

Extreme dependence for multivariate data

Alfred Galichon, Damien Bosc

▶ **To cite this version:**

Alfred Galichon, Damien Bosc. Extreme dependence for multivariate data. 2010. hal-03588294

HAL Id: hal-03588294

https://sciencespo.hal.science/hal-03588294

Preprint submitted on 24 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Quantitative Finance

ISSN: 1469-7688 (Print) 1469-7696 (Online) Journal homepage: http://www.tandfonline.com/loi/rquf20

Extreme dependence for multivariate data

Damien Bosc & Alfred Galichon

To cite this article: Damien Bosc & Alfred Galichon (2014) Extreme dependence for multivariate data, Quantitative Finance, 14:7, 1187-1199, DOI: 10.1080/14697688.2014.886777

To link to this article: http://dx.doi.org/10.1080/14697688.2014.886777

Published online: 31 Mar 2014.

Submit your article to this journal

Article views: 136

View related articles

View Crossmark data

Citing articles: 2 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=rquf20

Download by: [FNSP Fondation National des Sciences Politiques] **Date:** 06 July 2017, At: 06:09

Quantitative Finance, 2014
Vol. 14, No. 7, 1187–1199, http://dx.doi.org/10.1080/14697688.2014.886777

Routledge
Taylor & Francis Group

Extreme dependence for multivariate data

DAMIEN BOSC*† and ALFRED GALICHON‡

†EDF R&D, 1 avenue du Général de Gaulle, Clamart 92140, France
‡Sciences Po, Department of Economics, 28 rue des Saints-Pères, Paris 75007, France

(*Received 16 June 2013; accepted 16 January 2014*)

This article proposes a generalized notion of extreme multivariate dependence between two random vectors which relies on the extremality of the cross-covariance matrix between these two vectors. Using a partial ordering on the cross-covariance matrices, we also generalize the notion of positive upper dependence. We then propose a means to quantify the strength of the dependence between two given multivariate series and to increase this strength while preserving the marginal distributions. This allows for the design of stress-tests of the dependence between two sets of financial variables that can be useful in portfolio management or derivatives pricing.

Keywords: Multivariate dependence; Extreme dependence; Covariance set

JEL Classification: C58, C02

1. Introduction

Extreme dependence and the closely related notion of comonotonicity are important concepts in various fields. They are central in the economics of insurance (following the seminal work of Borch (1962), Arrow (1963), and Arrow (1970)), in economic theory (see Yaari (1987), Landsberger and Meilijson (1994), and Schmeidler (1989)), in statistics (see Dall'Aglio (1956), Rüschendorf (1990), Rachev (1991), Zolotarev (1983)), as well as in financial risk management (see the book by Malevergne and Sornette (2006) and references therein).

The notion of extreme (positive) dependence or comonotonicity for univariate random variables goes back to the work of Hoeffding (1940) and Fréchet (1951). Two real random variables (X, Y) are comonotonic if their cumulative distribution function satisfy $F_{X,Y}(x, y) = \min(F_X(x), F_Y(y))$, or equivalently if their copula C is the upper Fréchet copula $C(u_1, u_2) = \min(u_1, u_2)$. Equivalently X and Y can be written as nondecreasing functions of a third random variable Z. As a consequence, comonotone variables maximize covariance over the set of pairs with fixed marginals:

$$\mathbf{E}(XY) = \sup_{\substack{\tilde{X}\sim X \\ \tilde{Y}\sim Y}} \mathbf{E}(\tilde{X}\tilde{Y}), \tag{1}$$

where $\tilde{X} \sim X$ denotes equality in distribution between $\tilde{X}$ and X. Similarly, X and Y are said to have extreme negative dependence when X and $-Y$ have extreme positive dependence. Their covariance is then minimal instead of maximal, and their copula is the lower Fréchet copula, $C(u, v) = \max(u + v - 1, 0)$.

The present article aims at proposing an operational theory of extreme dependence in the multivariate case, that is when X and Y are random vectors. Our contribution is twofold. First, we introduce (in definition 3.2) a generalization of the notion of extreme dependence to the multivariate case, and we investigate how extreme positive dependence generalizes in this setting. We also introduce a notion of positive extreme dependence (in definition 4.1). Next, we introduce a measure of the strength of dependence based on an entropic measure (in section 5). We then show how useful the concept of extreme dependence, either in risk management or in asset pricing, can be.

Generalizing extreme dependence. When dealing with the multivariate case, where X and Y are random vectors in $\mathbf{R}^d$, there is no canonical way to generalize this notion of (positive or negative) extreme dependence and Fréchet copula. One approach, based on the theory of Optimal Transport (see e.g. the books Rachev and Rüschendorf (1998) and Villani (2003)), would be to consider the following optimization problem

$$\max_{\substack{\tilde{X}\sim X \\ \tilde{Y}\sim Y}} \mathbf{E}(\tilde{X} \cdot \tilde{Y}), \tag{2}$$

where $\cdot$ is the scalar product in $\mathbf{R}^d$. This program is a multivariate extension of the covariance maximization problem (1) and defines as extreme the distribution of the pair $(\tilde{X}, \tilde{Y})$ solution to the above problem. However, it does not take into account the cross-dependence between X_i and Y_j for $i \neq j$.

A more satisfactory generalization is based on the idea that both positive and negative extreme dependences are obtained

*Corresponding author. Email: damien.bosc@polytechnique.edu

© 2014 Taylor & Francis

by the maximization of the expectation of a nonzero bilinear form in (X, Y) over the set of couplings of X and Y (i.e. joint distributions with fixed marginals). In other words, we consider solutions of (2), where the scalar product is replaced by a nonzero bilinear form. This will be our notion of *multivariate extreme dependence*: random vectors X and Y are said to exhibit extreme dependence if their cross-covariance matrix maximizes the expected value of a nonzero bilinear form over all the couplings of X and Y. These extreme couplings are proposed as a generalization of Fréchet (positive and negative) extreme dependence in the multivariate case. We provide a natural geometric characterization of this notion by considering the *covariance set*, that is the set of all cross-covariance matrices $\mathbf{E}(XY')$ for all the couplings of X and Y. We show that X and Y have extreme dependence if and only if their cross-covariance matrix lies on the boundary of the covariance set.

We then turn to generalizing the notion of extreme *positive* dependence. One natural way to generalize extreme positive dependence is to look for the couplings (X, Y) having a cross-covariance matrix $Cov(X, Y) = \mathbf{E}(XY') = (\mathbf{E}(X_i Y_j))_{i,j}$, which would be maximal for a certain partial (conical) ordering on matrices. As we shall see, it turns out that extreme positive dependence implies extreme dependence, and we characterize the geometric locus of extreme positive dependent vectors on the covariance set.

Stress-testing dependence. We give a method to associate any coupling, for example any empirical coupling, with an extreme coupling, by means of an entropic relaxation technique. An algorithm is described and results concerning its implementation are given. In particular, this algorithm provides a means to compute effectively the covariance set. We then apply these results to build stress-tests of multivariate dependence for portfolio management and to price derivatives on multiple underlyings. We also propose the construction of indices of maximal dependence, that is linear combinations of assets that have remarkable properties of extreme dependence.

The present article is organized as follows: the next section presents the notion of covariance set and the definition of couplings with extreme dependence, as well as a characterization of such couplings. The third section defines and characterizes couplings with positive extreme dependence, in relation to the notion of extreme dependence. The fourth section provides an algorithm to compute extreme couplings and the covariance set. An index of dependence, the affinity matrix, is utilized; a method to associate any coupling with an extreme coupling is described. We conclude with financial applications, namely stress-testing portfolio allocations and options pricing, as well as the computation of indices with extreme dependence. All proofs are collected in Appendix B.

Notation, definitions Let P and Q be two probability distributions on $\mathbf{R}^I$ and $\mathbf{R}^J$, with finite second-order moments. Without restricting the generality, we assume that P and Q have null expected values, so that the second-order moments $\mathbf{E}(X_i Y_j)$ are indeed covariances. $\Pi(P, Q)$ is the set of all probability distributions over $\mathbf{R}^I \times \mathbf{R}^J$ having marginals P and Q. We refer to an element of $\Pi(P, Q)$ as a *coupling*, understating the probabilities P and Q. If M and N belong to $\mathbf{M}_{I,J}(\mathbf{R})$, the set of real matrices of size $I \times J$, their scalar product is denoted by $M \cdot N = Tr\left(M'N\right)$. If $(X, Y) \sim \pi \in \Pi(P, Q)$, we denote indifferently $\sigma_{X,Y}$ or σ_π the matrix with general term $\mathbf{E}(X_i Y_j)$, which is the covariance between X_i and Y_j; it is the cross-covariance matrix between X and Y. Remark that $\sigma_{X,Y}$ is the upper-right block of the variance-covariance matrix of the vector $Z = (X, Y)'$, and that $\sigma_{X,Y}$ is neither a square matrix nor a symmetric matrix in general.

Moreover, we will say that a coupling π 'projects' onto σ_π, interpreting the function $\pi \mapsto \sigma_\pi$ as a projection operator.

Eventually, let us recall that the *subdifferential* $\partial f(x_0)$ of a convex function on $\mathbf{R}^n$ at a point x_0 is defined as set of vectors v such that $f(x) - f(x_0) \geq v \cdot (x - x_0)$ for all $x \in \mathbf{R}^n$. Here the dot is the usual scalar product. It reduces to $\{\nabla f(x_0)\}$ if f is differentiable at x_0, which is true for almost every x_0 according to the Rademacher theorem.

2. Related literature and contribution

As mentioned in the introduction, the extension to the multivariate setting of the correlation maximization problem (1) has been tackled by several authors in order to define notions of *multivariate comonotonicity*. Puccetti and Scarsini (2010) list several possible definitions of multivariate comonotonicity, among which one is directly related to the variational problem (2). Namely, *c-comonotonicity* refers to the couplings solving problem (2): these are the optimal quadratic couplings of Optimal Transport Theory, also called *maximum correlation couplings*. This variational approach to multivariate comonotonicity is also the basis of Ekeland *et al.* (2012) and Galichon and Henry (2012). They propose to extend the univariate notion of comonotonicity and define the *μ-comonotonicity* by stating that two vectors X and Y are μ-comonotone if there exists a random vector $U \sim \mu$ such that

$$\mathbf{E}(X \cdot U) = \max\{\mathbf{E}(X \cdot \tilde{U}),\ \tilde{U} \sim \mu\},$$
$$\mathbf{E}(Y \cdot U) = \max\{\mathbf{E}(Y \cdot \tilde{U}),\ \tilde{U} \sim \mu\}.$$

This notion of comonotonicity has the advantage of being transitive, unlike c-comonotonicity. Carlier *et al.* (2012) showed that this notion of comonotonicity appeared as more grounded in economic theory than alternative proposals because it is directly related to Pareto efficiency.

