy) \geqslant b^r_x > 0$, and therefore $\widehat{\theta}^0_\infty(x,y^r_x) > 0$ for some $y^r_x \in B^r_x$. From the arbitrariness of $r \in (0,r_x)$, We deduce (8.38) as a consequence of the convexity of $\widehat{\theta}^0(x,\cdot)$. □

Proof of Proposition 2.6 (iii) We need to prove the existence of some

$$\theta' \in \Theta \quad \text{such that} \quad \underline{\theta}_\infty = \theta', \quad \mu\otimes\text{pw}, \quad \text{and} \quad \underline{\theta}_\infty \geqslant \theta'. \tag{8.39}$$

For simplicity, we denote $\theta := \underline{\theta}_\infty$. Let

$$F^1 := \text{cl conv dom}\theta(X,\cdot), \quad F^k := \text{cl conv}\big(\text{dom}\theta(X,\cdot) \cap \text{Aff rf}_X F^{k-1}\big), \; k \geqslant 2,$$
$$\text{and} \quad F := \cup_{n\geqslant 1}(F^n \backslash \text{cl rf}_X F^n) \cup \text{cl dom}_X \theta.$$

Fix some sequence $\varepsilon_n \searrow 0$, and denote $\theta_* := \liminf_{n\to\infty} \theta\big(X, \varepsilon_n X + (1-\varepsilon_n)Y\big)$, and

$$\theta' := \big[\infty \mathbf{1}_{Y\notin F(X)} + \mathbf{1}_{Y \in \text{cl dom}_X \theta}\theta_*\big]\mathbf{1}_{X\notin N_\mu},$$

where $N_\mu \in \mathcal{N}_\mu$ is chosen such that $\mathbf{1}_{Y\in F^k(X)}\mathbf{1}_{X\notin N_\mu}$ are Borel measurable for all k from Lemma 6.1, and $\theta(x,\cdot)$ (resp. $\theta_n(x,\cdot)$) is convex finite on $\text{dom}_x\theta$ (resp. $\text{dom}_x\theta_n$), for $x \notin N_\mu$. Consequently, θ' is measurable. In the following steps, we verify that θ' satisfies (8.39).

Step 1: We prove that $\theta' \in \Theta$. Indeed, $\theta' \in \mathbb{L}^0_+(\Omega)$, and $\theta'(X,X) = 0$. Now we prove that $\theta'(x,\cdot)$ is convex for all $x \in \mathbb{R}^d$. For $x \in N_\mu$, $\theta'(x,\cdot) = 0$. For $x \notin N_\mu$, $\theta(x,\cdot)$ is convex finite on $\text{dom}_x\theta$, then by the fact that $\text{dom}_x\theta$ is a convex relatively open set containing x, it follows from Lemma 8.1 that $\theta_*(x,\cdot) = \lim_{n\to\infty}\theta\big(x, \varepsilon_n x + (1-\varepsilon_n)\cdot\big)$ is the lower semi-continuous envelop of $\theta(x,\cdot)$ on $\mathrm{cl\,dom}_x\theta$. We now prove the convexity of $\theta'(x,\cdot)$ on all $\mathbb{R}^d$. We denote $\widehat{F}(x) := F(x)\backslash \mathrm{cl\,dom}_x\theta$ so that $\mathbb{R}^d = F(x)^c \cup \widehat{F}(x) \cup \mathrm{cl\,dom}_x\theta$. Now, let $y_1, y_2 \in \mathbb{R}^d$, and $\lambda \in (0,1)$. If $y_1 \in F(x)^c$, the convexity inequality is verified as $\theta'(x,y_1) = \infty$. Moreover, $\theta'(x,\cdot)$ is constant on $\widehat{F}(x)$, and convex on $\mathrm{cl\,dom}_x\theta$. We shall prove in Steps 4 and 5 below that

$$F(x) \text{ is convex, and } \mathrm{rf}_x F(x) = \mathrm{dom}_x\theta. \tag{8.40}$$

In view of Proposition 2.1 (ii), this implies that the sets $\widehat{F}(x)$ and $\mathrm{cl\,dom}_x\theta$ are convex. Then we only need to consider the case when $y_1 \in \widehat{F}(x)$, and $y_2 \in \mathrm{cl\,dom}_x\theta$. By Proposition 2.1 (ii) again, we have $[y_1, y_2) \subset \widehat{F}(x)$, and therefore $\lambda y_1 + (1-\lambda)y_2 \in \widehat{F}(x)$, and $\theta'(x, \lambda y_1 + (1-\lambda)y_2) = 0$, which guarantees the convexity inequality.

