

## *Document de recherche*

# Agrégation et diversification des risques

## Commission sur la gestion des risques et le capital requis

**Avril 2016**

Document 2160037

*This document is available in English*  
© 2016 Institut canadien des actuaires

Les documents de recherche ne représentent pas nécessairement l'opinion de l'Institut canadien des actuaires. Les membres devraient connaître les documents de recherche. Les documents de recherche ne constituent pas des normes de pratique et sont donc de caractère non exécutoire. Il n'est pas obligatoire que les documents de recherche soient conformes aux normes de pratique. Le mode d'application de normes de pratique dans un contexte particulier demeure la responsabilité des membres.

## NOTE DE SERVICE

- À :** Tous les Fellows, affiliés, associés et correspondants de l'Institut canadien des actuaires
- De :** Pierre Dionne, président  
Direction de la pratique actuarielle  
Leonard Pressey, président  
Commission sur la gestion des risques et le capital requis
- Date :** Le 14 avril 2016
- Objet :** **Document de recherche sur l'agrégation et la diversification des risques**
- 

La Commission sur la gestion des risques et le capital requis (CGRCR) de l'Institut canadien des actuaires (ICA) a parrainé un document de recherche sur l'agrégation et la diversification des risques. Le fonds documentaire sur ce sujet croît rapidement et il faut s'attendre qu'à l'avenir, les nouvelles approches en matière d'agrégation des risques et leurs répercussions sur le risque de modélisation feront l'objet de nombreuses études. La recherche portait plus particulièrement sur les éléments suivants :

- La recherche de documents sur les méthodes utilisées à l'échelle internationale pour mesurer et reconnaître les effets du jumelage de risques divers (p. ex., le risque d'assurance, le risque de crédit, le risque de marché, le risque opérationnel et d'autres risques). Les sources peuvent englober les documents de recherche universitaire, les démarches en matière de réglementation des exigences de capital axées sur les risques, les ouvrages d'actuariat et les méthodes du secteur des assurances.
- Bien que l'accent soit placé sur les sociétés d'assurance-vie et les sociétés d'assurances autres que l'assurance-vie, les démarches effectuées dans d'autres secteurs peuvent être pertinentes pour le secteur des assurances.

Les chercheurs ont amorcé leurs travaux par un examen des documents de recherche universitaire sur l'agrégation des risques, de même que sur les démarches en matière de réglementation axées sur les risques. En outre, ils se sont appuyés sur leurs propres publications et travaux de recherche sur l'agrégation et la diversification des risques pour envisager l'avenir des méthodes d'agrégation des risques. Ces examens ont ensuite été utilisés pour résumer les pratiques actuelles et faire ressortir les avantages et les inconvénients des diverses démarches d'agrégation des risques, et leurs limites, notamment l'exactitude théorique, les problèmes d'étalonnage et la mise en œuvre pratique.

Conformément à la Politique sur le processus officiel d’approbation de matériel d’orientation autre que les normes de pratique et les documents de recherche de l’ICA, le présent document de recherche a été préparé par la Commission sur la gestion des risques et le capital requis, et il a été approuvé par la Direction de la pratique actuarielle à des fins de diffusion le 5 avril 2016.

La CGRCR tient à souligner l’assistance fournie par la Section conjointe de la gestion des risques (Society of Actuaries (SOA), ICA et Casualty Actuarial Society (CAS)) et le personnel du siège social de l’ICA.

Pour toute question ou commentaire à propos du présent document de recherche, veuillez communiquer avec Leonard Pressey, président, Commission sur la gestion des risques et le capital requis à [leonard.pressey@empire.ca](mailto:leonard.pressey@empire.ca).