This article aims at finding multivariate couplings which exhibit a form of strong dependence, just as the previously defined comonotonic couplings. In what follows, the couplings defined as 'extreme' are comonotonic couplings (in the sense of the c-comonotonicity) *up to a linear transform* of one marginal (the c-comonotonic coupling corresponds to the identity transform). In other words, an extreme coupling (X, Y) satisfies the variational problem

$$\mathbf{E}(X'MY) = \sup_{\pi \in \Pi(P,Q)} \mathbf{E}_\pi(X'MY). \tag{3}$$

This definition of extreme dependence is broad enough to encompass 'positive dependence' as c-comonotonicity as well as 'negative dependence' (counter-comonotonicity in the univariate case). Furthermore, it allows for a geometrical interpretation of extreme dependence: the cross-covariance matrix of an extreme coupling is located on the boundary of the compact and convex set of all possible cross-covariance

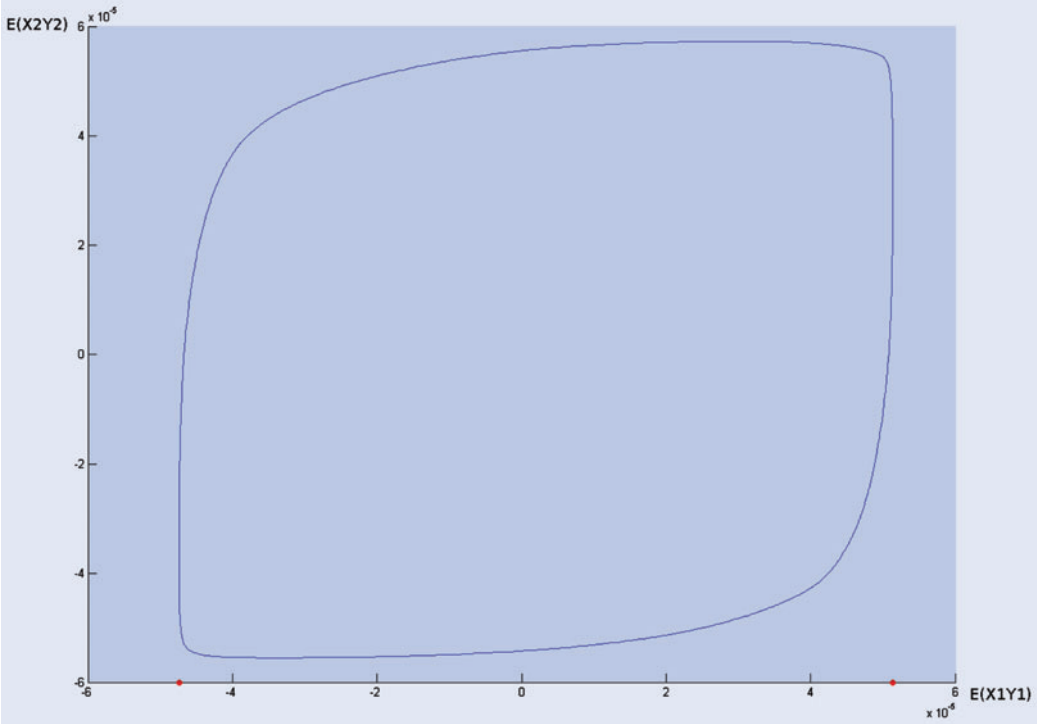


Figure 1. Example of a two-dimensional section of a covariance set.

matrices, called the covariance set. This set has been introduced in Galichon and Salanié (2010) in the case with discrete marginals, and generalized to the case with continuous marginals in Dupuy and Galichon (2013). Taking advantage of this simple interpretation, we then characterize the couplings π which have a cross-covariance matrix σ_π that are maximal for some partial orders $\succ$.

Although the idea of generating extreme dependence by solving problem (3) arises naturally from the theory of optimal transport—and more generally in the theory of distributions with given marginals, see e.g. Tiit (1992), the computation of the covariance set has remained a difficult point until now. The rest of the article proposes a method to compute extreme couplings, and for any given coupling $\hat{\pi}$, proposes a means to build a continuous sequence of couplings π_T with π_0 being extreme and $\sigma_{\pi_1} = \sigma_{\hat{\pi}}$. This is done by an entropic penalization technique, which allows for fast numerical computations and yields a measure of strength of dependence.

3. Multivariate extreme dependence

In this section we detail the notion of multivariate extreme dependence we propose. Consider the covariance set, the set of cross-covariance matrices of couplings $\pi \in \Pi(P, Q)$:

Definition 3.1 The covariance set $\mathcal{F}(P, Q)$ is defined as:

$$\mathcal{F}(P, Q) = \{\Sigma \in \mathbf{M}_{I,J}(\mathbf{R}) : \exists \pi \in \Pi(P, Q), \Sigma_{ij} = \mathbf{E}_\pi(X_i Y_j), \text{ for all } i, j\}.$$

As $\Pi(P, Q)$ is a convex and compact set (see e.g. Villani (2003), pp. 49–50), the covariance set is also a convex compact subset of $\mathbf{M}_{I,J}(\mathbf{R})$.

Figure 1 gives an example of the two-dimensional section of a covariance set, meaning that only the diagonal elements of the cross-covariance matrix are represented. P and Q are discrete distributions on $\mathbf{R}^2$ with equally weighted atoms and we look at the two component-wise covariances $\mathbf{E}(X_1 Y_1)$, $\mathbf{E}(X_2 Y_2)$. The solid curve is the boundary of the covariance set: every coupling between P and Q has component-wise covariances located within the convex hull of this curve. The independence coupling projects on the point $(0, 0)$. The dots on the x-axis represent, respectively, the minimal and maximal covariances between X_1 and Y_1. These covariances would be attained in terms of copulas by the lower and upper Fréchet copulas. This motivates our definition of *extreme dependence couplings* as couplings whose cross-covariance matrices are on the boundary of the covariance set.

Definition 3.2 A coupling $(X, Y) \sim \pi \in \Pi(P, Q)$ has extreme dependence if and only if $\left(\mathbf{E}_\pi(X_i Y_j)\right)_{ij}$ lies on the boundary of the covariance set $\mathcal{F}(P, Q)$.

The cross-covariance matrix between X and Y, $\sigma_{X,Y}$, satisfies

$$Tr\left(M' \sigma_{X,Y}\right) = \mathbf{E}(X' M Y), \quad \text{for all } M \in \mathbf{M}_{I,J}(\mathbf{R}), \tag{4}$$

which allows to reformulate the notion of extreme dependence as follows:

THEOREM 3.3 *The following conditions are equivalent:*

(i) $(X, Y) \sim \pi \in \Pi(P, Q)$ *have extreme dependence;*
(ii) *there exists* $M \in \mathbf{M}_{I,J}(\mathbf{R}) \backslash \{0\}$ *such that*

$$Tr\left(M' \sigma_\pi\right) = \sup_{\tilde{\pi} \in \Pi(P,Q)} Tr\left(M' \sigma_{\tilde{\pi}}\right),$$

or equivalently

$$\mathbf{E}_\pi(X' M Y) = \sup_{\tilde{\pi} \in \Pi(P,Q)} \mathbf{E}_{\tilde{\pi}}(X' M Y); \tag{5}$$

(iii) *there exists* $M \in \mathbf{M}_{I,J}(\mathbf{R})\backslash\{0\}$ *and a convex function* u *on* $\mathbf{R}^I$ *such that* $MY \in \partial u\,(X)$ *holds almost surely.*

This theorem is a corollary of the characterization of optimal couplings proved in Rachev and Rüschendorf (1990) and Brenier (1991).

Thus, a coupling (X, Y) is extreme if and only if there exists a linear transform, namely a nontrivial matrix M, such that (X, MY) is a maximum correlation coupling. In dimension 1, the interpretation is obvious: two real random variables have extreme dependence iff there exists a scalar $m \neq 0$ and a nondecreasing function u such that $mY = u(X)$. According to the classic terminology, X and Y are said to be comonotonic if $m > 0$, and anti-comonotonic otherwise.

When $M = Id$ in (5), the optimal coupling is the optimal transport coupling for the quadratic cost (it solves problem (2)).

4. Positive extreme dependence

The aim of this section is to propose a generalization of the concept of Fréchet copula of upper dependence to the multivariate case. Copula theory fails to handle this problem. Indeed, if C_P and C_Q are two copulas, the first in dimension I (associated with distribution P) and the second in dimension J (associated with distribution Q), a natural candidate for a copula modeling positive extreme dependence would be $C_\pi(x, y) = \min(C_P(x_1, \ldots, x_I), C_Q(x_1, ..., x_J))$. But according to an 'impossibility theorem' due to Schweizer and Sklar (1983), C_π is a copula function if and only if C_P and C_Q are themselves upper Fréchet copulas. We thus depart from the copula approach and aim at characterizing positive extreme dependence through the cross-covariance matrix of X and Y. Starting from the observation that in the univariate case, the positive extreme dependence attains maximum covariance between X and Y over all the couplings of P and Q, we introduce a conic order on the cross-covariance matrices $\sigma_{X,Y}$ and define positive extreme dependent couplings as the couplings whose cross-covariance matrices are maximal with respect to this order.

For a given compact convex set $B \subset \mathbf{M}_{I,J}(\mathbf{R})$ such that $0 \notin B$ (such a set is called a compact basis), a closed convex cone in $\mathbf{M}_{I,J}(\mathbf{R})$ is defined by setting:

$$K(B) = \{y \in \mathbf{M}_{I,J}(\mathbf{R}) | x \cdot y \geq 0,\ \forall x \in B\}. \tag{6}$$

Considering cones of this form might seem restrictive (Appendix A gives more details on such cones), yet we provide some examples that show that classic cones can be defined in such a manner.

Let M_1, M_2 be two matrices in $\mathbf{M}_{I,J}(\mathbf{R})$. A *strict* conic order on $\mathbf{M}_{I,J}(\mathbf{R})$ is defined by

$$M_1 \succ_{K(B)} M_2 \quad \text{if } M_1 - M_2 \in Int(K(B)),$$

where $Int(K(B)) = \{y \in \mathbf{M}_{I,J}(\mathbf{R}) | x \cdot y > 0,\ \forall x \in B\}$ is the interior of $K(B)$.