<u>*Step 2:*</u> We next prove that $\theta = \theta'$, $\mu\otimes$pw. By the second claim in (8.40), it follows that $\theta_*(X,\cdot)$ is convex finite on $\mathrm{dom}_X\theta$, $\mu-$a.s. Then as a consequence of Proposition 2.4 (ii), we have $\mathrm{dom}_X\theta' = \mathrm{dom}_X(\infty\mathbf{1}_{Y\notin F(X)}) \cap \mathrm{dom}_X(\theta_*\mathbf{1}_{Y\in \mathrm{cl\,dom}_X\theta})$, $\mu-$a.s. The first term in this intersection is $\mathrm{rf}_X F(X) = \mathrm{dom}_X\theta$. The second contains $\mathrm{dom}_X\theta$, as it is the dom_X of a function which is finite on $\mathrm{dom}_X\theta$, which is convex relatively open, containing X. Finally, we proved that $\mathrm{dom}_X\theta = \mathrm{dom}_X\theta'$, $\mu-$a.s. Then $\theta'(X,\cdot)$ is equal to $\theta_*(X,\cdot)$ on $\mathrm{dom}_X\theta$, and therefore, equal to $\theta(X,\cdot)$, $\mu-$a.s. We proved that $\theta = \theta'$, $\mu\otimes$pw.

<u>*Step 3:*</u> We finally prove that $\theta' \leqslant \theta$ pointwise. We shalll prove in Step 6 below that

$$\mathrm{dom}\theta(X,\cdot) \quad \subset \quad F. \tag{8.41}$$

Then, $\infty\mathbf{1}_{Y\notin F(X)}\mathbf{1}_{X\notin N_\mu} \leqslant \theta$, and it remains to prove that

$$\theta(x,y) \geqslant \theta_*(x,y) \quad \text{for all} \quad y \in \mathrm{cl\,dom}_x\theta, \quad x \notin N_\mu.$$

To see this, let $x \notin N_\mu$. By definition of N_μ, $\theta_n(x,\cdot) \longrightarrow \theta(x,\cdot)$ on $\mathrm{dom}_x\theta$. Notice that $\theta(x,\cdot)$ is convex on $\mathrm{dom}_x\theta$, and therefore as a consequence of Lemma 8.1,

$$\theta_*(x,y) = \lim_{\epsilon\searrow 0}\theta\big(x, \epsilon x + (1-\epsilon)y\big), \quad \text{for all} \quad y \in \mathrm{cl\,dom}_x\theta.$$

Then $y^\epsilon := (1-\epsilon)y + \epsilon x \in \mathrm{dom}_x\theta_n$, for $\varepsilon \in (0,1]$, and n sufficiently large by (i) of this Proposition, and therefore $(1-\epsilon)\theta_n(x,y) - \theta_n(x,y_\epsilon) \geqslant (1-\epsilon)\theta'_n(x,y) - \theta'_n(x,y_\epsilon) \geqslant 0$, for $\theta'_n \in \Theta$ such that $\theta'_n = \theta_n$, $\mu\otimes$pw, and $\theta_n \geqslant \theta'_n$. Taking the lim inf as $n\to\infty$, we get $(1-\epsilon)\theta(x,y) - \theta(x,y_\epsilon) \geqslant 0$, and finally $\theta(x,y) \geqslant \lim_{\epsilon\searrow 0}\theta\big(x, \epsilon x + (1-\epsilon)y\big) = \theta'(x,y)$, by sending $\epsilon \searrow 0$.

<u>*Step 4:*</u> (First claim in (8.40)) Let $x_0 \in \mathbb{R}^d$, let us prove that $F(x_0)$ is convex. Indeed, let $x, y \in F(x_0)$, and $0 < \lambda < 1$. Since $\mathrm{cl\,dom}_x\theta$ is convex, and $F^n(x_0)\backslash \mathrm{cl\,rf}_X F^n(x_0)$ is convex by Proposition 2.1 (ii), we only examine the following non-obvious cases:

- Suppose $x \in F^n(x_0)\backslash \mathrm{cl\,rf}_{x_0}F^n(x_0)$, and $y \in F^p(x_0)\backslash \mathrm{cl\,rf}_{x_0}F^p(x_0)$, with $n < p$. Then as $F^p(x_0)\backslash \mathrm{cl\,rf}_{x_0}F^p(x_0) \subset \mathrm{cl\,rf}_{x_0}F^n(x_0)$, and $\lambda x + (1-\lambda)y \in F^n(x_0)\backslash \mathrm{cl\,rf}_{x_0}F^n(x_0)$ by Proposition 2.1 (ii).
- Suppose $x \in F^n(x_0)\backslash \mathrm{cl\,rf}_{x_0}F^n(x_0)$, and $y \in \mathrm{cl\,dom}_{x_0}\theta$, then as $\mathrm{cl\,dom}_{x_0}\theta \subset \mathrm{cl\,rf}_{x_0}F^n(x_0)$, this case is handled similar to previous case.

