Pareto efficiency for the concave order and multivariate comonotonicity

G. Carlier *, R.-A. Dana †, A. Galichon ‡

December 2, 2009

Abstract

In this paper, we focus on efficient risk-sharing rules for the concave dominance order. For a univariate risk, it follows from a *comonotone dominance principle*, due to Landsberger and Meilijson [25], that efficiency is characterized by a comonotonicity condition. The goal of this paper is to generalize the comonotone dominance principle as well as the equivalence between efficiency and comonotonicity to the multi-dimensional case. The multivariate setting is more involved (in particular because there is no immediate extension of the notion of comonotonicity) and we address it using techniques from convex duality and optimal transportation.

Keywords: concave order, stochastic dominance, comonotonicity, efficiency, multivariate risk-sharing.

1 Introduction

In this paper, we study *Pareto efficient allocations* of risky consumptions of *several goods* in a contingent exchange economy. We shall consider a framework where goods are imperfect substitutes and agents have incomplete preferences associated with the *concave order*.

There is a distinguished tradition in modelling preferences by concave dominance. Introduced in economics by Rothschild and Stiglitz [32], the

*CEREMADE, UMR CNRS 7534, Université Paris IX Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, FRANCE carlier@ceremade.dauphine.fr

†CEREMADE, UMR CNRS 7534, Université Paris IX Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, FRANCE dana@ceremade.dauphine.fr

‡Département d'Economie, Ecole Polytechnique, alfred.galichon@polytechnique.edu

concave order has then been used in a wide variety of economic contexts. To give a few references, let us mention efficiency pricing (Peleg and Yaari [30], Chew and Zilcha [10]), measurement of inequality (Atkinson [3]), finance (Dybvig [15], Jouini and Kallal [23]).

With respect to most of the aforementioned literature, the novelty of this paper is to deal with the multivariate case, i.e. the case where consumption is denominated in several units which are imperfectly substitutable. These units can be e.g. consumption and labor; or future consumptions at various maturities; or currency units with limited exchangeability. The hypothesis of the lack of, or of the imperfect substituability between various consumption units arises in many different fields of the economic and financial literatures. This amounts to the fact that one can no longer model consumption as a *random variable*, but as a *random vector*.

The aim of the paper is to find testable implications of efficiency (for the concave order) on observable data (for instance, insurance contracts) and a tractable parametrization of efficient allocations. In the case of univariate risk, this tractable characterization exists: the *comonotonicity* property. Indeed, since the early work of Borch [5], Arrow [1], [2] and Wilson [36], it is well-known that efficient allocations of risk between expected utility maximizers fulfill *the mutuality principle* or equivalently are comonotone. It may easily be proven that these allocations are efficient for the concave order. It is also well-known that if agents have preferences compatible with the concave dominance (we shall refer to them as being *risk averse*), then efficient allocations must be comonotone, otherwise there would be mutually profitable transfers among agents (see le Roy and Werner [26]).

An important step in the theory of efficient risk-sharing was made by Landsberger and Meilijson [25] who proved (for two agents and a discrete setting) that any allocation of a given aggregate risk is dominated in the sense of concave dominance by a comonotone allocation. Moreover, this dominance can be made strict if the initial allocation is not itself comonotone. This result that we shall refer to as the *comonotone dominance principle* has been extended to the continuous case by limiting arguments (see [13] and [28]). It implies the comonotonicity of efficient allocations for the concave order. The equivalence between comonotonicity and efficiency was only proved recently by Dana [12] for the discrete case and by Dana and Meilijson [13] for the continuous case. Therefore comonotonicity fully characterizes efficiency and it is a testable and tractable property. Townsend [35] proposed to test whether the mutuality principle holds in three poor villages in southern India and did find that individual consumptions comove positively with village average consumption. From a theoretical point of view, comonotone allocations form a tractable class which is convex and almost compact, in a sense to be made precise later on. Existence results may then be obtained for many risk-sharing problems by restricting attention to comonotone allocations (see for instance [24] in the framework of risk measures, or [8], [9] for classes of law invariant and concave utilities). Furthermore, even though the comonotone dominance principle has been essentially used in the case of particular classes of utility functions, it is a very general principle that may be useful for incomplete preferences, which are compatible with the concave dominance.

In this paper, we will first revisit in detail the comonotone dominance result in the univariate case. In particular, we will give a direct proof under the assumption that the underlying probability space is non-atomic. Our proof does not rely on the discrete case and a limiting argument, but instead uses the theory of monotone rearrangements (see [7] for other applications). We will then prove that any efficient allocation for the concave order is a solution of a risk-sharing problem between expected utility maximizers and that efficient allocations coincide with comonotone allocations. Let us mention that a totally different proof of Landsberger and Meilisjon's result, based on a certain variational problem, will follow from our analysis of the multivariate case.

The remainder and central part of the paper will be devoted to the extension of the comonotone dominance result and its application to the characterization of efficient allocations to the multivariate setting. While the case of a univariate risk has been very much investigated, it is far less so for the multivariate case. This is mainly due to the fact that in the multivariate risk framework, it was not obvious until recently what the appropriate notion of comonotonicity should be. Ekeland, Galichon and Henry in [18] have introduced a notion of multivariate comonotonicity to characterize comonotonic multivariate risk measures, which they call μ-comonotonicity. We shall say that an allocation $(X_1, ..., X_p)$ (with each X_i being multivariate) is comonotone if there is a random vector Z and convex and differentiable maps φ_i such that $X_i = \nabla\varphi_i(Z)$ (in [18], μ is then the distribution of Z). While this is indeed an extension of the univariate definiton, this is by no means the only possible one. In a recent paper [31], Puccetti and Scarsini review various possible other multivariate extensions of the notion of comonotonicity and emphasize the fact that *naive extensions do not enjoy some of the main properties of the univariate concept*. In fact, it turns out that, as we show in this paper, the notion of comonotonicity which is related to efficient risk allocations is (up to some regularity subtleties), the one of [18]. We shall extend the comonotone dominance principle to the multivariate case and apply it to characterize Pareto efficiency. The statements will be however more complicated than in the univariate case and will involve taking weak closures at some points and also to introduce strict convexity in a quantified way. This stems from the fact that multivariate comonotone allocations of a given risk do not form neither a convex nor a compact (up to constants) set contrary to the univariate case (counterexamples will be given). While the results of [18] are strongly related to maximal correlation functionals and to the quadratic optimal transportation problem (and in particular Brenier's seminal paper [6]) the present approach will rely on a slightly different optimization problem that has some familiarities with the multi-marginals optimal transport problem of Gangbo and Święch [20].

The paper is organized as follows. Section 2 gives some definitions and tools from rearrangement theory. Section 3 deals with the univariate case, revisiting the dominance result of [25] with a new proof and applications to Pareto risk-sharing. Our notion of multivariate comonotoncity is introduced in section 4, an analogue of the comonotone dominance principle is stated and efficient sharing-rules are then characterized. The proofs of the multivariate results are given in section 5, they are based on convex duality for an optimization problem on measures that is in the spirit of optimal transportation theory (although slightly different). Finally, section 6 concludes the paper.

2 Preliminaries

2.1 Definitions and notations

Given as primitive is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For every (univariate or multivariate) random vector X on such space, the law of X is denoted $\mathcal{L}(X)$. Given X and Y two random vectors, we will say that X and Y are equivalent in distribution, which we denote $X \sim Y$, if $\mathcal{L}(X) = \mathcal{L}(Y)$.

Definition 2.1. *Let X and Y be bounded random vectors with values in $\mathbb{R}^d$, then X dominates Y (for the concave order), which we shall denote $X \succcurlyeq Y$, if and only if $\mathbb{E}(\varphi(X)) \leq \mathbb{E}(\varphi(Y))$ for every convex function $\varphi : \mathbb{R}^d \to \mathbb{R}$. If, in addition, $\mathbb{E}(\varphi(X)) < \mathbb{E}(\varphi(Y))$ for every strictly convex function φ, then X is said to dominate Y strictly.*

The concave order is usually defined with concave utilities rather than with convex loss functions. Clearly the definition above coincides with the standard one. We use convex functions to be consistent with the standard conventions from *convex* analysis, a field which we shall extensively use in the multivariate case (Legendre transforms, infimal convolutions, convex duality). Note that if $X \succcurlyeq Y$, then $\mathbb{E}(X) = \mathbb{E}(Y)$ so that comparing risks for $\succcurlyeq$ only makes sense for random vectors with the same mean. In words, $X \succcurlyeq Y$ means that $\mathbb{E}(X) = \mathbb{E}(Y)$ and X is riskier (or more spread) than Y. In this paper, we will focus on the concave order rather than on the second order stochastic dominance which is widely used in economics. For the sake of completeness, we recall that given two real-valued bounded random vectors X and Y, X is said to dominate Y for second-order stochastic dominance (notation $X \succeq_2 Y$) whenever $\mathbb{E}(u(X)) \geq \mathbb{E}(u(Y))$ for every concave and *nondecreasing* function $u : \mathbb{R}^d \to \mathbb{R}$. It is easy to see then that $X \succcurlyeq Y$ if and only if $X \succeq_2 Y$ and $\mathbb{E}(X) = \mathbb{E}(Y)$. We refer to Rothschild and Stiglitz [32] and Föllmer and Schied [19] for various characterizations of second-order stochastic dominance in the univariate case and to Müller and Stoyan [29] for the multivariate case.

In the univariate case, we recall that comonotonicity is defined by

Definition 2.2. *Let X_1 and X_2 be two real-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, then the pair (X_1, X_2) is comonotone if*

$$(X_1(\omega') - X_1(\omega))(X_2(\omega') - X_2(\omega)) \geq 0 \text{ for } \mathbb{P} \otimes \mathbb{P}\text{-a.e. } (\omega, \omega') \in \Omega^2.$$

An $\mathbb{R}^p$-valued random vector $(X_1, ..., X_p)$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be comonotone if (X_i, X_j) is comonotone for every $(i, j) \in \{1, ..., p\}^2$.

It is well-known that comonotonicity of $(X_1, ..., X_p)$ is equivalent to the fact that each X_i can be written as a nondecreasing function of the sum $\sum_i X_i$ (see for instance Denneberg [14]). The extension of this notion to the multivariate case (i.e when each X_i is $\mathbb{R}^d$-valued) is not immediate and will be addressed in section 4.

Given $X \in L^\infty(\Omega, \mathbb{R}^d)$ a random vector of aggregate risk of dimension $d \geq 1$, we denote by $\mathcal{A}(X)$ the corresponding set of admissible allocations or risk-sharing of X among p agents:

$$\mathcal{A}(X) := \{\mathbf{Y} = (Y_1, ..., Y_p) \in L^\infty(\Omega, \mathbb{R}^d) : \sum_{i=1}^{p} Y_i = X\}.$$

For simplicity we have not written explicitly the dependence of $\mathcal{A}(X)$ on the number p of agents.

For $d = 1$, we denote by $\text{com}(X)$ the set of comonotone allocations in $\mathcal{A}(X)$. Therefore $(X_1, ..., X_p) \in \text{com}(X)$ if and only if there are nondecreasing functions f_i summing to the identity such that $X_i = f_i(X)$. Note also that the functions f_i's are then all 1-Lipschitz and then allocations in $\text{com}(X)$ are 1-Lipschitz functions of X.