Definition 4.1 Let B be a compact basis. A coupling (X, Y) such that $\sigma_{X,Y}$ is a maximal element in $\mathcal{F}(P, Q)$ with respect to the strict conic order $\succ_{K(B)}$ is said to have *positive extreme dependence* with respect to $\succ_{K(B)}$.

The following results characterize couplings with positive extreme dependence in terms of maximal correlation couplings.

THEOREM 4.2 *The following conditions are equivalent:*

(i) $(X, Y) \sim \pi \in \Pi(P, Q)$ *have extreme positive dependence with respect to* $\succ_{K(B)}$

(ii) *there exists* $M \in B$ *such that*

$$Tr\left(M'\sigma_\pi\right) = \sup_{\tilde{\pi} \in \Pi(P,Q)} Tr\left(M'\sigma_{\tilde{\pi}}\right),$$

or equivalently

$$\mathbf{E}_\pi(X'MY) = \sup_{\tilde{\pi} \in \Pi(P,Q)} \mathbf{E}_{\tilde{\pi}}(X'MY); \tag{7}$$

(iii) *there exists* $M \in B$ *and a convex function* u *such that* $MY \in \partial u\,(X)$ *holds almost surely*

Hence, $\sigma_{X,Y}$ is maximal if and only if there exists $M \in B$ such that X and MY are maximally correlated. Obviously, this result is a close parallel to theorem 3.3 except that M is constrained to belong to B. As a consequence the positive extreme couplings are a particular case of extreme couplings. The interpretation in dimension 1 is again straightforward: X and Y have positive extreme dependence (w.r.t. the usual order on $\mathbf{R}$) iff they are comonotonic.

To better understand the relation between definitions 3.2 and 4.1, let us go back to the two-dimensional section of the covariance set discussed in the previous section, and consider that $K(B)$ is the positive orthant of $\mathbf{R}^2 \times \mathbf{R}^2$. The shaded region in figure 2 is the set of couplings dominating the coupling that projects on the square dot, with respect to the orthant order; as a consequence this coupling cannot have positive extreme dependence. This intuitively explains why maximal elements should be on the boundary of the covariance set, hence that positive extreme couplings should be extreme couplings. Maximal elements are represented on the bold curve in figure 3. Consequently, the couplings exhibiting positive extreme dependence project on this bold portion of the boundary of the covariance set. They form only a small part of the couplings of extreme dependence.

To demonstrate the applicability of this approach, here are three examples of partial orders on covariance matrices.

Example 4.3 Orthant order. Let $M^+_{I,J}(\mathbf{R})$ (resp. $M^{++}_{I,J}(\mathbf{R})$) denotes the set of real $I \times J$ matrices with nonnegative coefficients (resp. positive coefficients). The set $B = M^+_{I,J}(\mathbf{R}) \cap \{M | \sum_{i,j} M_{i,j} = 1\}$ is a compact basis. $K(B)$ is easily seen to be the set $M^+_{I,J}(\mathbf{R})$ and its interior is $M^{++}_{I,J}(\mathbf{R})$. Eventually $M_1 \succ M_2$ iff $M_1 - M_2$ has positive coefficients: this is the (strict) *orthant order* on matrices.

Example 4.4 Loewner order. Let S_n be the set of symmetric matrices of size $n \times n$ and let S^+_n and S^{++}_n denote, respectively, the set of nonnegative matrices in S_n and the set of definite positive matrices in S_n. If $B = \{S \in S^+_n(\mathbf{R}) | Tr(S) = 1\}$ is the set of semi-definite matrices with unit trace, B is a convex compact subset of $\mathbf{M}_n(\mathbf{R})$ and $K(B) = \{M \in \mathbf{M}_n(\mathbf{R}) | Tr(M'S) \geq 0, \forall S \in B\}$ is the set of matrices M whose symmetric part, $\frac{M+M'}{2}$, is semi-definite positive. The strict order $\succ_{K(B)}$ is then defined as: $M_1 \succ M_2$ iff the symmetric part of $M_1 - M_2$ is definite positive. This is an extension to $\mathbf{M}_n(\mathbf{R})$ of the classic *Loewner order* on symmetric matrices.

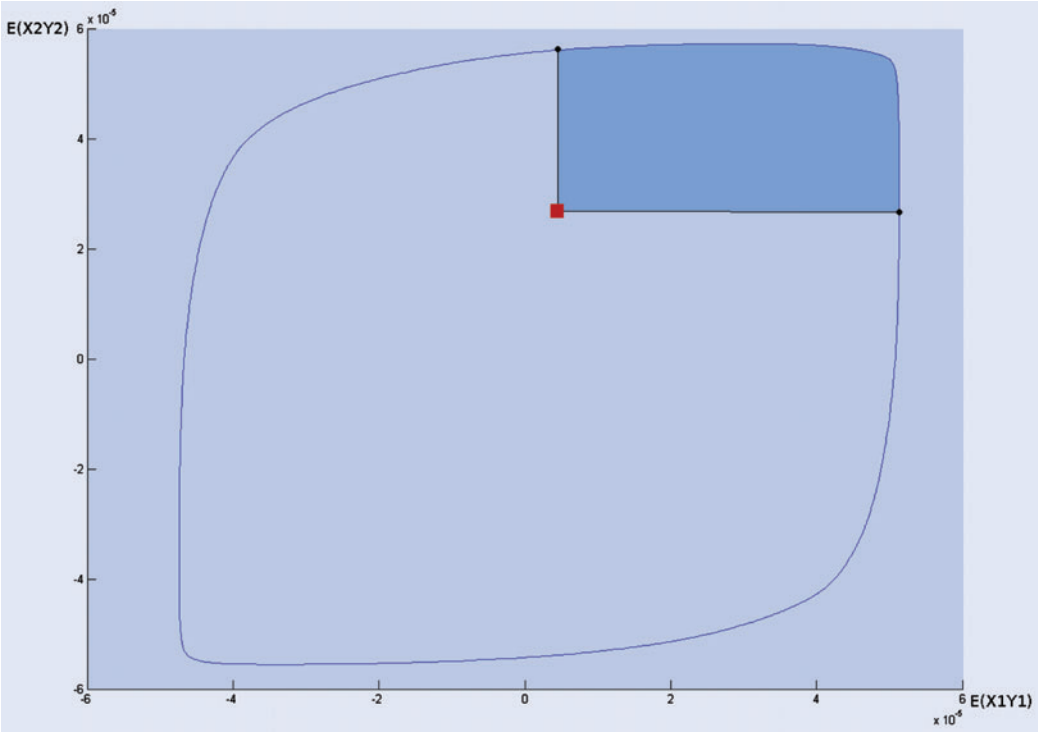


Figure 2. Shaded region: location of the couplings dominating the coupling materialized by the square dot.

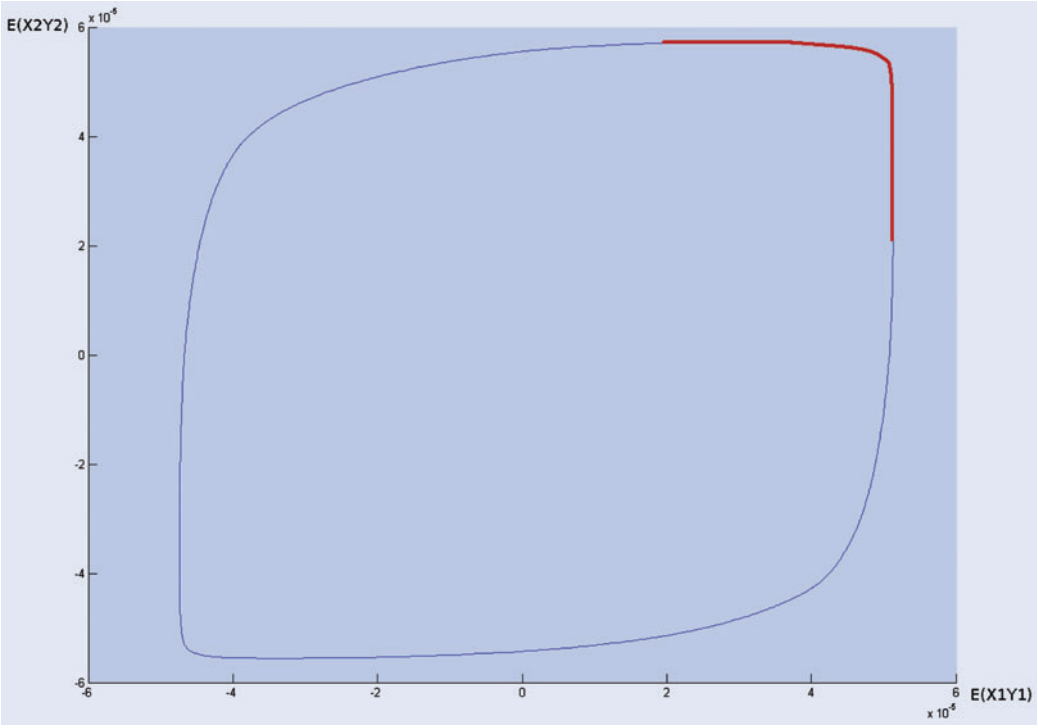


Figure 3. Maximal couplings on the boundary.

The following trivial example shows that the ordering induced by example 4.4 allows various positive extreme couplings. A first remark is that the maximum correlation coupling (the coupling solving (2)) is indeed positive extreme, by setting $M = Id$ in theorem 4.2. Consider $P \sim \mathcal{N}(0, I_2)$, the bivariate normal distribution, and $Q = \mathcal{N}(0, 1) \otimes \mathcal{U}_{(0,1)}$, the distribution of a vector whose first component is normal and whose second is the uniform distribution on $(0, 1)$, independent from the first component. Let $X \sim P$ and $Y = (X_1, U)'$, $U \sim \mathcal{U}_{(0,1)}$ be independent from (X_1, X_2), so that $Y \sim Q$. This coupling

does not have the maximum correlation even though $X_1 = Y_1$. However, it satisfies (7) with $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and qualifies as a maximal coupling.

5. An index of dependence

Suppose now we are observing (or simulating) a coupling $\hat{\pi} \in \Pi(P, Q)$, thereafter referred to as an empirical coupling. Even if this coupling is supposed to exhibit strong dependence, its cross-covariance matrix will never be exactly located on the boundary of the covariance set. Our problem is then to *associate an extreme coupling with* $\hat{\pi}$; more precisely, we propose to find a continuous sequence of nondeterministic couplings π_T such that $\pi_1 = \hat{\pi}$ and π_0 is an extreme coupling. In other words, we give a means to go smoothly from an empirical coupling to an extreme one by progressively increasing the strength of the dependence between the marginals. This is done by introducing an entropic penalization of (5), so that solutions of the penalized problem project on inner points of the covariance set.