Step 5: (Second claim in (8.40)). We have $\mathrm{dom}_X\theta \subset F(X)$, and therefore $\mathrm{dom}_X\theta \subset \mathrm{rf}_X F(X)$. Now we prove by induction on $k \geqslant 1$ that $\mathrm{rf}_X F(X) \subset \cup_{n\geqslant k}(F^n\backslash \mathrm{cl}\,\mathrm{rf}_X F^n) \cup \mathrm{cl}\,\mathrm{dom}_X\theta$. The inclusion is trivially true for $k = 1$. Let $k \geqslant 1$, we suppose that the inclusions holds for k, hence $\mathrm{rf}_X F(X) \subset \cup_{n\geqslant k}(F^n\backslash \mathrm{cl}\,\mathrm{rf}_X F^n) \cup \mathrm{cl}\,\mathrm{dom}_X\theta$. As $\cup_{n\geqslant k}(F^n\backslash \mathrm{cl}\,\mathrm{rf}_X F^n) \cup \mathrm{cl}\,\mathrm{dom}_X\theta \subset F^k$. Applying rf_X gives

$$\begin{aligned}
\mathrm{rf}_X F(X) &\subset \mathrm{rf}_X\Big[\cup_{n\geqslant k}(F^n\backslash \mathrm{cl}\,\mathrm{rf}_X F^n) \cup \mathrm{cl}\,\mathrm{dom}_X\theta\Big]\\
&= \mathrm{rf}_X\Big[F^k \cap \cup_{n\geqslant k}(F^n\backslash \mathrm{cl}\,\mathrm{rf}_X F^n) \cup \mathrm{cl}\,\mathrm{dom}_X\theta\Big]\\
&= \mathrm{rf}_X F^k \cap \mathrm{rf}_X\Big[\cup_{n\geqslant k}(F^n\backslash \mathrm{cl}\,\mathrm{rf}_X F^n) \cup \mathrm{cl}\,\mathrm{dom}_X\theta\Big]\\
&\subset \mathrm{cl}\,\mathrm{rf}_X F^k \cap \cup_{n\geqslant k}(F^n\backslash \mathrm{cl}\,\mathrm{rf}_X F^n) \cup \mathrm{cl}\,\mathrm{dom}_X\theta\\
&\subset \cup_{n\geqslant k+1}(F^n\backslash \mathrm{cl}\,\mathrm{rf}_X F^n) \cup \mathrm{cl}\,\mathrm{dom}_X\theta.
\end{aligned}$$

Then the result is proved for all k. In particular we apply it for $k = d + 1$. Recall from the proof of Lemma 3.1 that for $n \geqslant d + 1$, F^n is stationary at the value $\mathrm{cl}\,\mathrm{dom}_X\theta$. Then $\cup_{n\geqslant d+1}(F^n\backslash \mathrm{cl}\,\mathrm{rf}_X F^n) = \emptyset$, and $\mathrm{rf}_X F(X) \subset \mathrm{rf}_X \mathrm{cl}\,\mathrm{dom}_X\theta = \mathrm{dom}_X\theta$. The result is proved.
Step 6: We finally prove (8.41). Indeed, $\mathrm{dom}\theta(X,\cdot) \subset F^1$ by definition. Then

$$\begin{aligned}
\mathrm{dom}\theta(X,\cdot) &\subset F^1\backslash \mathrm{Aff}F^1 \cup \big(\cup_{2\leqslant k\leqslant d+1}(\mathrm{dom}\theta(X,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_X F^{k-1})\backslash \mathrm{Aff}F^k\big) \cup F^{d+1}\\
&\subset F^1\backslash \mathrm{cl}\,F^1 \cup \big(\cup_{k\geqslant 2}\ \mathrm{cl}\ \mathrm{conv}(\mathrm{dom}\theta(X,\cdot) \cap \mathrm{Aff}\,\mathrm{rf}_X F^{k-1})\backslash \mathrm{cl}\,F^k\big) \cup\ \mathrm{cl}\,\mathrm{dom}_X\theta\\
&= \cup_{k\geqslant 1}F^k\backslash \mathrm{cl}\,F^k \cup \mathrm{cl}\,\mathrm{dom}_X\theta = F.
\end{aligned}$$

□

9 Some convex analysis results

We start by proving the required properties of the notion of relative face.

Proof of Proposition 2.1 (i) is an easy exercise. As for (ii), we first prove that $\mathrm{rf}_a A = \mathrm{ri}A \neq \emptyset$ iff $a \in \mathrm{ri}A$. We suppose that $\mathrm{rf}_a A = \mathrm{ri}A \neq \emptyset$. The non-emptiness implies by (i) that $a \in A$, and therefore $a \in \mathrm{rf}_a A = \mathrm{ri}A$. Now we suppose that $a \in \mathrm{ri}A$. Then for $x \in \mathrm{ri}A$, $\big[x, a - \epsilon(x - a)\big] \subset \mathrm{ri}A \subset A$, for some $\epsilon > 0$, and therefore $\mathrm{ri}A \subset \mathrm{rf}_a A$. On the other hand, by (9.43), $\mathrm{ri}A = \{x \in \mathbb{R}^d : x \in (x', x_0],\ \text{for some } x_0 \in \mathrm{ri}A,\ \text{and } x' \in A\}$. Taking $x_0 := a \in \mathrm{ri}A$, we have the remaining inclusion $\mathrm{rf}_a A \subset \mathrm{ri}A$.