Definition 2.3. *Let* $\mathbf{X} = (X_1, ..., X_p)$ *and* $\mathbf{Y} := (Y_1, ..., Y_p)$ *be in* $\mathcal{A}(X)$*, then* $\mathbf{X}$ *is said to dominate* $\mathbf{Y}$ *if* $X_i \succcurlyeq Y_i$ *for every* $i \in \{1, ..., p\}$*. If, in addition there is an* $i \in \{1, ..., p\}$ *such that* X_i *strictly dominates* Y_i *then* $\mathbf{X}$ *is said to strictly dominate* $\mathbf{Y}$*. An allocation* $\mathbf{X} \in \mathcal{A}(X)$ *is called Pareto-efficient (for the concave order) if there is no allocation in* $\mathcal{A}(X)$ *that strictly dominates* $\mathbf{X}$*.*

Remark 2.4. Dominance of allocations can also be defined as follows. Let $\mathbf{X} = (X_1, ..., X_p)$ and $\mathbf{Y} := (Y_1, ..., Y_p)$ be in $\mathcal{A}(X)$, then $\mathbf{X}$ dominates $\mathbf{Y}$ if and only if

$$\mathbb{E}(\sum_i \varphi_i(X_i)) \leq \mathbb{E}(\sum_i \varphi_i(Y_i))$$

for every collection of convex functions $\varphi_i : \mathbb{R}^d \to \mathbb{R}$. Moreover $\mathbf{X}$ strictly dominates $\mathbf{Y}$ if and only if the previous inequality is strict for every collection of strictly convex functions $\varphi_i : \mathbb{R}^d \to \mathbb{R}$.

Remark 2.5. Note that (for $d = 1$) the concave order coincides with second order stochastic dominance on $\mathcal{A}(X)$. Indeed if $(X_1, ..., X_p)$ and $(Y_1, ..., Y_p)$ belong to $\mathcal{A}(X)$ and if $X_i \succeq_2 Y_i$ for every i, then for all i, $\mathbb{E}(X_i) \geq \mathbb{E}(Y_i)$. Since $\sum_i \mathbb{E}(X_i) = \sum_i \mathbb{E}(Y_i) = \mathbb{E}(X)$, we obtain that $\mathbb{E}(X_i) = \mathbb{E}(Y_i)$ for all i and, as recalled above, the two dominances coincide on random variable with same expectations.

2.2 Rearrangement inequalities and comonotonicity in the univariate case

A fundamental tool for the univariate analysis is a supermodular version of Hardy-Littlewood's inequality which we now restate. We thus need to recall the concepts of nondecreasing rearrangement of $f : [0, 1] \to \mathbb{R}$ with respect to the Lebesgue measure and that of a submodular function.

Two Borel functions on $[0, 1]$, f and g, are equimeasurable with respect to the Lebesgue measure denoted λ, if, for any uniformly distributed (on $[0, 1]$) random variable U, $f(U)$ and $g(U)$ have same distribution. Given f an integrable function on $[0, 1]$, there exists a unique right-continuous nondecreasing function denoted $\widetilde{f}$ which is equimeasurable to f, $\widetilde{f}$ called the *nondecreasing rearrangement* of f (with respect to the Lebesgue measure).

A function $L : \mathbb{R}^2 \to \mathbb{R}$ is submodular if for all $(x_1, y_1, x_2, y_2) \in \mathbb{R}^4$ such that $x_2 \geq x_1$ and $y_2 \geq y_1$:

$$L(x_2, y_2) + L(x_1, y_1) \leq L(x_1, y_2) + L(x_2, y_1). \tag{2.1}$$

It is strictly submodular if for all $(x_1, y_1, x_2, y_2) \in \mathbb{R}^4$ such that $x_2 > x_1$ and $y_2 > y_1$:

$$L(x_2, y_2) + L(x_1, y_1) < L(x_1, y_2) + L(x_2, y_1). \tag{2.2}$$

A function $L \in C^2$ is submodular if and only if $\partial^2_{xy} L(x, y) \le 0$ for all $(x, y) \in \mathbb{R}^2$. If $\partial^2_{xy} L(x, y) < 0$ for all $(x, y) \in \mathbb{R}^2$, then L is strictly submodular.

Important examples of submodular functions are as follows. If $\varphi : \mathbb{R} \to \mathbb{R}$ is convex (strictly convex) and $L(x, y) = \varphi(x - y)$, then L is submodular (strictly submodular). Similarly $(x, y) \mapsto u(x + y)$ is submodular for any concave u.

The submodular version of Hardy-Littlewood's inequality then reads as:

Lemma 2.6. *Let f and g be in $L^\infty([0, 1], \mathcal{B}, \lambda)$ and $\widetilde{f}$, $\widetilde{g}$ be their nondecreasing rearrangements and L be submodular. We then have, for any random variable U uniformly distributed on $[0, 1]$*

$$\mathbb{E}(L(\widetilde{f}(U), \widetilde{g}(U))) \le \mathbb{E}(L(f(U), g(U))).$$

Moreover if L is continuous and strictly submodular, then the inequality is strict unless f and g are comonotone, that is fulfill:

$$(f(t) - f(t'))(g(t) - g(t')) \ge 0 \ \lambda \otimes \lambda\text{-a.e. } .$$

Let us give simple applications of Lemma 2.6 that will be very useful for the construction of comonotone allocations dominating a given allocation.

Lemma 2.7. *Let f be in $L^\infty([0, 1], \mathcal{B}, \lambda)$ and $\widetilde{f}$ be its nondecreasing rearrangement. Then, for any uniformly distributed random variable U and any increasing and bounded function g on $[0, 1]$, one has*

1. *$g(U) - \widetilde{f}(U) \succcurlyeq g(U) - f(U)$, with strict dominance if f is not nondecreasing,*
2. *$\|g(U) - \widetilde{f}(U)\|_{L^p} \le \|g(U) - f(U)\|_{L^p}$ for any $p \in [1, \infty]$.*
3. *If $0 \le f \le \mathrm{id}$, then $0 \le \widetilde{f} \le \mathrm{id}$.*

Proof. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be convex (strictly) and $L(x, y) = \varphi(x - y)$, then L is submodular (strictly). From Lemma 2.6, we have $\mathbb{E}(\varphi(g(U) - \widetilde{f}(U))) \le \mathbb{E}(\varphi(g(U) - f(U)))$ with a strict inequality whenever φ is strictly convex and f is not nondecreasing, proving the first assertion. To prove the second assertion, we take $\varphi(x) = |x|^p$ for any $p \in [1, \infty[$, the case $p = \infty$ is obtained by passing to the limit. To prove the last statement, we first remark that if $f \geq 0$ then $\widetilde{f} \geq 0$ since it is equimeasurable to f. We then define the submodular function $(x, y) \mapsto (x - y)_+$. If $f \leq \text{id}$, it follows from lemma 2.6 that

$$0 = \mathbb{E}((f(U) - U)_+) \geq \mathbb{E}((\widetilde{f}(U) - U)_+)$$

so that $\widetilde{f} \leq \text{id}$. □

2.3 Characterization of comonotonicity by maximal correlation

We now provide another characterization of comonotonicity based on the notion of maximal correlation. From now on, we assume that the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic which means that there is no $A \in \mathcal{F}$ such that for every $B \in \mathcal{F}$ if $\mathbb{P}(B) < \mathbb{P}(A)$ then $\mathbb{P}(B) = 0$. It is well-known that $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic if and only if a random variable $U \sim \mathcal{U}([0, 1])$ (that is U is uniformly distributed on $[0, 1]$) can be constructed on $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and define for every $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ (both Z and X being univariate here) the maximal correlation functional:

$$\varrho_Z(X) := \sup_{\tilde{X} \sim X} \mathbb{E}(Z\tilde{X}) = \sup_{\tilde{Z} \sim Z} \mathbb{E}(\tilde{Z}X) = \sup_{\tilde{Z} \sim Z,\, \tilde{X} \sim X} \mathbb{E}(\tilde{Z}\tilde{X}). \tag{2.3}$$

The functional ϱ_Z has extensively been discussed in economics and in finance, therefore we only recall a few useful facts. Let F_X^{-1} be the quantile function of X, that is the pseudo-inverse of distribution function F_X,

$$F_X^{-1}(u) := \inf \{y : F_X(y) > u\}.$$

From Hardy-Littlewood's inequalities, we have

$$\varrho_Z(X) = \int_0^1 F_X^{-1}(t) F_Z^{-1}(t) dt$$

and the supremum is achieved by any pair $(\tilde{Z}, \tilde{X})$ of comonotone random variables $(F_Z^{-1}(U), F_X^{-1}(U))$ for U uniformly distributed. By symmetry, one can either fix Z or fix X. Fixing for instance Z, the supremum is achieved by $F_X^{-1}(U)$ where $U \sim \mathcal{U}([0, 1])$ and satisfies $Z = F_Z^{-1}(U)$. When Z is non-atomic, there exists a unique $U = F_Z(Z)$ such that $Z = F_Z^{-1}(U)$ and the supremum is uniquely attained by the non-decreasing function of Z, $F_X^{-1} \circ F_Z(Z)$:

$$\varrho_Z(X) = \mathbb{E}(Z F_X^{-1} \circ F_Z(Z)) \tag{2.4}$$

Clearly ϱ_μ is a subadditive functional and we have

$$\varrho_Z\left(\sum_i X_i\right) \leq \sum_i \varrho_Z(X_i). \tag{2.5}$$

Proposition 2.8. *Let $(X_1, ..., X_p)$ be in $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. The following assertions are equivalent:*

1. *$(X_1, ..., X_p)$ are comonotone,*
2. *for any $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ nonatomic,*
$$\varrho_Z\left(\sum_i X_i\right) = \sum_i \varrho_Z\left(X_i\right), \tag{2.6}$$
3. *For some $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ nonatomic, (2.6) holds true.*

Proof. For the sake of simplicity, we restrict ourselves to $p = 2$ and set $(X_1, X_2) = (X, Y)$.

1 implies 2 for any Z since $F^{-1}_{X+Y} = F^{-1}_X + F^{-1}_Y$ for X and Y comonotone. To show that 3 implies 1, assume that for some Z non atomic, we have (2.6) or equivalently from (2.5) that

$$\varrho_Z(X + Y) \geq \varrho_Z(X) + \varrho_Z(Y)$$

Let Z_{X+Y} (resp Z_X and Z_Y) be distributed as Z and solve $\sup_{\tilde{Z}\sim Z} E(\tilde{Z}X)$ (resp $\varrho_Z(X)$ and $\varrho_Z(Y)$). We then have:

$$\mathbb{E}(Z_{X+Y}(X + Y)) \geq \mathbb{E}(Z_X X) + \mathbb{E}(Z_Y Y)$$

As $\mathbb{E}(Z_{X+Y}X) \leq \mathbb{E}(Z_X X)$ and $\mathbb{E}(Z_{X+Y}Y) \leq \mathbb{E}(Z_Y Y)$, we obtain that $E(Z_{X+Y}X) = E(Z_X X) = \varrho_Z(X)$ and $\mathbb{E}(Z_{X+Y}Y) = \mathbb{E}(Z_Y Y) = \varrho_Z(Y)$, hence from (2.4), $X = F^{-1}_X \circ F_{Z_{X+Y}}(Z_{X+Y}))$ and $Y = F^{-1}_Y \circ F_{Z_{X+Y}}(Z_{X+Y}))$ proving their comonotonicity. □

Proposition 2.8 was the starting point of Ekeland, Galichon and Henry [18], for providing a multivariate generalization of the concept of comonotonicity. In the sequel we shall further discuss this multivariate extension and compare it with the one we propose in the present paper.