5.1. Entropic relaxation

Consider the following problem, which is the entropic penalization of (5):

$$W(M, T) := \max_{\pi \in \Pi(P,Q)} \left(\mathbf{E}_\pi (X' M Y) + T\, Ent(\pi) \right), \qquad (8)$$

where the entropy of a coupling π is defined as

$$Ent(\pi) = \begin{cases} -\int \log \pi(x, y) d\pi(x, y), & \text{if } \pi \ll dx \otimes dy \\ \qquad \text{and the integral exists and is finite,} \\ -\infty \quad \text{otherwise.} \end{cases}$$

This problem has been considered by Kosowsky and Yuille (1994) as a simulated annealing method to solve the assignment problem, and in econometrics by Galichon and Salanié (2010) and Dupuy and Galichon (2013).

The parameter T can be thought of as a 'temperature' parameter which controls the strength of the entropic penalization. The problem (5) corresponds to $T = 0$, while letting T to $+\infty$ amounts to maximizing the entropy, in which case the solution of problem (8) is the independence coupling.

Let $\pi_{M,T}$ denote a solution of (8); a proof of its existence can be found in Rüschendorf (1995) and references therein. We assume furthermore that the entropy of $\hat{\pi}$ is finite.

Fixing the temperature at 1, our aim in the first place is to find a matrix M such that $\hat{\pi}$ and $\pi_{M,1}$ have the same cross-covariance matrix: $\sigma_{\hat{\pi}} = \sigma_{\pi_{M,1}}$. By a property of the subdifferential of a maximum function, the gradient of W with respect to M is: $\nabla_M W(\cdot, 1) = \sigma_{\pi_{M,1}}$. This implies that M is the solution of the following variational problem

$$\min_{M \in M_{I,J}(\mathbf{R})} W(M, 1) - \sigma_{\hat{\pi}} \cdot M. \qquad (9)$$

$W(\cdot, 1)$ is a convex function as a supremum of affine functions in M, and consequently the objective function in (9) is convex as well: this is a classic unconstrained convex minimization problem. Moreover, (9) is bounded below, which yields the existence of a global minimizer. A detailed proof and a discussion of uniqueness in (9) is given in Appendix C.

Figure 4 shows the diagonal of $\sigma_{\pi_{M,1}}$ in the coordinates $(\mathbf{E}(X_1 Y_1), \mathbf{E}(X_2 Y_2))$ for a large number of randomly generated matrices M. This graph is obtained by sampling many matrices M with coefficients uniformly distributed in the interval $[-1, 1]$, and then solving for each simulated M the problem (8), in order to find $\sigma_{\pi(M,T)}$. T is taken small enough to obtain near from extreme couplings. The solution of (8) is computed thanks to the algorithm presented in section 5.2. The bullet point has coordinates $(\mathbf{E}_{\hat{\pi}}(X_1 Y_1), \mathbf{E}_{\hat{\pi}}(X_2 Y_2))$. One sees that any inner point of the covariance set can be attained by a properly chosen π_M. This is a noticeable advantage of the entropic relaxation: not only are the optimal couplings solving (8) easily computed (at least when the marginals are discrete, see section 5.2), but also changing the temperature parameter allows one to reach any cross-covariance matrix inside the covariance set.

5.2. Numerical solution

The optimal $\pi_{M,1}$ in (8) obeys the following equation (see e.g. Rüschendorf (1995) for a proof):

$$\log \pi_{M,1}(x, y) = x' M y + u(x) + v(y),$$

for some $u \in L^1(dP)$, $v \in L^1(dQ)$. In other words, the log-likelihood of $\pi_{M,1}$ is the sum of a quadratic term $x' M y$ and of an additively separable function in x and y. The solution is found by setting u and v such that $\pi_{M,1}$ has the marginals P and Q. This is the purpose of the Iterative Projection Fitting Algorithm (Deming and Stephan 1940).

This algorithm consists in building a sequence π_n such that π_{2n} has first marginal P and π_{2n+1} has second marginal Q. It turns out that π_n converges towards a probability π with correct marginals P and Q. When the marginals P and Q are discrete distributions with atoms $P(x)$ and $Q(y)$, respectively, the algorithm is straightforward, as it consists in solving a series of linear systems:

$$\begin{cases} e^{v_{n+1}(y)} = \dfrac{Q(y)}{\sum_x r(x, y) e^{u_n(x)}}, \\ e^{u_{n+1}(y)} = \dfrac{P(x)}{\sum_y r(x, y) e^{v_{n+1}(y)}}, \end{cases}$$

where $r(x, y) = \frac{e^{x' M y}}{\sum_{x,y} e^{x' M y}}$.

The convex unconstrained minimization problem (9) can be solved by a Quasi-Newton algorithm (we used the BFGS method in the examples below). Of course, this algorithm can be used for any temperature T by replacing M by $\frac{M}{T}$ in the previous equations.

5.3. Derivation of the extreme coupling

We recall that our aim is to associate an inner coupling (i.e. a coupling whose cross-covariance matrix is inside the covariance set) to some extreme coupling which projects onto the boundary of the covariance set, by finding a trajectory of couplings that goes smoothly from the inner one to the extremal one.

The previous algorithm yields a particular matrix $\hat{M}$ and a coupling $\pi_{\hat{M}}$ such that $\sigma_{\hat{\pi}} = \sigma_{\pi_{\hat{M},1}}$. This coupling was found by setting arbitrarily the temperature at 1; the entropy penalization

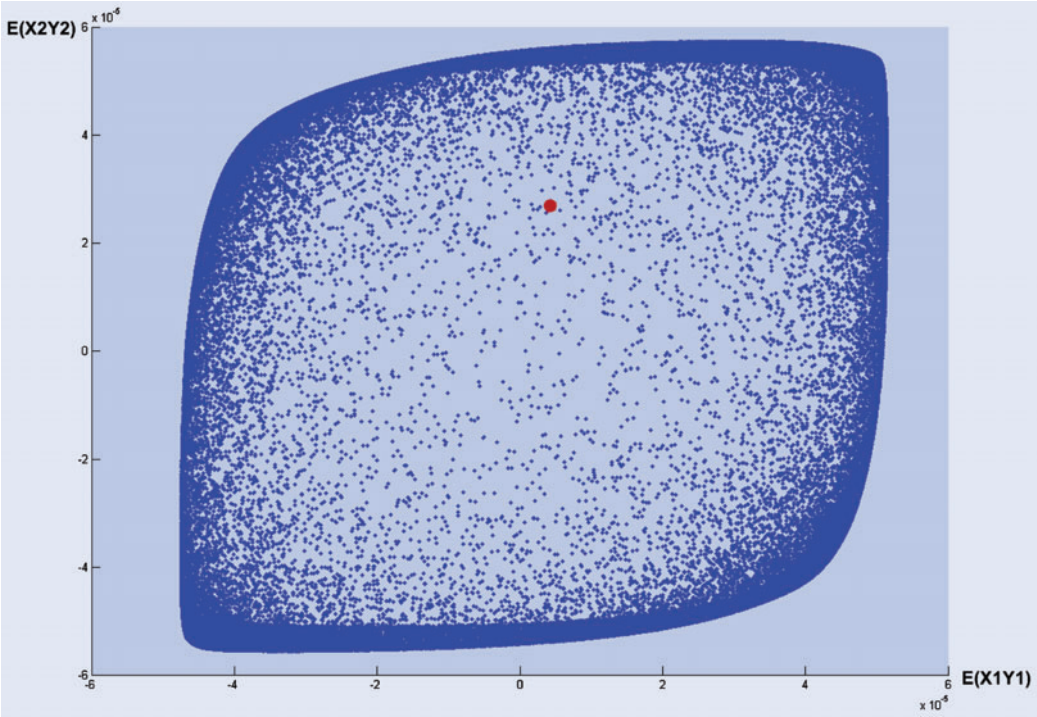


Figure 4. Projection of various π_M.

was thus effective and this allowed to reach inner points in the covariance set. This temperature parameter is easily explained. When it goes to $+\infty$, the entropy penalization is predominant in (8). Intuitively, the solution is the coupling with maximal 'disorder': this is the independence coupling. On the contrary, the lower the temperature, the closer (8) is to the nonpenalized problem. Hence, the lower T, the more $\pi_{\hat{M},T}$ projects near the boundary of the covariance set. Hence, associating $\hat{\pi}$ with an extreme coupling can be done in the following way: once $\hat{M}$ is found by solving (9), a sequence of $\pi_{\hat{M},T_n}$, $T_n \downarrow 0$ yields a trajectory of cross-covariance matrices which tends to the boundary.

Figure 5 summarizes this idea: each point on the curve is the projection of some $\pi_{\hat{M},T_n}$. As $T \to +\infty$, we recover the independence coupling whose projection is located at (0,0). When the temperature decreases, the trajectory passes on $\hat{\pi}$ at $T = 1$, and gradually approaches the boundary of the covariance set. The entropy is decreasing along this trajectory, as $Ent(\pi_T)$ decreases as $T \downarrow 0$ (thanks to the convexity of $W(M, T)$ in T), and thus lowering the temperature corresponds to going away from the independence coupling (maximal entropy). Thus, the temperature can be seen as a means to control the strength of the dependence. The matrix $\hat{M}$ can be seen as an *affinity matrix* (see Dupuy and Galichon (2013)): in the limit of $T \to 0$, the extreme coupling $\pi_{\hat{M},0}$ achieves the supremum of $\mathbf{E}_\pi(X'\hat{M}Y)$. Thus, $\hat{M}$ is the linear transform that makes X the most dependent with $\hat{M}Y$ under $\pi_{\hat{M},0}$.

This can be used to define formally an *index of dependence*, (if $\hat{\pi}$ different from the independence coupling): choosing a norm $||\cdot||$ over the set of matrices $M_{I,J}(\mathbf{R})$ and using the homogeneity of W, namely $W(\lambda M, \lambda T) = \lambda W(M, T)$ for all $\lambda \in \mathbf{R}$, we have $\pi_{\hat{M},1} = \pi_{\hat{M}/||\hat{M}||,1/||\hat{M}||}$ and the temperature $\frac{1}{||\hat{M}||}$ appears as an indicator of the strength of the dependence between the marginals of $\hat{\pi}$.