We now prove that $\mathrm{rf}_a A$ is convex. We consider $x, y \in \mathrm{rf}_a A$ and $\lambda \in [0, 1]$. We consider an $\epsilon > 0$ such that $\big(a - \epsilon(x - a), x + \epsilon(x - a)\big) \subset A$ and $\big(a - \epsilon(y - a), y + \epsilon(y - a)\big) \subset A$. Then if we write $z = \lambda x + (1 - \lambda)y$, $\big(a - \epsilon(z - a), z + \epsilon(z - a)\big) \subset A$ by convexity of A, because $a, x, y \in A$.

In order to prove that $\mathrm{rf}_a A$ is relatively open, we consider $x, y \in \mathrm{rf}_a A$, and we verify that $\big(x - \epsilon(y - x), y + \epsilon(y - x)\big) \subset \mathrm{rf}_a A$ for some $\epsilon > 0$. Consider the two alternatives:
Case 1: If a, x, y are on a line. If $a = x = y$, then the required result is obvious. Otherwise,

$$\big(a - \epsilon(x - a), x + \epsilon(x - a)\big) \cup \big(a - \epsilon(y - a), y + \epsilon(y - a)\big) \subset \mathrm{rf}_a A$$

This union is open in the line and x and y are interior to it. We can find $\epsilon' > 0$ such that $\big(x - \epsilon'(y-x), y + \epsilon'(y-x)\big) \subset \mathrm{rf}_a A$.

<u>*Case 2:*</u> If a, x, y are not on a line. Let $\epsilon > 0$ be such that $\big(a - 2\epsilon(x-a), x + 2\epsilon(x-a)\big) \subset A$ and $\big(a - 2\epsilon(y-a), y + 2\epsilon(y-a)\big) \subset A$. Then $x + \epsilon(x-a) \in \mathrm{rf}_a A$ and $a - \epsilon(y-a) \in \mathrm{rf}_a A$. Then, if we take $\lambda := \frac{\epsilon}{1+2\epsilon}$,

$$\lambda(a - \epsilon(y-a)) + (1-\lambda)(x + \epsilon(x-a)) = (1-\lambda)(1+\epsilon)x - \lambda\epsilon y = x + \lambda\epsilon(y-x)$$

Then $x + \lambda\epsilon(x-y) \in \mathrm{rf}_a A$ and symmetrically, $y + \lambda\epsilon(y-x) \in \mathrm{rf}_a A$ by convexity of $\mathrm{rf}_a A$. And still by convexity, $\big(x - \epsilon'(y-x), y + \epsilon'(y-x)\big) \subset \mathrm{rf}_a A$ for $\epsilon' := \frac{\epsilon^2}{1+2\epsilon} > 0$.

(v) We will prove these two results by a recurrence on the dimension of the space d. First if $d = 0$ the results are trivial. Now we suppose that the result is proved for any $d' < d$, let us prove it for dimension d.

<u>*Case 1:*</u> $a \in \mathrm{ri}A$. This case is trivial as $\mathrm{rf}_a A = \mathrm{ri}A$ and $A \subset \mathrm{cl}\,\mathrm{ri}A = \mathrm{cl}\,\mathrm{rf}_a A$ because of the convexity of A. Finally $A \backslash \mathrm{cl}\,\mathrm{rf}_a A = \varnothing$ which makes it trivial.

<u>*Case 2:*</u> $a \notin \mathrm{ri}A$. Then $a \in \partial A$ and there exists a hyperplan support H to A in a because of the convexity of A. We will write the equation of the corresponding half-space containing A, $E : c \cdot x \leqslant b$ with $c \in \mathbb{R}^d$ and $b \in \mathbb{R}$. As $x \in \mathrm{rf}_a A$ implies that $[a - \epsilon(x-a), x + \epsilon(x-a)] \subset A$ for some $\epsilon > 0$, we have $(a - \epsilon(x-a)) \cdot c \leqslant b$ and $(x + \epsilon(x-a)) \cdot c \leqslant b$. These equations are equivalent using that $a \in H$ and thus $a \cdot c = b$ to $-\epsilon(x-a) \cdot c \leqslant 0$ and $(1+\epsilon)(x-a) \cdot c \leqslant 0$. We finally have $(x-a) \cdot c = 0$ and $x \in H$. We proved that $\mathrm{rf}_a A \subset H$.