PD, LP

## Agrégation et diversification des risques

C. Bernard

S. Vanduffel

Le 23 février 2016

### Table des matières

Vue d'ensemble .....	6
Introduction .....	6
1. Mesure du risque agrégé .....	8
1.1 Mesures de risque cohérentes.....	10
1.2 Contrôle ex post et robustesse des mesures de risque.....	11
2. Agrégation et diversification .....	13
2.1 Avantages de la diversification et sous-additivité.....	13
2.2 L'erreur de n'utiliser que les corrélations .....	15
2.3 L'impact des microcorrélations.....	16
2.4 Ajustement d'une distribution multivariée.....	18
2.5 Résumé.....	19
3. Survol de la réglementation à l'heure actuelle .....	19
3.1 Dispositifs de réglementation .....	20
3.2 Comparaison des dispositifs de réglementation internationaux et commentaires.....	21
4. Risque de modélisation de la dépendance .....	23
5. Conclusions.....	26
Annexe A – Risque de modélisation de la dépendance au moment de l'agrégation des risques	27
A.1 Limites convexes supérieures et inférieures.....	28
A.2 Algorithme de réarrangement .....	29
A.3 Exemple d'application de l'algorithme de réarrangement.....	30
A.4 La variance et le risque de modélisation de la dépendance.....	31
A.5 La VaR et le risque de modélisation de la dépendance.....	32
Annexe B – Risque de modélisation de la dépendance lors de l'agrégation des risques et quantification du risque de modélisation .....	36

---

B.1	Fondements théoriques et hypothèses.....	38
B.2	Cadre pratique pour la méthode non paramétrique.....	40
B.3	Limites d'une mesure de risque donnée.....	43
B.3.1	<i>Limites théoriques de la variance et de la TVaR</i> .....	44
B.3.2	<i>Limites pratiques de l'écart-type et de la TVaR</i> .....	45
B.3.3	<i>Limites théoriques de la VaR</i> .....	47
B.3.4	<i>Limites pratiques de la VaR</i> .....	48
B.4	Exemples de limites des mesures de risque des portefeuilles avec modélisation par Monte Carlo de l'incertitude au sujet de la dépendance.....	51
B.4.1	<i>Limites de la variance</i> .....	51
B.4.2	<i>Limites de la TVaR</i> .....	54
B.4.3	<i>Limites de la VaR</i> .....	55
B.4.4	<i>Autres considérations relatives au risque de modélisation</i> .....	57
	Annexe C – Définitions des notations mathématiques.....	60
	Bibliographie.....	63

## Vue d'ensemble

Le présent rapport passe en revue la littérature théorique portant sur l'agrégation et la diversification des risques et sur les approches réglementaires. Nous montrerons les avantages et les inconvénients des différentes approches, en insistant sur le risque de modélisation.

Tout d'abord, à la section [1](#), nous exposerons les principes de base de la mesure du risque agrégé. Nous étudierons plus particulièrement l'idée d'une mesure du risque comme moyen adéquat pour mesurer le risque agrégé. Nous traiterons des propriétés souhaitées des mesures de risque et nous exemplifierons notre analyse par l'étude de la valeur à risque (VaR) et de la valeur à risque conditionnelle (TVaR).

À la section [2](#), nous explorerons la question des avantages de la diversification associée à l'agrégation des risques et les limites éventuelles des corrélations comme seule statistique de mesure de la dépendance. Nous pousserons plus loin et expliquerons l'importance de recourir à un modèle complet multivarié pour obtenir une description juste du niveau de risque agrégé.

Ensuite, à la section [3](#), nous explorerons l'approche que les organismes de réglementation emploient en matière d'agrégation et de diversification des risques et nous ferons quelques observations sur l'hypothèse implicite sur laquelle s'appuient les organismes de réglementation internationaux, et sur les différentes approches pouvant être adoptées.

Nous terminerons notre revue en insistant sur le fait que le risque de modélisation constitue un enjeu essentiel de la mesure de l'agrégation et de la diversification des risques. À la section [4](#), nous analyserons un cadre qui permet une quantification du risque de modélisation et qui a été élaboré récemment par [Bernard et Vanduffel \[2015a\]](#)<sup>1</sup>, qui se sont inspirés des idées d'[Embrechts et al. \[2013\]](#). De plus amples détails sont donnés aux annexes [A](#) et [B](#). Les définitions des notations mathématiques utilisées dans le présent document sont présentées à l'annexe [C](#).