6. Applications

In the following financial applications below, we use the technique described in the previous section with times series of linear daily returns on industrial sectors of mainstream indices: S&P 500 and DJ Eurostoxx. We consider Health Care, Financial and Food & Beverage sectors of these indices: P and Q are distributions on $\mathbf{R}^3$. The historical data span 5 years between September 2004 and September 2009. Table 1 gives summary statistics (the first three variables correspond to S&P sectors, and the last three to Eurostoxx). In particular, the correlations between sectors belonging to different indices are mild ($< 35\%$). Inside an index, correlation is well higher, but remains below 80%; this motivated our choice for these sectors: the marginal distributions are not degenerated.

6.1. *Numerical results*

P and Q are discrete distributions with equally weighted atoms in $\mathbf{R}^3$, each atoms being a vector of the returns at some date of the three sectors.

$$P = \frac{1}{N}\sum_{t=1}^{N} \delta_{r^{X_t}},$$

$r_t^X =$ vector of the linear returns at date t on the three sectors of the S&P500.

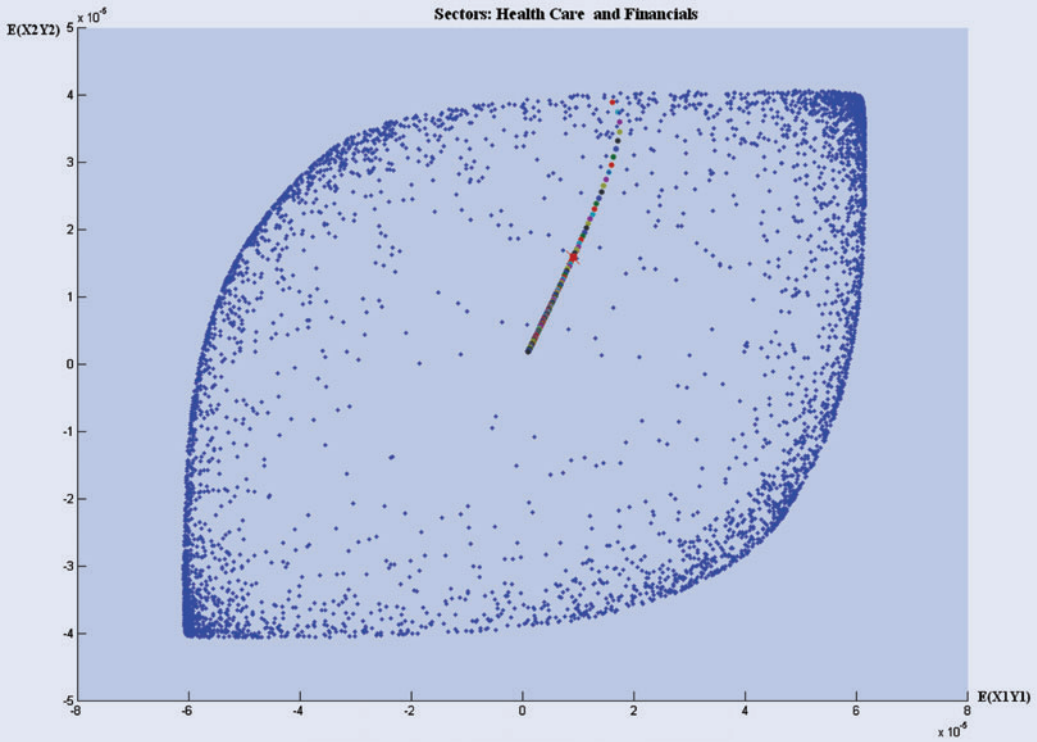


Figure 5. A trajectory toward an extreme coupling when the sectors are Health Care and Financials.

Table 1. Summary statistics.

Mean returns	$10^{-4}.(1.03\ -1.13\ 1.67\ 1.16\ -1.37\ 3.99)$
Variance	$10^{-4}.(1.36\ 7.65\ 1.16\ 1.14\ 4.15\ 1.12)$
Correlation matrix	$\begin{pmatrix} 1 & & & & & \\ 0.66 & 1 & & & & \\ 0.76 & 0.62 & 1 & & & \\ 0.22 & 0.10 & 0.19 & 1 & & \\ 0.26 & 0.33 & 0.25 & 0.49 & 1 & \\ 0.22 & 0.16 & 0.22 & 0.67 & 0.58 & 1.00 \end{pmatrix}$
Cross-covariance	$10^{-5}.\begin{pmatrix} 2.74 & 3.05 & 2.13 \\ 6.04 & 1.8 & 5.52 \\ 2.66 & 4.62 & 2.56 \end{pmatrix}$

The optimal $\hat{M}$ for these three sectors are:

# of components	2	3
Optimal M	$\begin{pmatrix} 0.23 & -0.14 \\ -0.10 & 0.40 \end{pmatrix}$	$\begin{pmatrix} 0.25 & -0.139 & -0.37 \\ -0.39 & 0.44 & -0.80 \\ -0.57 & -0.15 & 0.86 \end{pmatrix}$
Error = $\frac{\|\|\sigma_M - \sigma_{\hat{\pi}}\|\|}{\|\|\sigma_{\hat{\pi}}\|\|}$	$\approx 0.1\%$	$< 0.2\%$

The linear returns are expressed in percentage. The error is computed as the percentage of difference between $\sigma_{\hat{\pi}}$, the cross-covariance targeted, and $\sigma_{\pi_{M,1}}$, the covariance matrix of the optimal coupling. They should be perfectly equal in theory and this percentage measures the convergence of the gradient algorithm.

6.2. *Financial applications*

First, we use the trajectory of couplings $T \mapsto \pi_{\hat{M},T}$ as a continuous family of scenarios of increasing dependence. Theses scenarios are used to build scenarios of stress-tests involving multivariate variables, with obvious applications to risk management. By stress-testing, we mean increasing the index of dependence defined above (that is, lowering the temperature parameter), thus shifting away continuously from some coupling $\hat{\pi}$ to the extreme coupling $\pi_{\hat{M},0}$. This is to be compared to the method that consists in selecting the maximum correlation coupling as the 'strongest dependence scenario'; indeed this coupling might be less in line with the cross-covariance structure of the empirical coupling $\hat{\pi}$, yielding unexpected and undesired results when managing risky portfolios or options on several assets.

Then, we exploit further the affinity matrix $\hat{M}$ in order to exhibit indices of maximal correlation, based on an analysis of its singular value decomposition.

6.2.1. Portfolios stress-testing. In order to underline the necessity of accounting properly for the multivariate dependence, the problem of one-period portfolio allocation is considered. Suppose an investor chooses to allocate his wealth between assets $X_1, \ldots, X_n, Y_1, \ldots, Y_m$. The problem is to study the impact of the change of the dependence between $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_m)$ on the investor's portfolio.

In the numerical examples below, the assets are S&P Sector Indices: X is composed of Materials, Construction and Retail indices, while Y is composed of Food and Beverage, Health Care, Financials and Utilities indices. The corresponding summary statistics are given in table 2. Correlation is higher than in the above examples as the sectors are industrial sectors on a single index, the S&P500.

Table 2. Summary statistics.

Mean returns	$10^{-4}.\,(2.89\ 1.67\ 1.03\ -1.13\ 1.97\ 2.01\ 1.85)$
Variance	$10^{-4}.\,(3.59\ 1.16\ 1.36\ 7.65\ 1.92\ 0.984\ 3.25)$
Correlation matrix	$\begin{pmatrix} 1 & & & & & & \\ 0.72 & 1 & & & & & \\ 0.71 & 0.76 & 1 & & & & \\ 0.69 & 0.86 & 0.65 & 1 & & & \\ 0.69 & 0.85 & 0.69 & 0.76 & 1 & & \\ 0.69 & 0.67 & 0.75 & 0.62 & 0.66 & 1 & \\ 0.70 & 0.76 & 0.60 & 0.72 & 0.74 & 0.56 & 1 \end{pmatrix}$
Cross-covariance	$10^{-4}.\begin{pmatrix} 1.41 & 1.53 & 3.62 & 1.85 \\ 0.921 & 0.979 & 1.83 & 1.05 \\ 1.27 & 1.45 & 3.73 & 1.50 \end{pmatrix}$

It is assumed that the investor chooses his portfolio allocation according to the Markowitz allocation problem (over a one-year horizon), meaning that the weights ω that determine the allocation are chosen by solving the problem $\max_{\sum_i \omega_i = 1} \mu \cdot \omega - \frac{\lambda}{2}\omega' \Sigma \omega$. μ are the expected yearly returns of the stocks, Σ the covariance matrix of the returns, and λ a risk aversion parameter specific to the investor. We assume that both μ and Σ are the standard empirical estimators computed over a period of one year, the in-sample period. The risk aversion parameter λ is set at 3. The solution to the Markowitz allocation problem with these parameters is denoted by w. The risk of a portfolio is here identified by its variance, and is known as soon as the covariance between the assets is specified. When performing the allocation at time 0, the investor is expecting a risk of $\omega' \Sigma w$. The dependence stress-test consists in considering that the market conditions change after the investment decision: the strength of dependence between X and Y increases.

The affinity matrix is computed with respect to the in-sample data and the whole trajectory of couplings of increasing dependence is obtained, parameterized by the temperature T. These couplings π_T yield stressed covariance matrices $\Sigma_T = \mathbf{E}_{\pi_T}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y))')$. Σ_T represents a scenario where the marginals of X and Y are left unchanged, while the realized dependence between X and Y has increased, compared to the initial covariance matrix Σ.

The unexpected risks the investor might face when the dependence varies are materialized in the variance $w' \Sigma_T w$, plotted in figure 6. The variance obtained at temperature 1 is $w' \Sigma w$; in the worst case (which corresponds to temperature 0.1 in figure 6), the investor chooses a portfolio that has a variance 4% higher than expected.

When the dependence is properly accounted for, the investor determines the optimal weights w_T according to the covariance Σ_T. The opportunity cost $\mu \cdot w_T - \mu \cdot w$ is the loss on the return when the dependence increases while the investor sticks to the initial allocation w. This cost is more and more significant as the temperature lowers, reaching 6% in this case, as shown in figure 7.

A comparison with the maximum correlation coupling is enlightening. First of all, this coupling is not defined when the dimension of X and Y are different. Consequently, an asset is removed from Y (namely the Food and Beverages index) and the same computations as above are performed: a covariance matrix Σ_B that would be the realized covariance if the assets were in maximum correlation dependence is computed. On this particular example, the variance $w' \Sigma_B w$ is 60% *lower* than the expected variance $w' \Sigma w$. Other examples can yield a significantly higher covariance. This shows that the maximum correlation coupling might not be always adapted as a means of stress-testing the dependence.