Now using (iii) together with the fact that $\mathrm{rf}_a A \subset H$ and $a \in H$ affine, we have

$$\mathrm{rf}_a(A \cap H) = \mathrm{rf}_a A \cap \mathrm{rf}_a H = \mathrm{rf}_a A \cap H = \mathrm{rf}_a A.$$

Then we can now have the recurrence hypothesis on $A \cap H$ because $\dim H = d - 1$ and $A \cap H \subset H$ is convex. Then we have $A \cap H \backslash \mathrm{cl}\,\mathrm{rf}_a A$ which is convex and if $x_0 \in A \cap H \backslash \mathrm{cl}\,\mathrm{rf}_a(A \cap H)$, $y_0 \in A \cap H$ and if $\lambda \in (0,1]$ then $\lambda x_0 + (1-\lambda)y_0 \in A \backslash \mathrm{cl}\,\mathrm{rf}_a(A \cap H)$.

First $A \backslash \mathrm{cl}\,\mathrm{rf}_a A = (A \backslash H) \cup (A \cap H \backslash \mathrm{cl}\,\mathrm{rf}_a A)$, let us show that this set is convex. The two sets in the union are convex ($A \backslash H = A \cap (E \backslash H)$), so we need to show that a non trivial convex combination of elements coming from both sets is still in the union. We consider $x \in A \backslash H$, $y \in A \cap H \backslash \mathrm{cl}\,\mathrm{rf}_a A$ and $\lambda > 0$, let us show that $z := \lambda x + (1-\lambda)y \in (A \backslash H) \cup (A \cap H \backslash \mathrm{cl}\,\mathrm{rf}_a A)$. As $x, y \in A$ ($\mathrm{cl}\,\mathrm{rf}_a A \subset A$ because A is closed), $z \in A$ by convexity of A. We now prove $z \notin H$,

$$z \cdot c = \lambda x \cdot c + (1-\lambda) y \cdot c = \lambda x \cdot c + (1-\lambda)b < \lambda b + (1-\lambda)b = b.$$

Then z is in the strict half space: $z \in E \backslash H$. Finally $z \in A \backslash H$ and $A \backslash \mathrm{cl}\,\mathrm{rf}_a A$ is convex.

Let us now prove the second part: we consider $x_0 \in A \backslash \mathrm{cl}\,\mathrm{rf}_a A$, $y_0 \in \mathrm{cl}\,\mathrm{rf}_a A$ and $\lambda \in (0,1]$ and write $z_0 := \lambda x_0 + (1-\lambda)y_0$.

<u>*Case 2.1:*</u> $x_0, y_0 \in H$. We apply the induction hypothesis.

<u>*Case 2.2:*</u> $x_0, y_0 \in A \backslash H$. Impossible because $\mathrm{rf}_a A \subset H$ and $\mathrm{cl}\,\mathrm{rf}_a A \subset \mathrm{cl}\,H = H$. $y_0 \in H$.

<u>*Case 2.3:*</u> $x_0 \in A \backslash H$ and $y_0 \in H$. Then by the same computation than the proof of convexity,

$$z_0 \in A \backslash H \subset A \backslash \mathrm{cl}\,\mathrm{rf}_a A.$$

(vi) We first prove the equivalence, we suppose that $a \in \mathrm{ri}A$. As by the convexity of A, $\mathrm{ri}A = \mathrm{ri}\,\mathrm{cl}\,A$, $\mathrm{rf}_a\mathrm{cl}\,A = \mathrm{ri}\,\mathrm{cl}\,A$, and therefore $\mathrm{cl}\,\mathrm{rf}_a\mathrm{cl}\,A = \mathrm{cl}\,A$. Finally, taking the dimension, we have $\dim(\mathrm{cl}\,\mathrm{rf}_a\mathrm{cl}\,A) = \dim(A)$. In this case we proved as well that $\mathrm{cl}\,\mathrm{rf}_a\mathrm{cl}\,A = \mathrm{cl}\,\mathrm{ri}\,\mathrm{cl}\,A = \mathrm{cl}\,A = \mathrm{cl}\,\mathrm{rf}_a A$, the last equality coming from the fact that $\mathrm{ri}A = \mathrm{rf}_a A$ as $a \in \mathrm{ri}A$.