3 The univariate case

3.1 An extension of Landsberger and Meilijson's dominance result

A landmark result, originally due to Landsberger and Meilijson [25] states that any allocation is dominated by a comonotone one. The original proof was carried in the discrete case for two agents and the results were extended to the general case by approximation. We give a different proof based on rearrangements, which is of interest per se since it does not require approximation arguments and slightly improves on the original statement by proving strict dominance of non-comonotone allocations. Like in Landsberger's and Meilijson's work, our argument is constructive in the case of two agents – but the two constructions are different. Contrary to Landsberger and Meilijson, we need however to assume, as before that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic.

Theorem 3.1. *Let X be a bounded real-valued random variable on the non-atomic probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathbf{X} = (X_1, ..., X_p) \in \mathcal{A}(X)$ be an allocation. There exists a comonotone allocation in $\mathcal{A}(X)$ that dominates $\mathbf{X}$. Moreover, if $\mathbf{X}$ is not comonotone, then there exists such an allocation that strictly dominates $\mathbf{X}$.*

Proof. Let us start with the case $p = 2$ and let $(X_1, X_2) \in \mathcal{A}(X)$. It follows from Ryff's polar factorization theorem (see [34]) that X can be written as $X = F(U)$ with U uniformly distributed and $F := F_X^{-1}$ is nondecreasing. By Jensen's conditional inequality $\mathbb{E}(X_i|X)$ dominates X_i and thus we may assume that the X_i's are functions of X hence of U:

$$X_1 = f_0(U),\ X_2 = g_0(U),\ f_0(x) + g_0(x) = F(x),\ \forall x \in [0, 1]$$

for Borel and bounded functions f_0 and g_0. Let us then define

$$X_1^1 = f_1(U),\ X_2^1 = g_1(U) := F(U) - f_1(U),\ \text{with } f_1 := \widetilde{f_0}.$$

By construction, $X_1^1 \sim X_1$ and it follows from lemma 2.7 that X_2^1 dominates X_1^1. Let us also remark that if (X_1, X_2) is not comonotone, then either f_0 or g_0 is not nondecreasing. Let us assume without loss of generality that f_0 is not nondecreasing. It thus follows again from lemma 2.7 that X_2^1 strictly dominates X_2. We then define a sequence (X_1^k, X_2^k) by taking alternated rearrangements as follows:

$$(X_1^k, X_2^k) = (f_k(U), g_k(U))$$

with for every $k \in \mathbb{N}$:

$$f_{2k+1} = \widetilde{f_{2k}},\ g_{2k+1} = F - f_{2k},\ g_{2k+2} = \widetilde{g_{2k+1}},\ f_{2k+2} = F - g_{2k+2}.$$

By construction, the sequence (X_1^k, X_2^k) belongs to $\mathcal{A}(X)$ and is monotone for the concave order. Moreover, the sequences f_k and g_k are bounded in L^∞ (use lemma 2.6 again). It thus follows from Helly's theorem that the sequences of nondecreasing functions (f_{2k+1}) and (g_{2k}) admit (pointwise and in L^p) converging subsequences. Moreover since $(f_k(U))$ is monotone for the concave order if f and f' are two cluster points of (f_{2k+1}), then $f(U)$ and $f'(U)$ have same law hence $f = f'$ since both are nondecreasing. This proves that the whole sequence (f_{2k+1}) converges to some nondecreasing f and similarly, the whole sequence (g_{2k}) converges to some nondecreasing g. Since $(f(U), F(U) - f(U))$ and $(F(U) - g(U), g(U))$ are limit points of the sequence $(X_1^k, X_2^k))$ that is monone for the concave order then one has $f(U) \sim F(U) - g(U)$ and $g(U) \sim F(U) - f(U)$ and then

$$f = \widetilde{F - g},\ g = \widetilde{F - f}.$$

Now, if we had $F - g \neq f$ then by lemma 2.7, $F - f$ would strictly dominate g which is absurd. We then have $f = F - g$ and the whole sequence (X_1^k, X_2^k) therefore converges to the comonotone allocation $(f(U), g(U))$ that dominates (X_1, X_2). Moreover, this dominance is strict if (X_1, X_2) is not itself comonotone since in this case (up to switching the role of X_1 and X_2) we have seen that (X_1^1, X_2^1) already strictly dominates (X_1, X_2).

Let us now treat the case $p = 3$, the case $p \geq 4$ generalizes straightforwardly by induction. Let $(X_1, X_2, X_3) \in \mathcal{A}(X)$, it follows from the previous step that there are F_1 and F_2 in $\mathcal{A}(X)$ with F_i being a nondecreasing functions of X and such that F_1 dominates $X_1 + X_2$ and F_2 dominates X_3. Since F_1 dominates $X_1 + X_2$, there is a bistochastic linear operator T (see [11]) such that $F_1 = T(X_1 + X_2) = T(X_1) + T(X_2)$. Let us then define $Y_1 = T(X_1)$ and $Y_2 = T(X_2)$, we have $Y_1 + Y_2 = F_1$ and Y_i dominates X_i $i = 1,\ 2$. It follows from the previous step that there are Z_1 and Z_2 summing to F_1, comonotone (hence nondecreasing in F_1 hence in X) such that Z_i dominates Y_i for $i = 1,\ 2$. Set then $Z_3 := F_2$, we then have (Z_1, Z_2, Z_3) is comonotone, belongs to $\mathcal{A}(X)$ and dominates (X_1, X_2, X_3).

Let us finally prove that dominance can be made strict if the initial allocation is not comonotone. Let $\mathbf{X} = (X_1, ..., X_p) \in \mathcal{A}(X)$ and let us assume that $\mathbf{X}$ is not comonotone. It follows from the previous steps that there is a $\mathbf{Y} = (Y_1, ..., Y_p) \in \mathcal{A}(X)$ that is comonotone and dominates $\mathbf{X}$. Since $\mathbf{X}$ is not comonotone there is an i for which $X_i \neq Y_i$, and then the allocation

$(\mathbf{X}+\mathbf{Y})/2$ strictly dominates $\mathbf{X}$. There is finally a comonotone allocation $\mathbf{Z} \in \mathcal{A}(X)$ that dominates $(\mathbf{X}+\mathbf{Y})/2$ and thus strictly dominates $\mathbf{X}$.

□

Remark 3.2. Theorem 3.1 may be also applied if aggregate risk is non negative and allocations are restricted to be non negative. Indeed, from lemma 2.7, if $X_1 \geq 0$ and $X_2 \geq 0$, then for each k, $X_1^k \geq 0$, $X_2^k \geq 0$ and non negativity holds true for their pointwise limit. As far as the second step of the proof of theorem 3.1 is concerned, it is enough to remark that, if $Z \succcurlyeq X$ and $X \geq 0$, then $\mathbb{E}(Z_-) \leq \mathbb{E}(X_-) = 0$ so that $Z \geq 0$.

Since the first step of the proof of theorem 3.1 consists in constructing in an algorithmic way an $\succcurlyeq$-monotone sequence of allocations that converge to a dominating one, let us now illustrate this construction on an example.

Example 3.3 (The normal case). Assume that $X \sim N(0, \sigma_X^2)$ and let $\mathbf{X} = (X_1, ..., X_p) \in \mathcal{A}(X)$ be normal (we only consider normal allocations of X in this example). Let us first remark that $\mathbf{X}$ is $\succcurlyeq$-dominated by $\mathbf{Y} = (Y_1, ..., Y_p)$ where $Y_i = \mathbb{E}(X_i \mid X)$ (strictly if some Y_i is not a function of X). As $\mathbb{E}(X_i \mid X) = \alpha_i X + \beta_i$ for all i, we may restrict attention to affine functions of X, $\alpha_i X + \beta_i$ which sum to X.

Let us then consider an allocation of the form $(\alpha X, (1-\alpha)X)$ (which is comonotone if and only if $\alpha \in [0,1]$). If this pair is not comonotone, without loss of generality, we may assume that $\alpha > 1$. Let us apply the first part of the algorithm, using the fact that for any $\gamma \in \mathbb{R}$, $\mathcal{L}(\gamma X) = \mathcal{L}(-\gamma X)$. $X_1^1 = \alpha X$, $X_1^2 = (1-\alpha)X$, $X_2^2 = (\alpha - 1)X$, $X_2^1 = (2-\alpha)X$, $X_3^1 = (\alpha - 2)X$, $X_3^2 = (2-\alpha)X$, $X_{2n}^1 = (2n - \alpha)X$, $X_{2n}^2 = (\alpha - (2n-1))$, $X_{2n+1}^1 = (\alpha - 2n)X$, $X_{2n+1}^2 = (2n+1-\alpha)X$. If $2n \leq \alpha \leq 2n+1$, the algorithm terminates after $2n$ iterations, the pair $X_{2n+1}^1 = (\alpha - 2n)X$, $X_{2n+1}^2 = (2n+1-\alpha)X$ is comonotone.

3.2 Application to efficiency for the concave order

Theorem 3.4. *Let X be a bounded real-valued random variable on the non-atomic probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathbf{X} = (X_1, ..., X_p) \in \mathcal{A}(X)$, then the following statements are equivalent:*

1. $\mathbf{X}$ *is efficient,*
2. $\mathbf{X} \in \mathrm{com}(X)$,

3. *there exist continuous and strictly convex functions* $(\psi_1, ..., \psi_p)$ *such that* $\mathbf{X}$ *solves*

$$\inf\{\sum_i \mathbb{E}(\psi_i(Y_i)) \ : \ \sum_i Y_i = X\},$$

4. *for every* $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ *nonatomic one has*

$$\varrho_Z\left(\sum_i X_i\right) = \sum_i \varrho_Z\left(X_i\right),$$

where ϱ_Z *is the maximal correlation functional defined by (2.3).*

Proof. 1 implies 2 : the fact that efficient allocations of X are comonotone directly follows from theorem 3.4. 2 implies 3: if $\mathbf{X} = (X_1, ..., X_p) \in \text{com}(X)$, let us write $X_i = f_i(X)$ for some nondecreasing and 1-Lipschitz functions f_i: $[m, M] \to \mathbb{R}$ (with $M := \text{Esssup}X$, $m := \text{Essinf}X$) summing to the identity map. Extending the f_i's by $f_i(x) = f_i(M) + (x - M)/p$ for $x \geq M$ and $f_i(x) = f_i(m) + (x - m)/p$ for $x \leq m$ we get 1-Lipschitz nondecreasing functions summing to the identity everywhere. Now let $\varphi(x) := \int_0^x f_i(s)ds$ for every x, the functions φ_i are by construction, convex and $C^{1,1}$ and have quadratic growth at ∞. The convex conjugates $\psi_i := \varphi_i^*$ are strictly convex and continuous functions and by construction one has for every i, $X \in \partial\psi_i(X_i)$ a.s., which implies that $(X_1, ..., X_p)$ minimizes $\mathbb{E}(\sum_i \psi_i(Y_i))$ subject to $\sum_i Y_i = X$ which proves 3. 3 implies 1 since the functions ψ_i's are strictly convex, if $(X_1, ..., X_p)$ satisfies 3 then it is an efficient allocation of X. Finally, the equivalence between 2 and 4 follows from proposition 2.8. □

As an immediate consequence, we have the following properties of efficient allocations:

Corollary 3.5. *Under the same assumptions as above, the set of efficient allocations of* X *is convex and compact in* L^∞ *up to zero-sum translations (which means that it can be written as* $\{(\lambda_1, ..., \lambda_p) \ : \ \sum_{i=1}^p \lambda_i = 0\} + A_0$ *with* A_0 *compact in* L^∞*). In particular, the set of efficient allocations of* X *is closed in* L^∞.