## Introduction

Pour les institutions financières, l'évaluation des risques de portefeuilles multidimensionnels  $(X_1, X_2, \dots, X_d)$  est l'une des principales difficultés de la gestion des risques. Bien entendu, ce problème se pose plus particulièrement pour les sociétés d'assurances. Habituellement, l'assureur est exposé à divers facteurs de risque (p. ex., risque de dommages, risque de longévité, risque de crédit, risque de marché, risque opérationnel) et possède différentes branches d'assurance ou plusieurs portefeuilles de clients. Dans ce cas, on cherche habituellement à mesurer le risque d'une somme aléatoire,  $S = \sum_{i=1}^d X_i$ , dans laquelle les risques individuels  $X_i$  représentent les pertes (réclamations des différents clients, variations de la valeur marchande des différents facteurs de risque, etc.), au moyen d'une mesure de risque comme la variance, la VaR ou la TVaR<sup>2</sup>. À l'évidence, la solution de ce problème passe

<sup>1</sup> Cet ouvrage a reçu en 2014 le prix *New Frontiers in Risk Management* de la PRMIA.

<sup>2</sup> Dans la littérature, cette mesure porte aussi entre autres noms celui de valeur à risque conditionnelle, d'insuffisance attendue et de perte de queue.

essentiellement par une méthode numérique une fois que la distribution conjointe  $(X_1, X_2, \dots, X_d)$  est entièrement spécifiée. Malheureusement, l'estimation d'une distribution à plusieurs variables n'est pas une mince affaire. Dans de nombreux cas, l'actuaire est en mesure d'appliquer des techniques mathématiques et statistiques pour décrire fructueusement les risques marginaux  $X_i$ , mais la dépendance entre les risques n'est pas spécifiée, ou ne l'est que partiellement. En d'autres termes, l'évaluation du risque d'un portefeuille est sujette à une erreur de spécification du modèle (risque de modélisation).

D'un point de vue mathématique, il est donc souvent pratique de supposer que les variables aléatoires  $X_i$  sont indépendantes les unes des autres, ce qui permet d'appliquer de puissantes méthodes exactes de calcul comme la formule récursive de Panjer ou la technique de convolution. Dans ce cas, on peut aussi tirer avantage du théorème central limite, qui énonce que la somme des risques,  $S$ , suit approximativement une loi normale si le nombre de risques est assez grand. De fait, l'existence même de l'assurance repose sur l'hypothèse de l'indépendance mutuelle des risques assurés, et parfois, cela est plus ou moins conforme à la réalité. Toutefois, dans la majorité des cas, les différents risques sont liés entre eux dans une certaine mesure. Par exemple, on obtient une somme  $S$  de risques dépendants lorsqu'on considère le montant agrégé des réclamations d'un portefeuille d'assurance de dommages, parce que les risques assurés sont soumis à des facteurs communs comme la géographie, le climat ou le contexte économique. La fonction de distribution cumulative de  $S$  ne peut plus être spécifiée facilement.

Les méthodes qui sont habituellement employées pour estimer une distribution multivariée représentant des risques dépendants utilisent une distribution gaussienne ou une loi  $t$  de Student multivariée, mais l'expérience a amplement démontré que ces modèles n'étaient pas toujours adéquats. Plus précisément, bien que la distribution gaussienne multivariée puisse bien s'ajuster, dans l'ensemble, à un ensemble de données, elle donne de piètres résultats lorsqu'il s'agit d'obtenir des estimations précises de la probabilité d'événements extrêmes (de queue) simultanés, ou, de manière équivalente, lorsqu'il s'agit d'estimer la VaR du portefeuille agrégé  $S = \sum_{i=1}^d X_i$  pour un niveau de confiance élevé; cf. [McNeil et al. \[2010\]](#). L'utilisation du modèle gaussien multivarié se fonde aussi sur la (fausse) intuition que les corrélations<sup>3</sup> suffisent lorsqu'il s'agit de modéliser la dépendance ([Embrechts et al. \[1999\]](#), [Embrechts et al. \[2002\]](#)). Cette fausse idée sous-tend aussi l'utilisation de la méthode standard des variances-covariances pour agréger le capital dans Bâle III et dans Solvabilité II, et elle est aussi présente dans de nombreux dispositifs de gestion du risque en usage dans l'industrie. Qui plus est, en pratique, le nombre d'observations considérées comme des événements extrêmes est trop faible. De fait, il existe toujours un niveau à partir duquel il n'y a plus d'observation. Par conséquent, si l'on décide de modéliser la dépendance dans la queue de la distribution, cette modélisation sera forcément arbitraire dans une certaine mesure, du moins elle ne sera pas basée sur des données observées.