Now we suppose that $a \notin \mathrm{ri}A$. Then $a \in \partial\mathrm{cl}\,A$, and $\mathrm{rf}_a\mathrm{cl}\,A \subset \partial\mathrm{cl}\,A$. Taking the dimension (in the local sense this time), and by the fact that $\dim\partial\mathrm{cl}\,A = \dim\partial A < \dim A$, we have $\dim(\mathrm{cl}\,\mathrm{rf}_a\mathrm{cl}\,A) < \dim(A)$ (as $\mathrm{cl}\,\mathrm{rf}_a\mathrm{cl}\,A$ is convex, the two notions of dimension coincide). □

We next report a result on the union of intersecting relative interiors of convex subsets which was used in the proof of Proposition 4.1. We shall use the following characterization of the relative interior of a convex subset K of $\mathbb{R}^d$:

$$\begin{aligned} \mathrm{ri}K &= \big\{x \in \mathbb{R}^d : x - \epsilon(x' - x) \in K \text{ for some } \epsilon > 0, \text{ for all } x' \in K\big\} && (9.42)\\ &= \big\{x \in \mathbb{R}^d : x \in (x', x_0], \text{ for some } x_0 \in \mathrm{ri}K, \text{ and } x' \in K\big\}. && (9.43) \end{aligned}$$

Lemma 9.1. *Let $K_1, K_2 \subset \mathbb{R}^d$ be convex with $\mathrm{ri}K_1 \cap \mathrm{ri}K_2 \neq \emptyset$. Then $\mathrm{conv}(\mathrm{ri}K_1 \cup \mathrm{ri}K_2) = \mathrm{ri}\,\mathrm{conv}(K_1 \cup K_2)$.*

Proof. We fix $y \in \mathrm{ri}K_1 \cap \mathrm{ri}K_2$.

Let $x \in \mathrm{conv}(\mathrm{ri}K_1 \cup \mathrm{ri}K_2)$, we may write $x = \lambda x_1 + (1-\lambda)x_2$, with $x_1 \in \mathrm{ri}K_1$, $x_2 \in \mathrm{ri}K_2$, and $0 \leqslant \lambda \leqslant 1$. If λ is 0 or 1, we will suppose then that $0 < \lambda < 1$. Then for $x' \in \mathrm{conv}(K_1 \cup K_2)$, we may write $x' = \lambda' x_1' + (1-\lambda')x_2'$, with $x_1' \in K_1$, $x_2' \in K_2$, and $0 \leqslant \lambda' \leqslant 1$. We will use y as a center as it is in both the sets. For all the variables, we add a bar on it when we subtract y, for example $\bar{x} := x - y$. The geometric problem is the same when translated with y,

$$\bar{x} - \epsilon(\bar{x}' - \bar{x}) \;=\; \lambda\left(\bar{x}_1 - \epsilon\left(\frac{\lambda'}{\lambda}\bar{x}_1' - \bar{x}_1\right)\right) + (1-\lambda)\left(\bar{x}_2 - \epsilon\left(\frac{1-\lambda'}{1-\lambda}\bar{x}_2' - \bar{x}_2\right)\right). \quad (9.44)$$

However, as $\bar{x}_1$ and $\bar{x}_1'$ are in $K_1 - y$, as it is a convex and 0 is an interior point, $\epsilon(\frac{\lambda'}{\lambda}\bar{x}_1' - \bar{x}_1) \in K_1 - y$ for ϵ small enough. Then as $\bar{x}_1$ is interior to $K_1 - y$ as well, $\bar{x}_1 - \epsilon(\frac{\lambda'}{\lambda}\bar{x}_1' - \bar{x}_1) \in K_1 - y$ as well. By the same reasoning, $\bar{x}_2 - \epsilon(\frac{1-\lambda'}{1-\lambda}\bar{x}_2' - \bar{x}_2) \in K_2 - y$. Finally, by (9.44), for ϵ small enough, $x - \epsilon(x' - x) \in \mathrm{conv}(K_1 \cup K_2)$. By (9.42), $x \in \mathrm{ri}\,\mathrm{conv}(K_1 \cup K_2)$.

Now let $x \in \mathrm{ri}\,\mathrm{conv}(K_1 \cup K_2)$. We use again y as an origin with the notation $\bar{x} := x - y$. As $\bar{x}$ is interior, we may find $\epsilon > 0$ such that $(1+\epsilon)\bar{x} \in \mathrm{conv}(K_1 \cup K_2)$. We may write $(1+\epsilon)\bar{x} = \lambda\bar{x}_1 + (1-\lambda)\bar{x}_2$, with $\bar{x}_1 \in K_1 - y$, $\bar{x}_2 \in K_2 - y$, and $0 \leqslant \lambda \leqslant 1$. Then $\bar{x} = \lambda\frac{1}{1+\epsilon}\bar{x}_1 + (1-\lambda)\frac{1}{1+\epsilon}\bar{x}_2$. By (9.43), $\frac{1}{1+\epsilon}\bar{x}_1 \in \mathrm{ri}K_1$, and $\frac{1}{1+\epsilon}\bar{x}_2 \in \mathrm{ri}K_2$. $\bar{x} \in \mathrm{conv}\big(\mathrm{ri}(K_1 - y) \cup \mathrm{ri}(K_2 - y)\big)$, and therefore $x \in \mathrm{conv}(\mathrm{ri}K_1 \cup \mathrm{ri}K_2)$. □

Now we use the measurable selection theory to establish the non-emptiness of ∂f.