A more classical way to stress the dependence is to suppose that the correlation between X_i and Y_j is fixed and equal to some parameter ρ for all i and j; the resulting cross-covariance matrix is denoted by Σ_ρ. A problem of this method is that it is known beforehand that, depending on the marginals, Σ_ρ might not be an admissible cross-covariance matrix for P and Q; the resulting variance-covariance matrix of the vector (X, Y) might fail to be semi-definite positive. This stress-test yields in this case underestimated risks. Indeed, while in our framework the variance $w' \Sigma w$ is at 1.91, this level of variance is attained only when ρ is above 95%, while the mean of the empirical cross-correlation is around 60%. Furthermore, even if ρ is set at 100 (disregarding the admissibility problem evoked above), the resulting variance is still lower than the one obtained with the extreme coupling.

It appears that the trajectory $T \mapsto \pi_T$ provides a coherent sequence of covariance matrices Σ_T that models an increase of the dependence between X and Y. This method respects both marginals and has the advantage of generating admissible matrices contrary to the usual method of parameterizing correlation matrices by a single parameter. Moreover, the maximum correlation coupling fails in this setting to properly account for increasing the risk of dependence, likely because it ignores the cross-correlation effects.

6.2.2. Options pricing. These couplings with increasing strength of dependence can be also used for the risk management and pricing of rainbow options (options on several underlyings). As a case study, consider the underlyings $X_1, \ldots, X_n$, $Y_1, \ldots, Y_m$. It is assumed that each one follows a log-normal martingale diffusion (i.e. we assume a null risk-free rate and write the risk-neutral dynamics):

$$\begin{cases} \dfrac{dX_t^i}{X_t^i} = \sigma_i^X dW_t^i, & d\langle W^i, W^j \rangle_t = \rho_{ij}^X dt,\ X_0^i = 1, \\ \dfrac{dY_t^i}{Y_t^i} = \sigma_i^Y dB_t^i, & d\langle B^i, B^j \rangle_t = \rho_{ij}^Y dt, \quad Y_0^i = 1. \end{cases}$$

The model is fully specified as soon as the correlation matrix between W and B is set. Consider the option that pays $\Phi = \min((\max_i X_T^i - K)_+, (\max_j Y_T^j - K)_+)$; it is the minimum between the payoffs of two best-of options on the X^i on the one hand and the Y^j on the other hand. It pays when the X_T^i and Y_T^i perform well, but mitigates the gain by selecting the lowest payoff between $(\max_i X_T^i - K)_+$ and $(\max_j Y_T^j - K)_+$.

Suppose an investor has sold this option and knows the distribution of the vector X and Y. In other words, he has been able to calibrate the volatilities σ_i^X and σ_i^Y, as well as the correlation matrices of $(W^1, \ldots, W^n)$ and of $(B^1, \ldots, B^m)$. The investor may have a guess on the dependence between X and Y (or equivalently between B and W), for instance an empirical estimation of the covariance matrix, but this guess is not sufficient to price the claim Φ in a conservative manner. A way to do this is to compute the price of this claim when the strength of the dependence between X and Y varies from the independence coupling to some extreme coupling and pick the highest value for the claim.

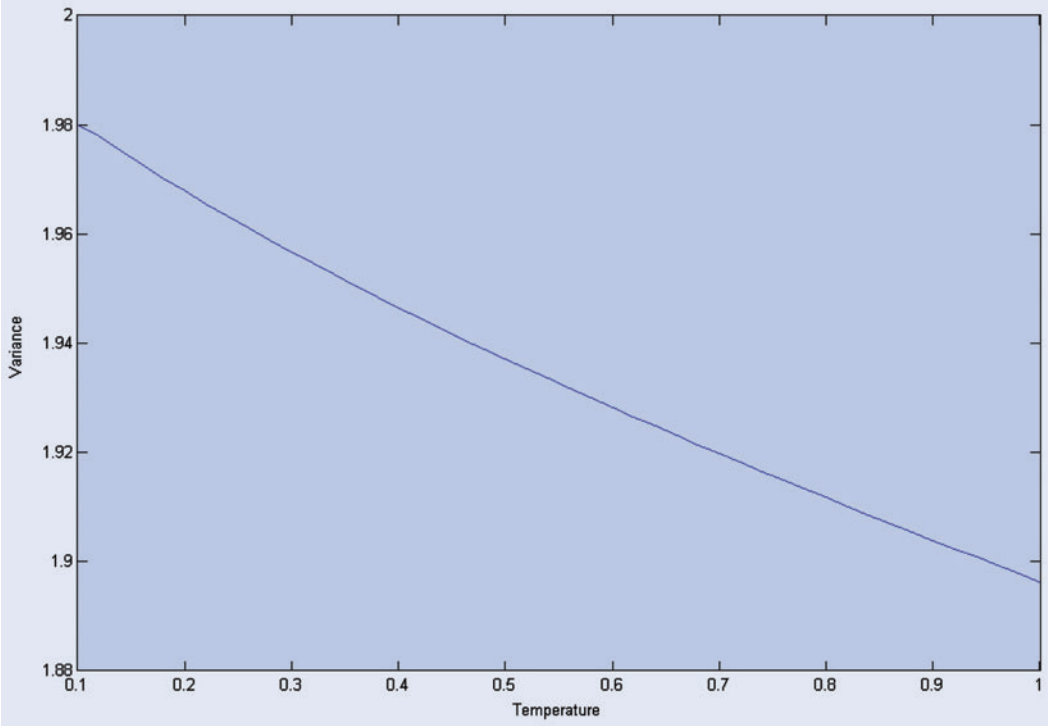


Figure 6. Plot of $T \mapsto w' \Sigma_T w$.

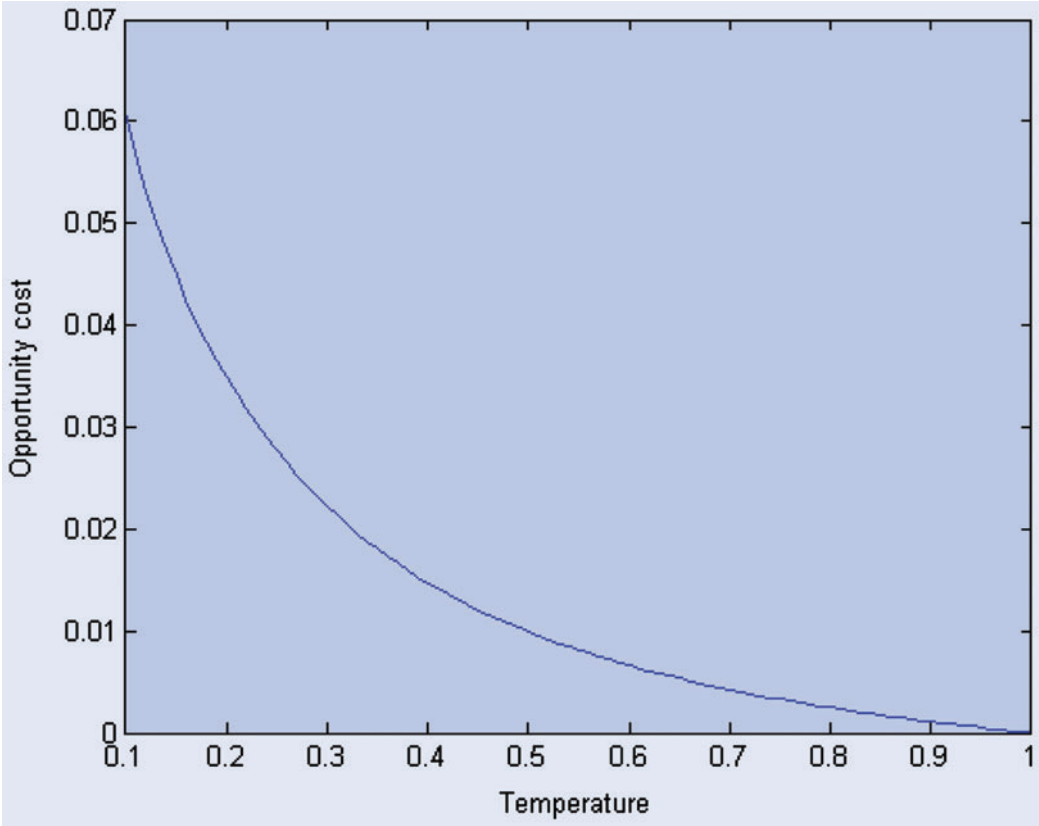


Figure 7. Opportunity cost as a function of the temperature.

For the purpose of numerical computations, the terminal distribution of the underlyings is discretized. The atoms of the discretized marginals are, respectively, denoted $x_T^{i,(k)}$ and $y_T^{j,(k')}$. For each specification of a cross-covariance matrix A

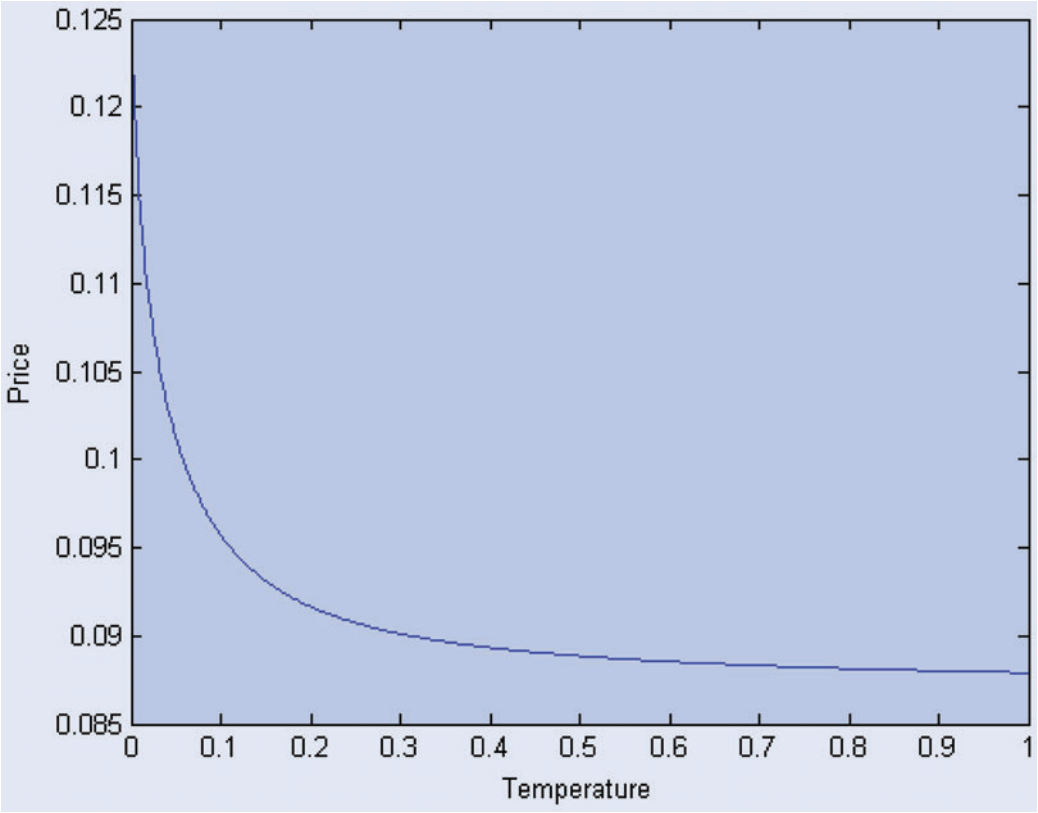


Figure 8. Price as a function of the temperature.