---

<sup>3</sup> Il est clair que l'utilisation de corrélations ne suffit pas pour modéliser la dépendance, puisqu'un seul nombre (à savoir la corrélation) ne suffit pas pour décrire l'interaction des variables, à moins qu'on ne pose d'autres hypothèses (p. ex. une structure de dépendance gaussienne).

Dernièrement, des ouvrages ont été publiés sur la création de modèles flexibles multivariés qui permettent un meilleur ajustement des données grâce, par exemple, à la construction de paires de copules ou de copules en vignes (pour en avoir un aperçu, cf., par exemple, [Aas et al. \[2009\]](#) ou [Czado \[2010\]](#)). Bien que ces modèles soient intéressants du point de vue théorique et intuitif, ils exigent, en pratique, un ensemble de données qui soit suffisamment riche pour pouvoir en tirer des résultats satisfaisants. Or, aucun modèle n'est parfait, et bien que ces nouvelles découvertes soient certainement nécessaires à l'évaluation précise du risque d'un portefeuille, elles ne sont utiles aux organismes de réglementation et aux gestionnaires du risque que si elles permettent de réduire de façon significative le risque de modélisation inhérent aux évaluations des risques.

Nous proposons ici un cadre permettant une quantification pratique du risque de modélisation et qui a été mis au point par [Bernard et Vanduffel \[2015a\]](#), en s'inspirant des idées d'[Embrechts et al. \[2013\]](#) et des ouvrages ci-inclus. En termes techniques, nous considérons  $N$  vecteurs observés  $\{(x_{1i}, \dots, x_{di})\}_{i=1, \dots, N}$  et nous supposons qu'un modèle à plusieurs variables a été ajusté à cet ensemble de données. Toutefois, on ne pourrait avoir confiance dans ce modèle ajusté dans les régions du support qui ne contiennent pas assez de points de données (p. ex., aux extrémités). L'idée est donc de subdiviser  $\mathbb{R}^d$  en deux sous-ensembles : le premier sous-ensemble, représenté par  $\mathcal{F}$ , porte le nom de « partie fixe », et le second, représenté par  $\mathcal{U}$ , est appelé « partie non fixe », ce dernier sous-ensemble englobant toutes les régions pour lesquelles le modèle ajusté ne donne pas de bons résultats. Cette approche comprend les deux directions que nous avons évoquées précédemment à l'égard de l'agrégation des risques. Plus précisément, si nous avons parfaitement confiance dans le modèle, toutes les observations se trouvent dans la partie « fixe » ( $\mathcal{U} = \emptyset$ ) et le risque de modélisation est nul. Si nous n'avons aucune confiance dans l'ajustement de la dépendance,  $\mathcal{F} = \emptyset$  et nous sommes dans le contexte décrit par [Embrechts et al. \[2013\]](#), qui calculent des limites de risque pour les portefeuilles lorsque les distributions marginales des composantes de risque sont connues mais que les données sur la dépendance sont absentes. L'approche préconisée par [Bernard et Vanduffel \[2015a\]](#) rend possible la prise en compte naturelle des données sur la dépendance et peut conduire à un rétrécissement des limites de risque. Ce cadre s'accompagne aussi d'un algorithme qui permet aux actuaires de prendre en compte le risque de modélisation d'une manière très pratique, comme nous le verrons en détail.