Lemma 9.2. *Let $f \in \mathfrak{C}$, $\partial f \neq \emptyset$.*

Proof. By the fact that f is continuous, we may write $\partial f(x) = \cap_{n\geqslant 1}F_n(x)$ for all $x \in \mathbb{R}^d$, with $F_n(x) := \{p \in \mathbb{R}^d : f(y_n) - f(x) \geqslant p \cdot (y_n - x)\}$ where $(y_n)_{n\geqslant 1} \subset \mathbb{R}^d$ is some fixed dense sequence. All F_n are measurable by the continuity of $(x,p) \longmapsto f(y_n) - f(x) - p \cdot (y_n - x)$ together with Theorem 6.4 in [12]. Therefore the mapping $x \longmapsto \partial f(x)$ is measurable by Lemma 7.1. Moreover, it is well known properties of the subgradient of finite convex functions that this mapping is closed nonempty-valued. Then the result holds by Theorem 4.1 in [21]. □

We conclude this section with the following result which has been used in our proof of Proposition 2.7. We believe that this is a standard convex analysis result, but we could not find precise references. For this reason, we report the proof for completeness.

Theorem 9.3. *Let $f_n, f : \mathbb{R}^d \to \bar{\mathbb{R}}$ be convex functions with* $\operatorname{int}\operatorname{dom} f \neq \emptyset$*. Then $f_n \longrightarrow f$ pointwise on $\mathbb{R}^d \backslash \partial \operatorname{dom} f$ if and only if $f_n \longrightarrow f$ pointwise on some dense subset $A \subset \mathbb{R}^d \backslash \partial \operatorname{dom} f$.*

Proof. We prove the non-trivial implication "if". We first prove the convergence on $\operatorname{int}\operatorname{dom} f$. f_n converges to f on a dense set. The reasoning will consist in proving that the f_n are Lipschitz, it will give a uniform convergence and then a pointwise convergence. First we consider $K \subset \operatorname{int}\operatorname{dom} f$ compact convex with nonempty interior. We can find $N \in \mathbb{N}$ and $x_1, ...x_N \in A \cap (\operatorname{int}\operatorname{dom} f \backslash K)$ such that $K \subset \operatorname{int}\operatorname{conv}(x_1, ..., x_N)$. We use the pointwise convergence on A to get that for n large enough, $f_n(x) \leqslant M$ for $x \in \operatorname{conv}(x_1, ..., x_N)$, $M > 0$ (take $M = \max_{1 \leqslant k \leqslant N} f(x_k) + 1$). Then we will prove that f_n is bounded from below on K. We consider $a \in A \cap K$ and $\delta_0 := \sup_{x \in K} |x - a|$. For n large enough, $f_n(a) \geqslant m$ for any $a \in A_0$ (take for example $m = f(a) - 1$). We write $\delta_1 := \min_{(x,y) \in K \times \partial \operatorname{conv}(x_1,...,x_N)} |x - y|$. Finally we write $\delta_2 := \sup_{x,y \in \operatorname{conv}(x_1,...,x_N)} |x - y|$. Now, for $x \in K$, we consider the half line $[x, a)$, it will cut $\partial\operatorname{conv}(x_1, ..., x_N)$ in one only point $y \in \partial\operatorname{conv}(x_1, ..., x_N)$. Then $a \in [x, y]$, and therefore $a = \frac{|a-y|}{|x-y|} x + \frac{|a-x|}{|x-y|} y$. By the convex inequality, $f_n(a) \leqslant \frac{|a-y|}{|x-y|} f_n(x) + \frac{|a-x|}{|x-y|} f_n(y)$. Then $f_n(x) \geqslant -\frac{|a-x|}{|a-y|} M + \frac{|x-y|}{|a-y|} m \geqslant -\frac{\delta_0}{\delta_1} M + \frac{\delta_2}{\delta_1} m$. Finally, if we write $m_0 := -\frac{\delta_0}{\delta_1} M + \frac{\delta_2}{\delta_1} m$,