Proof. Let $M := \text{Esssup}X$, $m := \text{Essinf}X$ and define K_0 as the set of functions $(f_1, ..., f_p) \in C([m, M], \mathbb{R}^p)$ such that each f_i nondecreasing, $f_i(0) = 0$ and $\sum_{i=1}^p f_i(x) = x$ for every $x \in [m, M]$ and let

$$K := K_0 + \{(\lambda_1, ..., \lambda_p) \ : \ \sum_{i=1}^p \lambda_i = 0\}.$$

Convexity directly follows from theorem 3.4 and the convexity of K. Let us remark that elements of K_0 have 1-Lipschitz components and are bounded, compactness of K in $C([m, M], \mathbb{R}^p)$ then follows from Ascoli's theorem. The compactness and closedness claims then directly follow. □

Convexity and compactness of efficient allocations is a quite remarkable feature and as we will show later it is no longer true in the multivariate case. Note also that efficient allocations are regular : they are 1-Lipschitz functions of aggregate risk.

4 The multivariate case

Our aim in this section is to generalize the previous results from the univariate to the multivariate case, by showing that:

1. any allocation is dominated by a comonotone one,
2. any non comonotone allocation is strictly dominated by a comonotone one,
3. efficient sharing allocations of X coincide with comonotone ones.

What is not clear a priori, to address these generalizations, is what the appropriate notion of comonotonicity is in the multivariate framework.

4.1 From random vectors to joint laws

From now on, we consider the situtation where there are p agents and risk is d-dimensional. We shall always assume in the sequel that the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic. X is a given $\mathbb{R}^d$-valued L^∞ random vector modelling an aggregate random multivariate risk, while $\mathbf{X} = (X_1, ..., X_p)$ is a given L^∞ sharing of X among the p agents that is

$$X = \sum_{i=1}^{p} X_i.$$

We set $\gamma_0 := \mathcal{L}(\mathbf{X})$ that is the joint law of the initial allocation of X and $m_0 := \mathcal{L}(X)$. As previously, we denote by $\mathcal{A}(X)$ the set of admissible allocations or individual endowments:

$$\mathcal{A}(X) := \{\mathbf{Y} = (Y_1, ..., Y_p) \in L^\infty(\Omega, \mathbb{R}^d) : \sum_{i=1}^{p} Y_i = X\}.$$

Since we will be dealing with law invariant utilities, we will not work with the set of feasible allocations $\mathcal{A}(X)$ but instead with the set of probability measures $\{\mathcal{L}(\mathbf{Y}),\ \mathbf{Y}\in\mathcal{A}(X)\}$. If γ is a probability measure on $(\mathbb{R}^d)^p$ then we denote by γ^i its i-th marginal (in particular for $\gamma=\mathcal{L}(\mathbf{Y})$ then $\gamma^i=\mathcal{L}(Y_i)$) and we denote by $\Pi_\Sigma\gamma$ the probability measure on $\mathbb{R}^d$ defined by

$$\int_{\mathbb{R}^d}\varphi(z)d\Pi_\Sigma\gamma(z)=\int_{\mathbb{R}^{d\times p}}\varphi(\sum_{i=1}^p x_i)d\gamma(x_1,...,x_p),\ \forall\varphi\in C_0(\mathbb{R}^d,\mathbb{R}). \quad (4.1)$$

(where C_0 denotes the space of continuous function that tend to 0 at ∞). It follows from this definition that if $\gamma=\mathcal{L}(\mathbf{Y})$ then $\Pi_\Sigma\gamma=\mathcal{L}(\sum Y_i)$. Hence, if $\mathbf{Y}\in\mathcal{A}(X)$ and $\gamma=\mathcal{L}(\mathbf{Y})$ then by definition $\Pi_\Sigma\gamma=m_0=\mathcal{L}(X)$. In other words, if $\gamma=\mathcal{L}(\mathbf{Y})$ with $\mathbf{Y}\in\mathcal{A}(X)$ then

$$\int\varphi(x_1+...+x_d)d\gamma(x_1,...,x_d)=\int\varphi(z)dm_0(z),\ \forall\varphi\in C_0(\mathbb{R}^d,\mathbb{R}). \quad (4.2)$$

Let us also remark that γ is compactly supported since $\mathbf{Y}$ is bounded. Now, the fact that $\{\mathcal{L}(\mathbf{Y}),\ \mathbf{Y}\in\mathcal{A}(X)\}$ is exactly the set of compactly supported probability measures γ on $(\mathbb{R}^d)^p$ such that $\Pi_\Sigma\gamma=m_0=\Pi_\Sigma\gamma_0$ (i.e that satisfy (4.2)) follows from the next lemma:

Lemma 4.1. *If $(\Omega,\mathcal{F},\mathbb{P})$ is non-atomic, γ is a compactly supported probability measure on $(\mathbb{R}^d)^p$ and satisfies (4.2) then there exists a random vector $\mathbf{Y}=(Y_1,...,Y_p)\in\mathcal{A}(X)$ such that $\mathcal{L}(\mathbf{Y})=\gamma$. Hence we have*

$$\{\mathcal{L}(\mathbf{Y}),\ \mathbf{Y}\in\mathcal{A}(X)\}=\mathcal{M}(m_0)$$

where $\mathcal{M}(m_0)$ is the set of compactly supported probability measures on $(\mathbb{R}^d)^p$ such that $\Pi_\Sigma\gamma=m_0=\Pi_\Sigma\gamma_0$.

Proof. For notational simplicity, let us assume $d=1$, $p=2$, X takes values in $[0,2]$ a.s. (so that m_0 has support in $[0,2]$) and γ is supported by $[0,1]^2$. For every $n\in\mathbb{N}^*$ and $k\in\{0,...,2^{n+1}\}$, set

$$X^n:=\sum_{k=0}^{2^{n+1}}\frac{k}{2^n}\mathbf{1}_{A_{k,n}},\ \text{ where } A_{k,n}:=\Big\{\omega\in\Omega\ :\ X(\omega)\in\Big[\frac{k}{2^n},\frac{k+1}{2^n}\Big[\Big\}$$

and

$$C_{k,n}:=\Big\{(y_1,y_2)\in[0,1]^2\ :\ y_1+y_2\in\Big[\frac{k}{2^n},\frac{k+1}{2^n}\Big[\Big\}.$$

Let us decompose the strip $C_{k,n}$ into a partition by triangles

$$C_{k,n}=\bigcup_{k\leq i+j\leq k+1}T_{k,n}^{i,j},\ T_{k,n}^{i,j}:=C_{k,n}\cap\Big[\frac{i}{2^n},\frac{i+1}{2^n}\Big[\times\Big[\frac{j}{2^n},\frac{j+1}{2^n}\Big[.$$

Since $\Pi_\Sigma(\gamma) = m_0$ we have:

$$\mathbb{P}(A_{k,n}) = \gamma(C_{k,n}) = \sum_{k \le i+j \le k+1} \gamma(T^{i,j}_{k,n})$$

and since $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic, it follows from Lyapunov's convexity theorem (see [27]) that there exists a partition of $A_{k,n}$ into measurable subsets $A^{i,j}_{k,n}$ such that

$$\gamma(T^{i,j}_{k,n}) = \mathbb{P}(A^{i,j}_{k,n}), \forall (i,j) \in \{0, ..., 2^n\} \ : \ k \le i+j \le k+1. \tag{4.3}$$

Choose $(y_1, y_2)^{i,j}_{k,n} \in T^{i,j}_{k,n}$ and define

$$\mathbf{Y}^n = (Y^n_1, Y^n_2) := \sum_{k=0}^{2^{n+1}} \sum_{k \le i+j \le k+1} (y_1, y_2)^{i,j}_{k,n} \mathbf{1}_{A^{i,j}_{k,n}}.$$

We may also choose inductively the partition of $A_{k,n}$ by the $A^{i,j}_{k,n}$ to be finer and finer with respect to n. By construction, we then have

$$\max\Big(\|X^n - X\|_{L^\infty}, \|X^n - Y^n_1 - Y^n_2\|_{L^\infty}, \|\mathbf{Y}^{n+1} - \mathbf{Y}^n\|_{L^\infty} \Big) \le \frac{1}{2^n}$$

so that $\mathbf{Y}^n$ is a Cauchy sequence in L^∞ thus converging to some $\mathbf{Y} = (Y_1, Y_2)$. One then has $Y_1 + Y_2 = X$ and passing to the limit in (4.3), we easily get that $\mathcal{L}(\mathbf{Y}) = \gamma$. □

In the sequel, we shall work with the set of joint laws $\mathcal{M}(m_0)$ instead of the set of admissible allocations $\mathcal{A}(X)$. Doing so, we will gain both some linearity and some compactness. Slightly abusing notations, we will call allocations joint laws in $\mathcal{M}(m_0)$. Also, for compactness issues, we fix some closed ball B (centered at 0) in $\mathbb{R}^d$ such that m_0 is supported by B^p and we will restrict ourselves to the set of elements of $\mathcal{M}(m_0)$ supported by pB (that is we will only consider risk-sharings of X whose components take value in B). We shall then denote

$$\mathcal{M}_B(m_0) := \{\gamma \in \mathcal{M}(m_0) \ : \ \gamma(B^p) = 1\}.$$

4.2 Efficiency and comonotonicity in the mutivariate case

In this section, we shall define several notions of dominance (large and strict), efficiency and of comonotonicity in terms of joint law.

Let $\mathcal{C}$ be the cone of convex and continuous functions on B.

Definition 4.2. *Let γ and π be in $\mathcal{M}_B(m_0)$ then γ dominates π whenever*

$$\int_{B^p} \sum_i \varphi_i(x_i) d\gamma(x_1, ..., x_p) \le \int_{B^p} \sum_i \varphi_i(x_i) d\pi(x_1, ..., x_p) \qquad (4.4)$$

for every functions $(\varphi_1, ..., \varphi_p) \in \mathcal{C}^p$. If, in addition, inequality (4.4) is strict whenever the functions φ_i are further assumed to be stricly convex, then γ is said to dominate strictly π. The allocation $\gamma \in \mathcal{M}_B(m_0)$ is said to be efficient if there is no other allocation in $\mathcal{M}_B(m_0)$ that strictly dominates it.