between X and Y, a trajectory $\pi_T(A)$ is obtained. The claim is priced as the expected value of Φ under the distribution $\pi_T(A)$:

$$P_T(A) = \mathbf{E}_{\pi_T(A)}\left(\min\left((\max_i X_T^i - K)_+, (\max_j Y_T^j - K)_+\right)\right)$$
$$= \sum_{k,k'} \min\left((\max_i x_T^{i,(k)} - K)_+, (\max_j y_T^{j,(k')} - K)_+\right)\pi_T(A)(x_T^{i,(k)}, y_T^{j,(k')}).$$

In the following example, X has 3 components and Y has 4, $\sigma^X = (0.15, 0.20, 0.22)'$ and $\sigma^Y = (0.13, 0.10, 0.16, 0.18)'$. For the sake of the exposition W and B are standard Brownian motions ($\rho^X = Id_n$ and $\rho^Y = Id_m$) while the cross-correlation matrix between W and B is randomly generated, and set at

$$\begin{pmatrix} 0.087 & 0.126 & 0.068 & 0.100 \\ 0.490 & 0.438 & 0.006 & 0.149 \\ 0.136 & 0.369 & 0.447 & 0.331 \end{pmatrix}.$$

The strike is set at 1, i.e. at time 0 the option is at-the-money.

As seen on figure 8, the price increases as the temperature lowers; this is an expected behavior, as when the dependence between the assets increases, so do the dependence between their respective maxima and thus the minimum of these maxima increases too, which yields a higher price. In this setting, the stress-test increases the price by more than 30% (i.e. between the price found with the independence coupling and the price found with the extreme coupling). This must be compared to the price that is obtained when the cross-correlation matrix Σ_ρ is taken of the form $\begin{pmatrix} \rho & \dots & \rho \\ \vdots & & \vdots \\ \rho & \dots & \rho \end{pmatrix}$. As a matter of fact, the stress-test of the cross-correlation fails, as the resulting correlation matrix $\begin{pmatrix} Id & \Sigma_\rho \\ \Sigma_\rho & Id \end{pmatrix}$ is no longer definite positive when $\rho > \frac{1}{2\sqrt{3}}$ which is lower than 30%. And even in the limit $\rho \to \frac{1}{2\sqrt{3}}$, the price does not reach 0.075, and is still lower than the nonstressed price.

6.2.3. Indices of maximal correlation. In order to better understand the link between the extreme coupling $\pi_{\hat{M},0}$ and the maximum correlation coupling (the one that corresponds to $M = Id$ in (5)), we use a singular value decomposition of the affinity matrix $\hat{M}$ of the coupling (X, Y). It writes $\hat{M} = USV'$, with U and V two orthogonal matrices and S a diagonal matrix with nonnegative entries. In particular,

$$\mathbf{E}_{\pi_{\hat{M},0}}\left((\sqrt{S}U'X)'(\sqrt{S}V'Y)\right) = \max_{\pi\in\Pi(P,Q)} \mathbf{E}_\pi\left((\sqrt{S}U'X)'(\sqrt{S}V'Y)\right).$$

In other words, if $(\tilde{X}, \tilde{Y}) = (\sqrt{S}U'X, \sqrt{S}V'Y)$, then this linear transform of (X, Y) has maximum covariance (under the distribution $\pi_{\hat{M},0}$).

Thus if $\tilde{P}$ is the distribution of $\sqrt{S}U'X$ with $X \sim P$, $\tilde{Q}$ is defined likewise from Q, and $\tilde{\pi}_{\hat{M},0}$ is the distribution of $(\sqrt{S}U'X, \sqrt{S}V'Y)$ where $(X, Y) \sim \pi_{\hat{M},0}$, then $\mathbf{E}_{\tilde{\pi}_{\hat{M},0}}(X'Y) = \max_{\pi\in\Pi(\tilde{P},\tilde{Q})} \mathbf{E}_\pi(X'Y)$. Therefore, the singular value decomposition of the affinity matrix provides linear transforms of the marginals that makes the extreme coupling $\pi_{\hat{M},0}$ the maximum correlation coupling after a scaling of the marginals by these transforms.

As an example, in the case of the 3 components described in the introduction of section 6, this transform writes

$$\tilde{X} = \begin{pmatrix} -0.42\ X_1 & +0.95\ X_2 & -0.019\ X_3 \\ -0.64\ X_1 & -0.27\ X_2 & +0.26\ X_3 \\ 0.11\ X_1 & +0.06\ X_2 & +0.35\ X_3 \end{pmatrix},$$

$$\tilde{Y} = \begin{pmatrix} -0.30\ Y_1 & +0.99\ Y_2 & -0.13\ Y_3 \\ -0.67\ Y_1 & -0.16\ Y_2 & +0.28\ Y_3 \\ 0.12\ Y_1 & +0.08\ Y_2 & +0.34\ Y_3 \end{pmatrix}.$$

This result states that $\tilde{X}$ and $\tilde{Y}$ are most correlated to one another under the distribution of the extreme coupling. These two vectors are composed of portfolios involving the components of the original index and can be viewed as new indices: we speak of *indices of maximal correlation*. When the strength of dependence is maximal ($T = 0$), they maximize the correlation $\mathbf{E}(\tilde{X}\tilde{Y})$ among all the couplings with same marginals.

This analysis can be seen as dual to the canonical correlation analysis, which consists, for two random vectors X and Y, in finding vectors a and b such that the correlation between $a'X$ and $b'Y$ is maximal. In the multivariate setting, $\sqrt{S}U'$ and $\sqrt{S}V'$ are the analogue of the optimal a and b. The technique described in this section has been introduced in the very different context of matching markets by Dupuy and Galichon (2013) under the name *saliency analysis*.

7. Conclusion

A recurring complaint in Applied Statistics is the 'curse of dimensionality': models that have a simple, computationally tractable form in dimension 1 become very complex, both computationally and conceptually in higher dimension. We show here that convex analysis, along with the theory of Optimal Transport, can lead to efficient solutions to the problem of extreme dependence. Building on a natural geometric definition of extreme dependence, we have introduced an index of dependence and used the latter to build stress-tests of dependence between two sets of economic variables. This is particularly relevant in the case of international finance, where the dependence between many economic variables in two countries is of interest.

Acknowledgements

The authors thank Rama Cont for a question which was the starting point of this article and Guillaume Carlier and Alexander Sokol for helpful conversation. Bosc acknowledges the support of the AXA Research Fund, AXA Investment Managers and the Investment Solutions Quantitative Team. Galichon's research in this area has been supported by FiME, Laboratoire de Finance des Marchés de l'Energie (www.fime-lab.org) .

References

Arrow, K., Uncertainty and the welfare of medical care. *Amer. Econ. Rev.*, 1963, **53**, 941–973.

Arrow, K., *Essays in the Theory of Risk-Bearing*, 1970 (North-Holland: Amsterdam).

Borch, K., Equilibrium in a reinsurance market. *Econometrica*, 1962, **30**, 424–444.

Brenier, Y., Polar factorization and monotone rearrangement of vector-valued functions. *Commun. Pure Appl. Math.*, 1991, **44**, 375–417.

Carlier, G., Dana, R.A. and Galichon, A., Pareto efficiency for the concave order and multivariate comonotonicity. *J. Econ. Theory*, 2012, **147**, 207–229.

Dall'Aglio, G., Sugli estremi dei momenti delle funzioni di ripartizione doppia. *Ann. Sc. Norm. Super. Pisa*, 1956, **10**, 35–74.

Deming, W. and Stephan, F., On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Stat.*, 1940, **11**, 427–444.

Dupuy, A. and Galichon, A., Personality traits and the marriage market, 2013. Available online at: http://ftp.iza.org/dp6943.pdf (accessed 27 October 2012).

Ekeland, I., Galichon, A. and Henry, M., Comonotonic measures of multivariate risks. *Math. Finance*, 2012, **22**, 109–132.

Fan, K., Fixed-Point and minimax theorems in locally convex topological linear spaces. *Proc. Nat. Acad. Sci.*, 1951, **38**, 121–126.

Fréchet, M., Sur les tabealux de corrélation dont les marges sont données. *Ann. Univ. Lyon. Sect. A*, 1951, **9**, 53–57.

Galichon, A. and Henry, M., Dual theory of choice under multivariate risks. *J. Econ. Theory*, 2012, **147**, 1501–1516.

Galichon, A. and Salanié, B., Matching with trade-offs: Revealed preferences over competing characteristics, 2010. Available online at: http://ssrn.com/abstract=1487307(accessed 22 March 2010).

Hoeffding, W., Masstabinvariante Korrelationstheorie. *Schr. Math. Inst. Univ. Berlin*, 1940, **5**, 179–233.

Kosowsky, J.J. and Yuille, A.L., The invisible hand algorithm: Solving the assignment problem with statistical physics. *Neural Networks*, 1994, **7**, 477–490.

Landsberger, M. and Meilijson, I., Co-monotone allocations, Bickel-Lehmann dispersion and the Arrow-Pratt measure of risk aversion. *Ann. Oper. Res.*, 1994, **52**, 97–106.

Malevergne, Y. and Sornette, D., *Extreme Financial Risks, From Dependence to Risk Management*, 2006 (Springer: Berlin).

Puccetti, G. and Scarsini, M., Multivariate comonotonicity. *J. Multivariate Anal.*, 2010, **101**, 291–304.