## 1. Mesure du risque agrégé

Essentiellement, les sociétés d'assurances perçoivent des primes en échange de réclamations aléatoires (futurs). Considérons un portefeuille contenant  $d$  polices, et soit  $X_i$  ( $i = 1, 2, \dots, d$ ) la perte correspondant à la  $i^{\text{e}}$  police, définie comme étant la réclamation aléatoire après déduction de la prime. Afin de protéger les titulaires de police et les autres créanciers de l'insolvabilité, l'organisme de réglementation exige que la perte du portefeuille,  $S = X_1 + X_2 + \dots + X_d$  soit « assez faible » par rapport aux ressources disponibles, disons une exigence en capital de  $K$ , ce qui signifie que le capital disponible  $K$  doit être tel que  $S - K$

constitue un « bon pari » pour les créanciers, c.-à-d. qu'ils sont « raisonnablement sûrs » que l'événement «  $S > K$  » a peu d'importance ([Tsanakas et Desli \[2005\]](#), [Dhaene et al. \[2012\]](#)).

À l'évidence, la mesure du degré de risque de  $S = X_1 + X_2 + \dots + X_d$  revêt une importance capitale pour l'établissement des exigences en capital. Cela dit, il existe plusieurs autres raisons d'étudier les propriétés de la perte agrégée  $S$ . De fait, l'une des tâches importantes prévues par le cadre de gestion du risque d'entreprise (GRE) concerne l'allocation du capital (des risques), c.-à-d. la répartition du capital total de l'assureur entre ses divers éléments (sous-groupes), dont les branches d'assurance, les types de risque et les régions géographiques. Ce faisant, il est possible de redistribuer le coût de détention du capital entre les divers éléments de manière à ce qu'il puisse être transféré aux déposants ou aux titulaires de police sous forme de frais (primes). L'allocation des risques permet aussi d'évaluer la performance des différentes branches d'assurance en déterminant le rendement du capital alloué à chacune d'entre elles. Enfin, l'exercice d'agrégation et d'allocation des risques peut faciliter l'identification des secteurs à risque dans l'entreprise et soutenir la prise de décision concernant l'expansion, la réduction, voire la suppression d'activités; cf. [Panjer \[2001\]](#), [Tsanakas \[2009\]](#).

Au moment de mesurer le risque agrégé  $S$ , il importe aussi de tenir compte du contexte présent. En particulier, les parties prenantes peuvent ne pas toutes avoir la même perception du degré de risque. Par exemple, les déposants et les titulaires de police ne s'intéressent principalement qu'à la probabilité que la société honore ses obligations. Quant aux organismes de réglementation, ils ont les mêmes intérêts que les déposants et les titulaires de police et ils établissent des règles que les sociétés doivent appliquer pour calculer le capital requis qu'elles doivent détenir. Toutefois, ils s'intéressent aussi à l'excédent de la perte par rapport au capital détenu, car cette somme doit être financée par la société en cas de renflouement de la compagnie. En termes formels, ils s'intéressent à l'*insuffisance* du capital de solvabilité requis  $\varrho(S)$  par rapport à la perte du portefeuille,  $S$ ; c.-à-d. que

$$(S - \varrho(S))_+ := \max(0, S - \varrho(S)) \quad (1.1)$$

L'insuffisance représente donc cette partie de la perte totale que l'assureur n'est pas en mesure de rembourser. Elle est aussi désignée sous le nom de *perte sociétale* ou de *manque à gagner pour les titulaires de police*. En raison de leur responsabilité limitée, les actionnaires n'ont pas vraiment à s'inquiéter du risque résiduel, mais s'intéressent plutôt aux propriétés de la variable  $S - (S - \varrho(S))_+ := \min(S, \varrho(S))$ . En résumé, les parties prenantes peuvent ne pas toutes avoir la même perception et la même sensibilité quant à la signification du risque qu'elles encourent, et elles peuvent employer différents paradigmes pour le définir et le mesurer.

Pour ce qui est de mesurer le risque, les deux mesures les plus importantes sont la VaR et la TVaR<sup>4</sup>. Pour un niveau de probabilité  $p$  donné, elles sont respectivement indiquées par  $\text{VaR}_p$  et  $\text{TVaR}_p$  et se définissent comme suit :

$$\text{VaR}_p(S) = \min\{s \mid P[S \leq s] \geq p\}, \quad 0 < p < 1, \quad (1.2)$$

<sup>4</sup> De ces deux mesures, la VaR est de loin la plus utilisée à l'heure actuelle en pratique par les organismes de réglementation et les gestionnaires du risque; cf., par exemple, [Jorion \[2006\]](#).

et

$$\text{TVaR}_p(S) = \frac{1}{1-p} \int_p^1 \text{VaR}_q[S] dq, \quad 0 < p < 1. \quad (1.3)$$

$\text{VaR}_p$  représente donc simplement la perte minimale que l'on observe selon une probabilité  $1-p$ , tandis que  $\text{TVaR}_p$  représente la moyenne de toutes les VaR supérieures.