One mathematical drawback of the previous definition stems from the fact that strict convexity is not a closed property which makes it difficult to attack the notion of efficiency by means of functional analytic methods. That is why we propose in what follows a somehow quantified version of strict dominance. Let $\omega := (\omega_1, ..., \omega_p)$ be a family of C^1 and strictly convex functions on $\mathbb{R}^d$, ω-dominance is then defined by

Definition 4.3. *Let γ and π be in $\mathcal{M}_B(m_0)$ then γ ω-strictly dominates π if inequality (4.4) is strict whenever the functions φ_i are such that $\varphi_i - \omega_i \in \mathcal{C}$ for $i = 1, ..., p$. The allocation $\gamma \in \mathcal{M}_B(m_0)$ is said to be ω-efficient if there is no other allocation in $\mathcal{M}_B(m_0)$ that ω-strictly dominates it.*

At this point, let us make a few remarks concerning ω-strict dominance.

Remark 4.4. Let γ and π be in $\mathcal{M}_B(m_0)$ then γ ω-strictly dominates π if and only if γ dominates π and

$$\int_{B^p} \sum_i \omega_i(x_i) d\gamma(x_1, ..., x_p) < \int_{B^p} \sum_i \omega_i(x_i) d\pi(x_1, ..., x_p). \qquad (4.5)$$

In particular for any $\lambda > 0$, γ ω-strictly dominates π if and only if γ $\lambda\omega$-strictly dominates π. Of course, γ strictly dominates π if and only if γ ω-strictly dominates π for any family of strictly convex functions ω.

Remark 4.5. Given $\gamma_0 \in \mathcal{M}_B(m_0)$ it is easy to check (taking test functions $\varphi_i = |x_i|^n$ and letting $n \to \infty$) that any $\gamma \in \mathcal{M}(m_0)$ dominating γ_0 (without the a priori restriction that it is supported on B^p) actually belongs to $\mathcal{M}_B(m_0)$. Hence our choice to only consider allocations supported by B^p is in fact not a restriction. Indeed, if γ is supported by B^p then efficiency of γ in the usual sense, i.e. without restricting to competitors supported by B^p, is *equivalent* to efficiency among competitors supported by B^p.

Remark 4.6. To clarify the concept of ω-strict dominance, let us consider the benchmark case where $\omega_i(x) = \frac{1}{2}|x|^2$. Then for any φ in $C^2(B)$ such that $D^2\varphi > 0$ (in the sense of symmetric matrices) on B one has that $\varphi - \lambda\omega_i$

is convex for small enough $\lambda > 0$. In particular, for this quadratic choice of ω_i, if γ ω-strictly dominates π, then inequality (4.4) is strict whenever the functions φ_i belong to $C^2(B)$ and have a strictly positive hessian. More generally, γ ω-strictly dominates π if inequality (4.4) is strict whenever the functions φ_i are convex and their hessian matrix (in the sense of distributions) is bounded from below on B by a strictly positive multiple of the identity matrix (which is a slightly stronger condition than strict convexity).

We shall now define comonotonicity starting from the following remark, instead of trying directly to generalize the univariate concept. Let $\psi := (\psi_1, ..., \psi_p)$ be a family of strictly convex continuous functions (defined on B) for any $x \in pB$, let us consider the following infimal convolution problem:

$$\square_i \psi_i(x) := \inf \left\{ \sum_i \psi_i(y_i) \ : \ y_i \in B, \ \sum_i y_i = x \right\}.$$

This problem admits a unique solution which we shall denote

$$T_\psi(x) := (T_\psi^1(x), ..., T_\psi^p(x)).$$

Note that, by construction

$$\sum_i T_\psi^i(x) = x, \ \forall x \in pB. \tag{4.6}$$

The well-defined map $x \mapsto T_\psi(x)$ gives the optimal way to share x so as to minimize the total cost when each individual cost is ψ_i. It thus corresponds to assigning to each agent a strictly convex cost. Of course, this map defines an efficient allocation γ_ψ defined by:

$$\gamma_\psi := (T_\psi^1, ..., T_\psi^p) \sharp m_0$$

that is

$$\int_{B^p} f(y_1, ..., y_p) d\gamma_\psi(y) := \int_{pB} f(T_\psi(x)) dm_0(x)$$

for any $f \in C(B^p)$. We then define comonotonicity as follows:

Definition 4.7. *An allocation $\gamma \in \mathcal{M}_B(m_0)$ is strictly comonotone if there exists a family of strictly convex continuous functions $\psi := (\psi_1, ..., \psi_p)$ such that $\gamma = \gamma_\psi$. Given a family of strictly convex functions in $C^1(B)$ $\omega := (\omega_1, ..., \omega_p)$, an allocation $\gamma \in \mathcal{M}_B(m_0)$ is ω-strictly comonotone if there exists a family of convex continuous functions $\psi := (\psi_1, ..., \psi_p)$ such that $\psi_i - \omega_i \in \mathcal{C}$ for every i and $\gamma = \gamma_\psi$.*

By construction, any strictly comonotone allocation is efficient and any ω-strictly comonotone allocation is ω-efficient. To study a kind of converse implication, we shall need the following (which is natural since neither ω-efficient nor ω-strictly comonotone allocations form a closed set in general):

Definition 4.8. *An allocation $\gamma \in \mathcal{M}_B(m_0)$ is comonotone if there exists a sequence of strictly comonotone allocations that weakly star converges to γ. Given a family of strictly convex functions in $C^1(B)$, $\omega := (\omega_1, ..., \omega_p)$, an allocation $\gamma \in \mathcal{M}_B(m_0)$ is ω-comonotone, if there exists a sequence of ω-strictly comonotone allocations that weakly star converges to γ.*

Definitions 4.7 and 4.8 will be discussed with more details in paragraph 4.4. To understand the previous notions of comonotonicity it is important to understand precisely the structure of the maps T_ψ. Let us first ignore regularity issues and further assume that the functions ψ_i are smooth as well as their Legendre transforms ψ_i^*. If we ignore the constraints $x_i \in B$ then the optimality conditions implies that there is some multiplier $p = p(x)$ such that

$$\nabla\psi_i(T_\psi^i(x)) = p \text{ hence } T_\psi^i(x) = \nabla\psi_i^*(p)$$

using (4.6) we get

$$x = \sum_j \nabla\psi_j^*(p) \text{ hence } p = \nabla(\sum_j \psi_j^*)^*(x)$$

so that

$$T_\psi^i(x) = \nabla\psi_i^* \left(\nabla(\sum_j \psi_j^*)^*(x) \right).$$

The maps T_ψ^i are then composed of gradient of convex functions and sum up to the identity. In dimension 1, gradients of convex functions simply are monotone maps (and then so are composed of such maps), in higher dimensions, we see a richer and more complicated structure emerging.

If we do not ignore any more the constraints that $x_i \in B$ but still assume that the ψ_i's are smooth then the optimality conditions read as the existence of a p and a $\lambda_i \geq 0$ such that

$$\nabla\psi_i(T_\psi^i(x)) = p - \lambda_i T_\psi^i(x)$$

together with the complementary slackness conditions

$$\lambda_i = 0, \text{ if } T_\psi^i(x) \text{ lies in the interior of } B.$$

4.3 A multivariate dominance result and equivalence between efficiency and comonotonicity

Let us fix an allocation $\mathbf{X} \in \mathcal{A}(X)$ such that $\mathbf{X} \in B^p$ a.s. and set $\gamma_0 = \mathcal{L}(\mathbf{X})$ so that $\gamma_0 \in \mathcal{M}_B(m_0)$. We are also given $\omega := (\omega_1, ..., \omega_p)$ a family of C^1 and strictly convex functions on B as in section 4.2. Our first main result in the multivariate case is a dominance result that is very much in the spirit of what we have already seen in dimension 1, namely that every allocation is ω-dominated by an ω-comonotone one.

Theorem 4.9. *Let $\gamma_0 = \mathcal{L}(\mathbf{X})$ and ω be as above. Then there exists $\gamma \in \mathcal{M}(m_0)$ that is ω-comonotone and dominates γ_0. Moreover if γ_0 is not itself ω-comonotone then γ ω-strictly dominates γ_0.*

The full proof of this result will be given in section 5. It heavily relies on some linear optimization program and its dual from which we can characterize some γ satisfying the claim of the previous theorem. Hence in some sense our proof is constructive.

Theorem 4.9 states that any allocation is dominated by an ω-comonotone allocation and that this dominance is ω-strict if the initial allocation is not itself ω-comonotone. In terms of efficiency, we thus have the following

Theorem 4.10. *Let $\gamma \in \mathcal{M}_B(m_0)$ and ω be as before. Then*

1. *if γ is strictly ω-comonotone then it is ω-efficient and thus efficient,*
2. *if γ is ω-efficient then it is ω-comonotone,*
3. *the closure for the weak-star topology of ω-efficient allocations coincides with the set of ω-comonotone allocations.*

Proof. 1. follows from the definition. 2. follows from theorem 4.9 and 3. follows from 1. and 2.

□

If we further specify $\omega_i(x) := \frac{|x|^2}{2}$ for every i, then we obtain as an immediate corollary:

Theorem 4.11. *Let $\gamma \in \mathcal{M}_B(m_0)$ and ω be quadratic as above. Then there is an equivalence between:*

1. *γ is in the weak star closure of weakly efficient allocations, i.e. the set of allocations π for which there is no other allocation π' such that*

$$\int_{B^p} \sum_i \varphi_i(x_i) d\pi'(x) < \int_{B^p} \sum_i \varphi_i(x_i) d\pi(x)$$

for every $\varphi_1, ..., \varphi_p$ such for any i there exists $\alpha_i > 0$ such that $D^2\varphi_i \geq \alpha_i \operatorname{id}$ on B.