Rachev, S., *Probability Metrics and the Stability of Stochastic Models*, 1991 (Wiley: Hoboken, NJ).

Rachev, S.T. and Rüschendorf, L., A Characterization of Random Variables with Minimum L^2- distance. *J. Multivariate Anal.*, 1990, **32**, 48–54.

Rachev, S.T. and Rüschendorf, L., *Mass Transportation Problems. Volume I: Theory and Volume II: Applications*, 1998 (Springer: New York).

Rüschendorf, L., Fréchet-bounds and their applications. In *Advances in Probability Distributions with Given Marginals*, Vol. 67 of *Mathematics and its Applications*, edited by G. Dall'Aglio, S. Kotz, and G. Salinetti, 151–187, 1990 (Kluwer Academic: Dordrecht).

Rüschendorf, L., Convergence of the iterative proportional fitting procedure. *Ann. Stat.*, 1995, **23**, 1160–1174.

Schmeidler, D., Subjective probability and expected utility without additivity. *Econometrica*, 1989, **57**, 571–587.

Schweizer, B. and Sklar, A., *Probabilistic Metric Spaces*, 1983 (North-Holland: New York).

Tiit, E.M., Extremal multivariate distributions having given discrete marginals. *Acta Comment. Univ. Tartu. Math.*, 1992, **942**, 94–113.

Valadier, M., Sous-différentiels d'une borne supérieure et d'une somme continue de fonctions convexes. *C. R. Acad. Sci. Paris*, 1969, **268**, A39–A42.

Villani, C., *Topics in Optimal Transportation*, 2003 (American Mathematical Society: Providence, RI).

Yaari, M., The dual theory of choice under risk. *Econometrica*, 1987, **55**, 95–115.

Zolotarev, V., Probability metrics. *Theory Probab. Appl.*, 1983, **28**, 278–302.

Appendix A. Facts on conic orders

In the space $\mathbf{M}_{I,J}(\mathbf{R})$, a *basis* is a convex set B with $0 \notin \bar{B}$ (the closure of B). We assume that B is a compact basis. Let $K(B)$ be the *dual cone* of the cone generated by B, $\mathbf{R}_+.B = \{\lambda.b,\ \lambda \geq 0,\ b \in B\}$, which means that:

$$K(B) = \{\Sigma \in \mathbf{M}_{I,J}(\mathbf{R}) | \Sigma \cdot M \geq 0,\ M \in \mathbf{R}_+.B\}.$$

Its interior is

$$Int(K(B)) = \{\Sigma \in \mathbf{M}_{I,J}(\mathbf{R}) | \Sigma \cdot M > 0,\ M \in \mathbf{R}_+.B \backslash \{0\}\}.$$

It is important to note that in both definitions, $\mathbf{R}_+.B$ and $\mathbf{R}_+.B\backslash\{0\}$ can be replaced by the basis B.

A *strict partial order* is defined on $\mathbf{M}_{I,J}(\mathbf{R})$ by setting

$$M_1 \succ_K M_2 \overset{\text{def}}{\Leftrightarrow} M_1 - M_2 \in K_+^*.$$

If S is a subset of $\mathbf{M}_{I,J}(\mathbf{R})$, a *maximal* element of S for this order is a matrix $A \in S$ such that for all $B \in S$, $M_1 - M_2 \notin K_+^*$: M_1 cannot be 'strictly dominated' by any element in S.

The choice of $\mathbf{M}_{I,J}(\mathbf{R})$ is arbitrary here and it could be replaced by any Euclidean space.

Appendix B. Proof of the results

B.1. Proof of Theorem 3.3

As stated in section 3, this theorem is a direct corollary of the results proven in Rachev and Rüschendorf (1990) and Brenier (1991) which characterize optimal couplings.

Proof As the covariance set is a closed convex set, a point $x \in \mathbf{M}_{I,J}(\mathbf{R})$ lies on its boundary if and only if there exists a nonzero $M \in \mathbf{M}_{I,J}(\mathbf{R})\backslash\{0\}$ such that $M \cdot x$ is maximal as a function of x. This translates to the fact that there exists a supporting hyperplane at x. Thus, σ_π is on the boundary of the covariance set iff there exists $M \in \mathbf{M}_{I,J}(\mathbf{R})\backslash\{0\}$ such that

$$M \cdot \sigma_\pi = \sup_{\tilde{\pi} \in \Pi(P,Q)} M \cdot \sigma_{\tilde{\pi}},$$

(recall that $M \cdot \sigma_\pi = Tr\left(M'\sigma_\pi\right)$).

Equivalence between (ii) and (iii) follows from a well-known result in Optimal Transport theory, the Knott-Smith optimality criterion (see Villani (2003), Th. 2.12). □

B.2. Proof of theorem 4.2

Before we give the proof of the theorem, we state and prove a number of auxiliary results which are of interest per se. Let B be a compact basis ; we have a crucial, although technical, variational characterization of the maximality of σ_π with respect to $\succ_{K(B)}$:

PROPOSITION B.1 (Variational characterization of maximality)

$$\sigma_\pi \text{ maximal iff } \inf_{M \in B} \sup_{\tilde{\pi} \in \Pi(P,Q)} (\sigma_{\tilde{\pi}} - \sigma_\pi) \cdot M = 0.$$

In other words, a coupling is maximal whenever there exists $M \in B$ such that σ_π maximizes $\sigma_{\tilde{\pi}} \cdot M$.

Proof [Proof of proposition B.1] Note that for every $\pi \in \Pi(P, Q)$, the function

$$f : (\tilde{\pi}, M) \in \Pi(P, Q) \times B \mapsto (\Sigma_{\tilde{\pi}} - \Sigma_\pi) \cdot M,$$

exhibits a saddlepoint $(\bar{\pi}, \bar{S})$:

$$\max_{\tilde{\pi} \in \Pi(P,Q)} \min_{M \in B} f(\tilde{\pi}, M) = f(\bar{\pi}, \bar{M}) = \min_{M \in B} \max_{\tilde{\pi} \in \Pi(P,Q)} f(\tilde{\pi}, M). \tag{B1}$$

This is a consequence of a classical minmax theorem by Fan (1951): a continuous function over a product of compact convex sets embedded in normed linear spaces, and which is linear in both arguments, exhibits a saddlepoint. Both $\Pi(P, Q)$ and B are compact and convex. The compactness of B is an hypothesis, and it is a well-known fact that $\Pi(P, Q)$ is compact, see Villani (2003). Moreover f is linear in M and $\tilde{\pi}$, and continuous in both arguments. Finally, $\Pi(P, Q)$ can be embedded in the space of Radon measures over $\mathbf{R}^I \times \mathbf{R}^J$ endowed with the bounded Lipschitz norm. We refer to Villani (2003, chapter 7). for more details on this point: the important thing is that $\Pi(P, Q)$ is a compact subset (for this norm) of this space.

Back to the proof of the theorem. If σ_π is maximal, then for all $\sigma_{\tilde{\pi}}$ one has $\sigma_{\tilde{\pi}} - \sigma_\pi \notin Int(K(B))$, which means that for some $M \in B$, $(\sigma_{\tilde{\pi}} - \sigma_\pi) \cdot M \leq 0$, hence

$$\sup_{\tilde{\pi} \in \Pi(P,Q)} \inf_{M \in B} (\sigma_{\tilde{\pi}} - \sigma_\pi) \cdot M \leq 0.$$

Therefore, the above quantity is necessarily zero, because one may choose $\tilde{\pi} = \pi$. Thanks to the compactness of B and $\Pi(P, Q)$, the minmax theorem applies and yield that the infimum of the supremum is zero.

On the contrary, if σ_π is not maximal then there exists some coupling $\tilde{\pi}$ such that $\sigma_{\tilde{\pi}} - \sigma_\pi \in Int(K(B))$. Thus, for all $M \in B$, $\sup_{\pi \in \Pi(p,q)} \sigma_{\tilde{\pi}} - \sigma_\pi \cdot M > 0$, and thanks to the compactness of B,

$$\inf_{M \in B} \sup_{\tilde{\pi} \in \Pi(p,q)} \sigma_{\tilde{\pi}} - \sigma_\pi \cdot M > 0.$$

□

As a consequence, we are now ready to prove theorem 4.2.

Proof [Proof of theorem 4.2]

(ii) ⇒ (i): If for some $M \in B$, a coupling π satisfies

$$\mathbf{E}_\pi(X \cdot MY) = \sup_{\tilde{\pi} \in \Pi(P,Q)} \mathbf{E}_{\tilde{\pi}}(X \cdot MY),$$

then $\sup_{\tilde{\pi} \in \Pi(P,Q)} (\sigma_{\tilde{\pi}} - \sigma_\pi) \cdot M = 0$ and so $\inf_{M \in B} \sup_{\tilde{\pi} \in \Pi(P,Q)} (\sigma_{\tilde{\pi}} - \sigma_\pi) \cdot M \leq 0$. But this is an infimum of quantities that are greater than zero, and eventually the 'inf sup' is zero.

(i) ⇒ (ii): if σ_π is maximal, then proposition B.1 entails $\inf_{M \in B} \sup_{\tilde{\pi} \in \Pi(p,q)} \sigma_{\tilde{\pi}} - \sigma_\pi \cdot M = 0$. Due to the compactness of B, there exists a matrix $M \in B$, such that the supremum is zero, which concludes the proof of this implication. □

Appendix C. More details on problem (9)

The objective function of the problem (9) is convex in M, because it is the sum of: a linear function of M,$-\sigma_{\hat{\pi}} \cdot M$; and of $W(M, 1)$, which is convex in M as the supremum over $\pi \in \Pi(p, q)$ of linear functions in M, namely $\mathbf{E}_\pi(X'MY) = \sigma_\pi \cdot M$.

Moreover, assuming that the entropy of the empirical coupling $\hat{\pi}$ is finite, then $W(M, 1) \geq \sigma_{\hat{\pi}} \cdot M + \text{Ent}(\hat{\pi})$. Thus $W(M, 1) - \sigma_{\hat{\pi}} \cdot M \geq \text{Ent}(\hat{\pi}) > -\infty$. A strictly convex function which is bounded from below admits a unique global minimizer.

Moreover, the objective function is differentiable as $W(M, 1)$ is differentiable and $\nabla_M W(M, 1) = \sigma_{\pi(M,1)}$. This is a consequence of a property of subdifferentials, see e.g. Valadier (1969). A global minimizer is necessarily a critical point, proving that the solution M of problem (9) satisfies $\sigma_{\pi(M,1)} = \sigma_{\hat{\pi}}$.