$$M \geqslant f_n \geqslant m_0, \quad \text{on } K.$$

This will prove that f_n is $\frac{M-m0}{\delta_1}$-Lipschitz. We consider $x \in K$ and a unit direction $u \in \mathcal{S}^{d-1}$ and $f'_n \in \partial f_n(x)$. For a unique $\lambda > 0$, $y := x + \lambda u \in \partial\operatorname{conv}(x_1, ..., x_N)$. As u is a unit vector, $\lambda = |y - x| \geqslant \delta_1$. By the convex inequality, $f_n(y) \geqslant f_n(x) + f'_n(x) \cdot (y - x)$. Then $M - m_0 \geqslant \delta_0 |f'_n \cdot u|$ and finally $|f'_n \cdot u| \leqslant \frac{M-m0}{\delta_1}$ as this bound does not depend on u, $|f'_n| \leqslant \frac{M-m0}{\delta_1}$ for any such subgradient. For n big enough the f_n are uniformly Lipschitz on K, and so in f. The convergence is uniform on K, it is then pointwise on K. As this is true for any such K, the convergence is pointwise on $\operatorname{int}\operatorname{dom} f$.
Now let us consider $x \in (\operatorname{cl}\operatorname{dom} f)^c$. The set $\operatorname{conv}(x, \operatorname{int}\operatorname{dom} f) \backslash \operatorname{dom} f$ has a nonempty interior because $\operatorname{dist}(x, \operatorname{dom} f) > 0$ and $\operatorname{int}\operatorname{dom} f \neq \emptyset$. As A is dense, we can consider $a \in A \cap \operatorname{conv}(x, \operatorname{int}\operatorname{dom} f) \backslash \operatorname{dom} f$. By definition of $\operatorname{conv}(x, \operatorname{int}\operatorname{dom} f)$, we can find $y \in \operatorname{int}\operatorname{dom} f$ such that $a = \lambda y + (1 - \lambda) x$. $\lambda > 0$ because $a \notin \operatorname{dom} f$. If $\lambda = 1$, $f_n(x) = f_n(a) \underset{n \to \infty}{\longrightarrow} \infty$. Otherwise, by the convexity inequality, $f_n(a) \leqslant \lambda f_n(y) + (1 - \lambda) f_n(x)$. Then, as $f_n(a) \underset{n \to \infty}{\longrightarrow} \infty$, and $f_n(y) \underset{n \to \infty}{\longrightarrow} f(y) < \infty$, we have $f_n(x) \underset{n \to \infty}{\longrightarrow} \infty$. □

References

[1] Gerald Beer. A polish topology for the closed subsets of a polish space. *Proceedings of the American Mathematical Society*, 113(4):1123–1133, 1991.

[2] Mathias Beiglböck, Pierre Henry-Labordère, and Friedrich Penkner. Model-independent bounds for option prices: a mass transport approach. *Finance and Stochastics*, 17(3):477–501, 2013.

[3] Mathias Beiglböck and Nicolas Juillet. On a problem of optimal transport under marginal martingale constraints. *The Annals of Probability*, 44(1):42–106, 2016.

[4] Mathias Beiglböck, Marcel Nutz, and Nizar Touzi. Complete duality for martingale optimal transport on the line. *arXiv preprint arXiv:1507.00671*, 2015.

[5] Dimitri P Bertsekas and Steven E Shreve. *Stochastic optimal control: The discrete time case*, volume 23. Academic Press New York, 1978.

[6] Alexander MG Cox and Jan Obłój. Robust pricing and hedging of double no-touch options. *Finance and Stochastics*, 15(3):573–605, 2011.

[7] Hadrien De March. Quasi-sure duality for multi-dimensional martingale optimal transport (work in progress). 2017.

[8] Ibrahim Ekren and H Mete Soner. Constrained optimal transport. *arXiv preprint arXiv:1610.02940*, 2016.

[9] Alfred Galichon, Pierre Henry-Labordere, and Nizar Touzi. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *The Annals of Applied Probability*, 24(1):312–336, 2014.

[10] Nassif Ghoussoub, Young-Heon Kim, and Tongseok Lim. Structure of optimal martingale transport plans in general dimensions. *arXiv preprint arXiv:1508.01806*, 2015.

[11] Christian Hess. *Contribution à l'étude de la mesurabilité, de la loi de probabilité et de la convergence des multifonctions*. PhD thesis, 1986.

[12] Charles J Himmelberg. Measurable relations. *Fundamenta mathematicae*, 87(1):53–72, 1975.

[13] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex analysis and minimization algorithms I: Fundamentals*, volume 305. Springer science & business media, 2013.

[14] David Hobson. The skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton Lectures on Mathematical Finance 2010*, pages 267–318. Springer, 2011.

[15] David Hobson and Martin Klimmek. Robust price bounds for the forward starting straddle. *Finance and Stochastics*, 19(1):189–214, 2015.

[16] David Hobson and Anthony Neuberger. Robust bounds for forward start options. *Mathematical Finance*, 22(1):31–56, 2012.

[17] Hans G Kellerer. Duality theorems for marginal problems. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 67(4):399–432, 1984.

[18] Jan Obłój and Pietro Siorpaes. Structure of martingale transports in finite dimensions. 2017.

[19] Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, 2015.

[20] Volker Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, pages 423–439, 1965.

[21] Daniel H Wagner. Survey of measurable selection theorems. *SIAM Journal on Control and Optimization*, 15(5):859–903, 1977.

[22] Danila A Zaev. On the monge–kantorovich problem with additional linear constraints. *Mathematical Notes*, 98(5-6):725–741, 2015.