2. *γ is ω-comonotone.*

4.4 Remarks on multivariate comonotonicity

Comparison with the notion of μ-comonotonicity of [18]. The notion of multivariate comonotonicity we consider in this paper is to be related to the notion of μ-comonotonicity proposed by Ekeland, Galichon and Henry in [18]. Recall the alternative characterization of comonotonicity given in the univariate case in proposition 2.8 : X_1 and X_2 are comonotone if and only if $\varrho_\mu(X_1 + X_2) = \varrho_\mu(X_1) + \varrho_\mu(X_2)$ for a measure μ that is regular enough. In dimension d, [18] have introduced the concept of μ-comonotonicity, based on this idea: if μ is a probability measure on $\mathbb{R}^d$ which does not give positive mass to small sets, two random vectors X_1 and X_2 on $\mathbb{R}^d$ are called *μ-comonotone* if and only if

$$\varrho_\mu(X_1 + X_2) = \varrho_\mu(X_1) + \varrho_\mu(X_2)$$

where the (multivariate) *maximum correlation functional* (see e.g. [33] or [18]) is defined by

$$\varrho_\mu(X) = \sup_{\tilde{Y} \sim \mu} \mathbb{E}\left(X \cdot \tilde{Y}\right).$$

The authors of [18] show that X_1 and X_2 are μ-comonotone if and only if there are two convex functions ψ_1 and ψ_2, and a random vector $U \sim \mu$ such that

$$X_1 = \nabla\psi_1(U) \text{ and } X_2 = \nabla\psi_2(U)$$

holds almost surely. Therefore, our present notion of multivariate comonotonicity approximately consists in declaring X_1 and X_2 comonotone if and only if there is some measure μ such that X_1 and X_2 are μ-comonotone. There are, however, qualifications to be added. Indeed, [18] require some regularity on the measure μ. In the current setting we do not impose regularity restrictions on μ; but we impose restrictions on the convexity of ψ_1 and ψ_2 to define our notion of ω-comonotonicity before passing to the limit. Although not equivalent, these two sets of restrictions originate from the same concern: two random vectors are always optimally coupled with very degenerate distributions, such as the distribution of constant vectors. Therefore we need to exclude these degenerate cases in order to avoid a definition which would be void of substance. This is the very reason why we introduced the strictly convex functions ω_i's.

Comonotone allocations do not form a bounded set. In the scalar case, comonotone allocations are parametrized by the set of nondecreasing functions summing to the identity map. This set of functions is convex and equilipschitz hence compact (up to adding constants summing to 0). This compactness is no longer true in higher dimension (at least when $\omega = 0$ and we work on the whole space instead of B) and we believe that this is a major structural difference with respect to the univariate case. For simplicity assume that $p = 2$, as outlined in paragraph 4.2, a comonotone allocation (X_1, X_2) of X is given by a pair of functions that are composed of gradient of convex functions and sum to the identity map. It is no longer true, in dimension 2 that this set of maps is compact (up to constants). Indeed let us take $n \in \mathbb{N}^*$, ψ_1 and ψ_2 quadratic

$$\psi_i(x) = \frac{1}{2} \left\langle S_i^{-1} x, x \right\rangle, \ i = 1, 2, \ x \in \mathbb{R}^2$$

with

$$S_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{8\sqrt{n}} \\ \frac{1}{8\sqrt{n}} & \frac{1}{2n} \end{pmatrix}, \qquad S_2 = \begin{pmatrix} \frac{1}{2} & \frac{-1}{8\sqrt{n}} \\ \frac{-1}{8\sqrt{n}} & \frac{1}{2n} \end{pmatrix},$$

the corresponding map T_ψ is linear and T_ψ^1 is given by the matrix

$$S_1(S_1 + S_2)^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{n}}{8} \\ \frac{1}{8\sqrt{n}} & \frac{1}{2} \end{pmatrix}$$

this is an unbouded sequence of matrices which proves the unboundedness claim.

Comonotone allocations do not form a convex set. Another difference with the univariate case is that the set of maps of the form T_ψ used to define comonotonicity is not convex. To see this (again in the case $p = d = 2$), it is enough to show that the set of pairs of 2×2 matrices

$K := (S_1(S_1{+}S_2)^{-1}, S_2(S_1{+}S_2)^{-1}),\ S_i$ symmetric, positive definite, $i = 1, 2\}$

is not convex. First let us remark that if $(M_1, M_2) \in K$ then M_1 and M_2 have a positive determinant. Now for $n \in \mathbb{N}^*$, and $\varepsilon \in (0, 1)$ consider

$$S_1 = \begin{pmatrix} 1 & \sqrt{1-\varepsilon} \\ \sqrt{1-\varepsilon} & 1 \end{pmatrix}, \qquad S_2 = \begin{pmatrix} 1 & -\sqrt{1-\varepsilon} \\ -\sqrt{1-\varepsilon} & 1 \end{pmatrix},$$

and

$$S_1' = \begin{pmatrix} 1 & \sqrt{n-\varepsilon} \\ \sqrt{n-\varepsilon} & n \end{pmatrix}, \qquad S_2' = \begin{pmatrix} 1 & -\sqrt{n-\varepsilon} \\ -\sqrt{n-\varepsilon} & n \end{pmatrix},$$

this defines two elements of K:

$$(M_1, M_2) = (S_1(S_1 + S_2)^{-1}, S_2(S_1 + S_2)^{-1}),$$

and

$$(M_1', M_2') = (S_1'(S_1' + S_2')^{-1}, S_2'(S_1' + S_2')^{-1}).$$

If K was convex then the matrix

$$M_1 + M_1' = \begin{pmatrix} 1 & \frac{\sqrt{1-\varepsilon}}{2} + \frac{\sqrt{n-\varepsilon}}{2n} \\ \frac{\sqrt{1-\varepsilon}}{2} + \frac{\sqrt{n-\varepsilon}}{2} & 1 \end{pmatrix},$$

would have a positive determinant which is obviously false for n large enough and ε small enough.

5 Proofs and variational characterization

Our proofs will very much rely on the following linear programming problem (which we believe to have its own interest):

$$(\mathcal{P}^*) \sup_{\gamma \in K(\gamma_0)} - \int_{B^p} \sum_i \omega_i(x_i) d\gamma(x)$$

where $K(\gamma_0)$ consists of all $\gamma \in \mathcal{M}_B(m_0)$ such that for each i the marginal γ^i dominates the corresponding marginal of γ_0 i.e.:

$$\int_{B^p} \varphi(x_i) d\gamma(x) \leq \int_{B^p} \varphi(x_i) d\gamma_0(x), \forall \varphi \text{ convex on } B.$$

Problem $(\mathcal{P}^*)$ presents some similarities with the multi-marginal Monge-Kantorovich problem solved by Gangbo and Święch in [20]. In the optimal transport problem considered in [20], one minimizes the average of some quadratic function over joint measures having prerscribed marginals whereas in $(\mathcal{P}^*)$ we have dominance constraints on the marginals. To shorten notations, let us set

$$\eta(x) := -\sum_i \omega_i(x_i), \forall x = (x_1, ..., x_p) \in B^p$$

$(\mathcal{P}^*)$ is the dual problem (see the next lemma for details) of

$$(\mathcal{P}) \inf \Big\{ \int_{B^p} \Big(\sum_i \varphi_i(x_i) - \varphi_0\Big(\sum_i x_i\Big)\Big) d\gamma_0(x), \ (\varphi_0, ..., \varphi_p) \in E \Big\}$$

where E consists of all families $\varphi := (\varphi_1, ..., \varphi_p, \varphi_0) \in C(B)^p \times C(pB)$ such $\varphi_i \in \mathcal{C}$ and

$$\sum_i \varphi_i(x_i) - \varphi_0\Big(\sum_i x_i\Big) \geq -\sum_i \omega_i(x_i).$$

It will also be convenient to consider

$$(\mathcal{Q}) \inf \Big\{ J(\psi)\ ,\ \psi = (\psi_1, ..., \psi_p) : \text{ each } \psi_i \text{ is such that } \psi_i - \omega_i \text{ is convex} \Big\}$$

with

$$J(\psi) := \int_{B^p} \Big(\sum_i \psi_i(x_i) - \Box_i \psi_i\Big(\sum_i x_i\Big)\Big) d\gamma_0(x).$$

Note that by construction $J(\psi) \geq 0$ for every admissible ψ and $J(\psi) = 0$ if and only if $\gamma_0 = \gamma_\psi$.

Lemma 5.1. *We have*

$$\max(\mathcal{P}^*) = \inf(\mathcal{P}) = \inf(\mathcal{Q}) - \int_{B^p} \sum_i \omega_i(x_i) d\gamma_0(x)$$

Proof. Let us write $(\mathcal{P})$ in the form

$$\inf_{\varphi = (\varphi_1, ..., \varphi_p, \varphi_0) \in C(B)^p \times C(pB)} F(\Lambda\varphi) + G(\varphi)$$

where Λ is the linear continuous map $C(B)^p \times C(pB) \to C(B^p)$ defined by

$$\Lambda\varphi(x) := \sum_i \varphi_i(x_i) - \varphi_0\Big(\sum_i x_i\Big),\ \forall x = (x_1, ..., x_p) \in B^p,$$

and F and G are the convex lsc (for the uniform norm of course) functionals defined respectively by

$$F(\theta) = \begin{cases} \int_{B^p} \theta d\gamma_0 \text{ if } & \theta \geq \eta \\ +\infty & \text{otherwise} \end{cases}.$$

for any $\theta \in C(B^p)$ and

$$G(\varphi) = \begin{cases} 0 \text{ if } & (\varphi_1, ..., \varphi_p) \in \mathcal{C}^p \\ +\infty & \text{otherwise.} \end{cases}$$

for any $\varphi = (\varphi_1, ..., \varphi_p, \varphi_0) \in C(B)^p \times C(pB)$. Since η is bounded on B^p, it is easy to see that $\inf(\mathcal{P})$ is finite and choosing φ of the form $(M, 0, ..., 0)$ with M constant such that $M \geq \eta + 1$ on B^p, we have $G(\varphi) = 0$ and F

continuous at $\Lambda\varphi$, it thus follows from Fenchel-Rockafellar's duality theorem (see for instance [17]) that one has

$$\inf(\mathcal{P}) = \max_{\gamma\in\mathcal{M}(B^p)} -F^*(\gamma_0-\gamma) - G^*(\Lambda^*(\gamma-\gamma_0)).$$

Note that the fact that the the sup is attained in the primal is part of the theorem. The adjoint of Λ, Λ^* is easily computed as : $\mathcal{M}(B^p)\to\mathcal{M}(B)^p\times\mathcal{M}(pB)$ (where $\mathcal{M}$ denotes the space of Radon measures):

$$\Lambda^*\gamma = (\gamma^1, ..., \gamma^p, -\Pi_\Sigma\gamma),\ \forall\gamma\in\mathcal{M}(B^p).$$

Direct computations give

$$F^*(\gamma-\gamma_0) = \begin{cases} -\int_{B^p}\eta d\gamma \text{ if } & \gamma\geq 0 \\ +\infty & \text{otherwise .}\end{cases}$$

and

$$G^*(\Lambda^*(\gamma-\gamma_0)) = \begin{cases} 0 \text{ if } & \gamma\in K(\gamma_0) \\ +\infty & \text{otherwise .}\end{cases}$$

We then have that $(\mathcal{P}^*)$ is the dual of $(\mathcal{P})$ in the usual sense of convex programming and

$$\max(\mathcal{P}^*) = \inf(\mathcal{P}).$$

To prove that

$$\inf(\mathcal{P}) = \inf(\mathcal{Q}) - \int_{B^p}\sum_i \omega_i(x_i)d\gamma_0(x)$$

let us take φ admissible for $(\mathcal{P})$ and set $\psi_i := \omega_i + \varphi_i$ for $i = 1, .., p$, the constraint then reads as

$$\sum_i \psi_i(x_i) \geq \varphi_0\Big(\sum_i x_i\Big),\ \forall x\in B^p.$$

Now in $(\mathcal{P})$, one wants to make φ_0 as large as possible without violating this constraint, the best φ_0 given $(\varphi_1, ..., \varphi_p)$ is then

$$\varphi_0 = \Box_i\psi_i,$$

this proves the desired identity.

□

Lemma 5.2. *Let ψ_i be such that $\psi_i-\omega_i\in\mathcal{C}$ for every i and $g=(g_1,...,g_p)\in\mathcal{C}^p$ then*

$$\lim_{\delta\to 0^+}\frac{1}{\delta}[J(\psi+\delta g)-J(\psi)] = \sum_i\int_B g_i(x_i)d(\gamma_0^i-\gamma_\psi^i)$$

$$= \int_{B^p}\sum_i g_i(x_i)d(\gamma_0-\gamma_\psi)(x)$$

Proof. For $\delta > 0$, we first have

$$\frac{1}{\delta}[J(\psi+\delta g)-J(\psi)] = \sum_i \int_B g_i(x_i)d(\gamma_0^i) - \int_{pB} \frac{1}{\delta}\Big(\square_i(\psi_i+\delta g_i)(x) - \square_i\psi_i(x)\Big)dm_0(x).$$

And let us remark that the integrand in the second term is bounded since g is. Let us then fix some $(x_1, ..., x_p) \in B^p$ and set $x = \sum_i x_i$, $y_i := T^i_\psi(x)$ and $y_i^\delta := T_{\psi+\delta g}(x)$. Since $\sum_i y_i = \sum_i y_i^\delta = x$, we have as a direct consequence of the definition of infimal convolutions:

$$\frac{1}{\delta}\Big(\square_i(\psi_i+\delta g_i)(x) - \square_i\psi_i(x)\Big) \leq \sum_i g_i(y_i) \tag{5.1}$$

and

$$\frac{1}{\delta}\Big(\square_i(\psi_i+\delta g_i)(x) - \square_i\psi_i(x)\Big) \geq \sum_i g_i(y_i^\delta). \tag{5.2}$$

Using the compactness of B and the strict convexity of ψ_i, it is easy to check that $y_i^\delta \to y_i$ as $\delta \to 0^+$. We thus deduce from (5.1) and (5.2) that

$$\lim_{\delta\to 0^+} \frac{1}{\delta}\Big(\square_i(\psi_i+\delta g_i)(x) - \square_i\psi_i(x)\Big) = \sum_i g_i(T^i_\psi(x))$$

and this holds for every $x \in pB$. It then follows from Lebesgue's dominated convergence Theorem that

$$\begin{aligned}\lim_{\delta\to 0^+} \frac{1}{\delta}[J(\psi+\delta g)-J(\psi)] &= \sum_i \int_B g_i(x_i)d(\gamma_0^i) - \sum_i \int_{pB} g_i(T^i_\psi(x))dm_0(x)\\ &= \sum_i \int_B g_i(x_i)d(\gamma_0^i - \gamma_\psi^i) = \int_{B^p} \sum_i g_i(x_i)d(\gamma_0 - \gamma_\psi)(x).\end{aligned}$$

□

It follows from the previous lemma that, if ψ solves $(\mathcal{Q})$, then γ_ψ dominates γ_0. Hence, if we knew that $(\mathcal{Q})$ possesses solutions, the existence of an ω-strictly comonotone allocation dominating γ_0 would directly follow. Unfortunately, it is not necessarily the case that the infimum in $(\mathcal{Q})$ is attained-or at least we haven't been able to prove without additional conditions- the difficulty coming from the fact that minimizing sequences need not be bounded (see paragraph 4.4). Maybe additional assumptions on γ_0 (here we recall that we made no assumption such as absence of atoms) would guarantee existence, but in the following, we shall overcome the difficulty by using Ekeland's variational principle:

Lemma 5.3. *Let $\varepsilon > 0$, there exists ψ_ε admissible for $(\mathcal{Q})$ such that*

1. $J(\psi_\varepsilon) \leq \inf(\mathcal{Q}) + \varepsilon$

2. $$\limsup_{\varepsilon\to 0^+} \int_{B^p} \sum_i \varphi_i(x_i) d(\gamma_{\psi_\varepsilon} - \gamma_0) \leq 0$$

   for every $(\varphi_1, ..., \varphi_p) \in \mathcal{C}^p$

3. $$\liminf_{\varepsilon\to 0^+} \int_{B^p} \sum_i \varphi_i^\varepsilon(x_i) d(\gamma_{\psi_\varepsilon} - \gamma_0) \geq 0$$

   for $\varphi_i^\varepsilon = \psi_{i,\varepsilon} - \omega_i$ (these are convex functions by definition).

Proof. For $\varepsilon > 0$, let f_ε be admissible for $(\mathcal{Q})$ and such that

$$J(f_\varepsilon) \leq \inf(\mathcal{Q}) + \varepsilon.$$

Let then $k_\varepsilon > 0$ be such that

$$\lim_{\varepsilon\to 0^+} \varepsilon k_\varepsilon [1 + \|f_\varepsilon\|] = 0 \;\; \text{(for instance } k_\varepsilon = \frac{1}{\varepsilon^{1/2}(1 + \|f_\varepsilon\|)}\text{)}. \tag{5.3}$$

It follows from Ekeland's variational principle (see [16] and [4]) that for every $\varepsilon > 0$, there is some ψ_ε admissible for $(\mathcal{Q})$ such that

$$\|\psi_\varepsilon - f_\varepsilon\| \leq \frac{1}{k_\varepsilon}, \;\; J(\psi_\varepsilon) \leq J(f_\varepsilon) \leq \inf(\mathcal{Q}) + \varepsilon \tag{5.4}$$

(where $\|h\|$ stands for the sum of the uniform norms of the h_i's) and:

$$J(\psi) \geq J(\psi_\varepsilon) - k_\varepsilon \varepsilon \|\psi - \psi_\varepsilon\|, \; \forall \psi = (\psi_1, ..., \psi_p) \; : \; \psi_i - \omega_i \in \mathcal{C}, \; \forall i. \tag{5.5}$$

Taking $\psi = \psi_\varepsilon + \delta\varphi$ with $\delta > 0$ and $\varphi \in \mathcal{C}^p$ in (5.5), dividing by δ and letting $\delta \to 0^+$, we thus get thanks to lemma 5.2

$$\int_{B^p} \sum_i \varphi_i(x_i) d(\gamma_0 - \gamma_{\psi_\varepsilon}) \geq -k_\varepsilon \varepsilon \|\varphi\|. \tag{5.6}$$

Using (5.3) and letting $\varepsilon \to 0^+$ we then obtain:

$$\limsup_{\varepsilon\to 0^+} \int_{B^p} \sum_i \varphi_i(x_i) d(\gamma_{\psi_\varepsilon} - \gamma_0) \leq 0 \tag{5.7}$$

for every $(\varphi_1, ..., \varphi_p) \in \mathcal{C}^p$. Let us finally prove the last assertion of the lemma; let us write $\psi_\varepsilon = \varphi^\varepsilon + \omega$ with $\varphi^\varepsilon \in \mathcal{C}^p$, then for $\delta \in (0,1)$ we have $\psi_\varepsilon - \delta\varphi^\varepsilon = (1-\delta)\varphi^\varepsilon + \omega$ and then we may apply (5.5) to $\psi_\varepsilon - \delta\varphi^\varepsilon$, this yields

$$\frac{1}{\delta}[J(\psi_\varepsilon - \delta\varphi^\varepsilon) - J(\psi_\varepsilon)] \geq -k_\varepsilon \varepsilon \|\varphi^\varepsilon\|$$

letting $\delta \to 0^+$ and arguing as in lemma 5.2, we obtain:

$$\int_{B^p} \sum_i \varphi_i^\varepsilon(x_i) d(\gamma_{\psi_\varepsilon} - \gamma_0) \geq -k_\varepsilon \varepsilon \|\varphi^\varepsilon\|.$$

Thanks to (5.3) and (5.4), we have

$$k_\varepsilon \varepsilon \|\varphi^\varepsilon\| \leq k_\varepsilon \varepsilon (\|\omega\| + \|\psi_\varepsilon - f_\varepsilon\| + \|f_\varepsilon\|) \leq k_\varepsilon \varepsilon \|\omega\| + \varepsilon + k_\varepsilon \varepsilon \|f_\varepsilon\| \to 0 \text{ as } \varepsilon \to 0^+.$$

This enables us to conclude that

$$\liminf_{\varepsilon \to 0^+} \int_{B^p} \sum_i \varphi_i^\varepsilon(x_i) d(\gamma_{\psi_\varepsilon} - \gamma_0) \geq 0. \tag{5.8}$$

□

Lemma 5.4. *Let ψ_ε be as in lemma 5.3 and set $\gamma_\varepsilon := \gamma_{\psi_\varepsilon}$ then up to some subsequence, γ_ε weakly star converges to some γ (ω-comonotone by construction) such that $\gamma \in \mathcal{M}_B(m_0)$ and γ dominates γ_0. Moreover γ solves $(\mathcal{P}^*)$.*

Proof. By the Banach-Alaoglu-Bourbaki theorem, we may indeed assume that γ_ε weakly star converges to some γ. Obviously, γ is ω-comonotone and $\Pi_\Sigma \gamma = \Pi_\Sigma \gamma_0 = m_0$ hence $\gamma \in \mathcal{M}_B(m_0)$. The fact that γ dominates γ_0 directly follows from letting $\varepsilon \to 0^+$ in (5.7). Let us finally prove that γ solves $(\mathcal{P}^*)$. Defining $\varphi^\varepsilon := \psi_\varepsilon - \omega$ as in lemma 5.3 we have:

$$J(\psi_\varepsilon) = \int_{B^p} \sum_i \varphi_i^\varepsilon(x_i) d(\gamma_0 - \gamma_\varepsilon) + \int_{B^p} \eta d(\gamma_\varepsilon - \gamma_0) \to \inf(\mathcal{Q}) \text{ as } \varepsilon \to 0^+.$$

Passing to the limit in (5.8) thus yields

$$\inf(\mathcal{Q}) \leq \int_{B^p} \eta d(\gamma - \gamma_0)$$

which with lemma 5.1 gives:

$$\int_{B^p} \eta d\gamma \geq \inf(\mathcal{Q}) + \int_{B^p} \eta d\gamma_0 = \max(\mathcal{P}^*)$$

so that γ solves $(\mathcal{P}^*)$.

□

Lemma 5.5. *Let γ be as in lemma 5.4, then:*

1. *if γ_0 solves $(\mathcal{P}^*)$ then γ_0 is ω-comonotone,*

2. *γ ω-strictly dominates γ_0 unless γ_0 is itself ω-comonotone.*

Proof. If γ_0 solves $(\mathcal{P}^*)$, it follows from lemma 5.1 that $\inf(\mathcal{Q}) = 0$. For any minimizing sequence ψ_ε (not necessarily the one constructed in lemma 5.3) of $(\mathcal{Q})$ we thus have

$$0 = \lim_{\varepsilon\to 0^+} J(\psi_\varepsilon) = \lim_{\varepsilon\to 0^+} \int_{B^p} \Big(\sum_i \psi_{i,\varepsilon}(x_i) - \Box_i \psi_{i,\varepsilon}\Big(\sum_i x_i\Big)\Big) d\gamma_0(x)$$
$$= \lim_{\varepsilon\to 0^+} \int_{B^p} \Big(\sum_i \psi_{i,\varepsilon}(x_i) - \sum_i \psi_{i,\varepsilon}\Big(T^i_{\psi_\varepsilon}\Big(\sum_i x_i\Big)\Big)\Big) d\gamma_0(x).$$

By density, we may consider a minimizing sequence ψ_ε such that each ψ_ε belongs to $C^1(B)$. Let us fix $(x_1, ..., x_p)$ and set $x := \sum_i x_i$, $y^\varepsilon := T_{\psi_\varepsilon}(x)$ can be characterized as follows : there is a $p \in \mathbb{R}^d$ and nonnegative λ_i's such that

$$\nabla \psi_{i,\varepsilon}(y_i^\varepsilon) = p - \lambda_i y_i^\varepsilon,\ \lambda_i = 0 \text{ if } y_i^\varepsilon \notin \partial B,\ \sum_i y_i^\varepsilon = x. \tag{5.9}$$

On the other hand since ω_i is strictly convex and $\psi_{i,\varepsilon} - \omega_i \in \mathcal{C}$ for any a and b in B^2 one has

$$\psi_{i,\varepsilon}(b) - \psi_{i,\varepsilon}(a) \geq \nabla \psi_{i,\varepsilon}(a) \cdot (b-a) + \theta_i(|b-a|) \tag{5.10}$$

where the function θ_i is defined by, for any $t \in [0, \mathrm{diam}(B)]$

$$\theta_i(t) := \inf\{\omega_i(b) - \omega_i(a) - \nabla\omega_i(a)\cdot(b-a),\ (a,b) \in B^2,\ |a-b| \geq t\}.$$

The function θ_i (modulus of strict convexity of ω_i) is a nondecreasing function such that $\theta_i(0) = 0$ and $\theta_i(t) > 0$ for $t > 0$. Combining (5.9) and (5.10), we get

$$\sum_i \psi_{i,\varepsilon}(x_i) - \sum_i \psi_{i,\varepsilon}(y_i^\varepsilon) \geq \sum_i \nabla\psi_{i,\varepsilon}(y_i^\varepsilon)\cdot(x_i - y_i^\varepsilon) + \sum_i \theta_i(|x_i - y_i^\varepsilon|)$$
$$= p \cdot \sum_i (x_i - y_i^\varepsilon) - \sum_i \lambda_i y_i^\varepsilon (x_i - y_i^\varepsilon) + \sum_i \theta_i(|x_i - y_i^\varepsilon|)$$
$$\geq \sum_i \theta_i(|x_i - y_i^\varepsilon|).$$

Hence the fact that $J(\psi_\varepsilon) \to 0$ as $\varepsilon \to 0^+$ gives

$$\lim_{\varepsilon\to 0^+} \int_{B^p} \sum_i \theta_i(|x_i - T^i_{\psi_\varepsilon}(\sum_j x_j)|) d\gamma_0(x) = 0$$

so that

$$T_{\psi_\varepsilon}(\sum_j x_j) - x \to 0 \text{ as } \varepsilon \to 0^+ \text{ for } \gamma_0\text{-a.e. } x.$$

By Lebesgue's dominated convergence theorem, we thus have for all $f \in C(B^p)$:

$$\int_{B^p} f(x)d\gamma_0(x) = \lim_{\varepsilon\to 0^+} \int_{B^p} f(T_{\psi_\varepsilon}(\sum_j x_j))d\gamma_0(x) = \lim_{\varepsilon\to 0^+} \int_{pB} f(T_{\psi_\varepsilon}(x))dm_0(x)$$
$$= \lim_{\varepsilon\to 0^+} \int_{B^p} f d\gamma_{\psi_\varepsilon}.$$

Hence, $\gamma_{\psi_\varepsilon}$ weakly star converges to γ_0 which proves that γ_0 is ω-comonotone.

Let us now prove 2. If γ_0 is not ω-comonotone then by 1., it does not solve $(\mathcal{P}^*)$ and thus $\int \eta d(\gamma - \gamma_0) > 0$ so that

$$\int_{B^p} \sum_i \omega_i(x_i)d\gamma < \int_{B^p} \sum_i \omega_i(x_i)d\gamma_0$$

and then γ ω-strictly dominates γ_0 (see remark 4.4).

□

6 Concluding remarks

In this paper, we have first revisited Landsberger and Meilijson's comonotone dominance principle. We gave a self-contained proof using monotone rearrangements. Actually, a second possible proof (based on the variational scheme we introduced in the multivariate setting), which directly covers the many agents case works as follows. Let $X \in L^\infty$, $\mathbf{X} = (X_1, ..., X_p) \in \mathcal{A}(X)$ and $\gamma_0 := \mathcal{L}(\mathbf{X})$. Let (ψ_n) be a minimizing sequence for problem $(\mathcal{Q})$. There is no reason for (ψ_n) to be bounded in general. However, in one dimension each maps $x \mapsto T^i_{\psi_n}$ is monotone (as a composition of two monotone maps). By Ascoli's theorem, one may therefore assume that $T_{\psi_n}(X)$ converges to some comonotone allocation of X. By the same arguments as in section 5, this comonotone allocation dominates X.

We have then extended the univariate theory of efficient risk-sharing to the case of several goods without perfect substituability, and we derived tractable implications. The main findings of this work are the following:

- *the intrinsic difficulty of the multivariate case,* as many features of the univariate case do not extend to higher dimensions: computational ease, the compactness and convexity of efficient risk-sharing allocations.

- *the relevance of specific techniques, in particular convex programming.* While the analysis in [18] rests upon optimal transportation techniques, the present paper heavily relies on convex programming and variational analysis (e.g. Ekeland's variational principle) to handle this type of questions. Indeed, most of the questions related to convex ordering can be restated in terms of cone programming problem and can be attacked by duality.

- *the need for qualification.* Contrary to the univariate case, the need to quantify strict convexity as we did in this paper comes by no coincidence. In fact, just as the authors of [18] impose regularity conditions on their "baseline measure" to avoid degeneracy, we work with cones which are strictly included in the cone of convex functions by quantifying the strict convexity of the functions used.

References

[1] K.J. Arrow, Uncertainty and the welfare of medical care, Amer. Econom. Rev. 53 (1963), 941–973.

[2] K.J. Arrow, Optimal insurance and generalized deductibles, Scand. Actuar. J. (1974), 1–42.

[3] A.B. Atkinson, On the measurement of Inequality, Journal of Economic Theory 2 (1970) 244–263.

[4] J.-P. Aubin, I. Ekeland, Applied Nonlinear Analysis, John Wiley ans Sons, New-York, 1984.

[5] K. Borch, Equilibrium in a reinsurance market, Econometrica, 30 (1962), 424–444.

[6] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, Comm. Pure Appl. Math. 44 (1991), 375–417.

[7] G. Carlier, R.-A. Dana, Rearrangement inequalities in nonconvex economic models, J. Math. Econ. 41 (2005), 483–503.

[8] G. Carlier, R.-A. Dana, Law invariant concave utility functions and optimization problems with monotonicity and comonotonicity constraints, Statistics and Decisions, 24 (2006), 127–152.

[9] G. Carlier, R.-A. Dana, Two-Persons Efficient Risk-Sharing and Equilibria for Concave Law-Invariant Utilities, Economic Theory, 36 (2008), 189–223.

[10] S.H. Chew, I. Zilcha, Invariance of the efficient sets when the expected utility hypothesis is relaxed, Journal of Economic Behaviour and Organisation 13 (1990) 125–131.

[11] K.M. Chong, N.M. Rice, Equimeasurable Rearrangements of functions, Queen's Papers in Pure and Applied Mathematics 28, 1971.

[12] R.-A. Dana, Stochastic Dominance, Risk, Dispersion and Equilibrium Asset Pricing , Journal of Mathematical Economic 40, (2004), 619-639.

[13] R.-A. Dana, I.I Meilijson, Modelling Agents' Preferences in Complete Markets by. Second Order Stochastic Dominance, working paper (2003)

[14] D. Denneberg, Non-additive Measures and Integral, Kluwer Academic Publishers, Holland, 1994.

[15] P. Dybvig, Distributional Analysis of Portfolio Choice, Journal of Business 61 (1988) 369–393.

[16] I. Ekeland, On the variational principle. J. Math. Anal. Appl. 47 (1974) 324–353.

[17] I. Ekeland, R. Temam, Convex analysis and variational problems, North-Holland, Amsterdam-Oxford, 1976.

[18] I. Ekeland, A. Galichon, M. Henry, Comonotonic measures of multivariate risk, available at SSRN: http://ssrn.com/abstract=1115729 (2009).

[19] H. Föllmer, A. Schied, Stochastic finance. An introduction in discrete time , De Gruyter editor, Berlin 2004.

[20] W. Ganbo, A. Święch, Optimal maps for the multidimensional Monge-Kantorovich problem, Comm. Pure Appl. Math., 51 (1998), 23–45.

[21] E. Jouini, C. Napp, Comonotonic processes, Insurance Math. Econom. 32 (2003), 255–265.

[22] E. Jouini, C. Napp, Conditional comonotonicity, Decis. Econ. Finance 27 (2004), 153–166.

[23] E. Jouini , H. Kallal, (2000) Efficient Trading Strategies in the Presence of Market Frictions, Review of Financial Studies 14 (2000) 343–369.

[24] E. Jouini, W. Schachermayer, N. Touzi, Optimal risk sharing for law invariant monetary utility functions, Math. Finance 18 (2008), 269–292.

[25] M. Landsberger, I.I. Meilijson, Comonotone allocations, Bickel Lehmann dispersion and the Arrow-Pratt measure of risk aversion, Annals of Operation Research 52 (1994) 97–106.

[26] S.F. Leroy, J. Werner, Principles of Financial Economics. Cambridge University Press, Cambridge, 2001.

[27] A. Lyapunov, Sur les fonctions-vecteurs completement additives, Bull. Acad. Sci. URSS (6) (1940), 465-478.

[28] M. Ludkovski, L. Rüschendorf, On comonotonicity of Pareto optimal risk sharing, Statistics and Probability Letters 78 (2008), 1181-1188.

[29] A. Müller, D. Stoyan, Comparaisons Methods for Stochastic Models and Risks , Wiley, New-York, 2002.

[30] B. Peleg, M.E. Yaari, A Price Characterisation of Efficient Random Variables, Econometrica 43 (1975) 283–292.

[31] G. Puccetti, M. Scarsini, Multivariate comonotonicity, Journal of Multivariate Analysis 101 (2010), 291–304.

[32] M. Rothschild, J. E. Stiglitz,. Increasing Risk, I. A Definition, Journal of Economic Theory 2 (1970) 225–243.

[33] L. Rüschendorf, Ordering of distributions and Rearrangement of functions, Annals of probability 9 (1980), 276–283.

[34] J.V. Ryff, Measure preserving Transformations and Rearrangements, Journal of Mathematical Analysis and Applications 31 (1970) 449- 458.

[35] R.M. Townsend, Risk and Insurance in village india, Econometrica 62 (1904) 539-592.

[36] R. Wilson, The theory of syndicates, Econometrica 55 (1968) 95-115.