### 1.1 Mesures de risque cohérentes

La VaR et la TVaR ne sont que deux cas particuliers de mesures de risque. De fait, nous pouvons utiliser toute fonctionnelle  $\varrho$  qui associe à la perte aléatoire  $X$  (appartenant à un ensemble  $\Gamma$  pertinent<sup>5</sup> de pertes aléatoires) un nombre  $\varrho[X]$ . Cela dit, il est logique d'imposer certaines propriétés (axiomes) à la mesure de risque  $\varrho$ . Nous définissons ci-après un ensemble typique (et intéressant) d'axiomes. D'un point de vue normatif, toutefois, il n'existe pas de « meilleur ensemble d'axiomes », puisque toute axiomatique normative repose sur la « conviction » dans la vérité des axiomes qui la composent. Nous posons :

- *Homogénéité positive* : pour tout  $X \in \Gamma$  et  $a > 0$ ,  $\varrho[aX] = a\varrho[X]$ .
- *Invariance par translation* : pour tout  $X \in \Gamma$  et  $b \in \mathbb{R}$ ,  $\varrho[X + b] = \varrho[X] + b$ .
- *Monotonie* : pour tout  $X, Y \in \Gamma$ ,  $X \leq Y$  implique que  $\varrho[X] \leq \varrho[Y]$ .
- *Sous-additivité* : pour tout  $X, Y \in \Gamma$ ,  $\varrho[X + Y] \leq \varrho[X] + \varrho[Y]$ .

Dans [Artzner et al. \[1999\]](#), lorsqu'une mesure de risque possède les quatre propriétés susmentionnées que sont la monotonie, l'homogénéité positive, l'invariance par translation et, notamment, la sous-additivité, elle est dite *cohérente*. On sait très bien que la VaR n'est pas sous-additive et que la TVaR l'est pour tout  $p$ . De fait, on voit tout de suite que la TVaR est la plus petite mesure de risque cohérente qui est plus prudente que la VaR (laquelle n'est pas cohérente) (pour la preuve, cf. [Artzner et al. \[1999\]](#) et aussi [Dhaene et al. \[2006\]](#)).

Alors que les trois premières propriétés ne sont pas sujettes à controverse, l'avantage pour une mesure de risque de posséder la propriété de sous-additivité a fait l'objet de nombreux débats et recherches (cf. aussi la section [2.1](#)). Dans la prochaine partie, nous expliquons que la sous-additivité est d'ordinaire une contrainte naturelle. À cet égard, nous soulignons que le terme « cohérente » prête à confusion, car il peut laisser penser que toute mesure de risque « incohérente » est inadéquate. On notera que le principe bien connu de l'écart-type, défini par  $\varrho(X) = \mathbb{E}(X) + k\sqrt{\text{var}(X)}$  pour une certaine constante  $k$ , ne satisfait pas à l'axiome de la monotonie et n'est donc pas cohérent<sup>6</sup>. Ci-après, nous supposons, à l'exemple de la littérature

<sup>5</sup> En particulier, l'ensemble  $\Gamma$  contient les pertes aléatoires  $X_i$  ( $i = 1, 2, \dots, d$ ) et nous supposons que  $X_i, X_j \in \Gamma$  implique que  $X_i + X_j \in \Gamma$ , et aussi que  $aX_i \in \Gamma$  pour tout  $a > 0$  et que  $X_i + b \in \Gamma$  pour tout nombre réel  $b$ .

<sup>6</sup> On dit qu'une mesure du risque de distorsion, définie par  $\varrho(X) = \int_0^1 F^{-1}(t)g'(1-t)dt$  pour une fonction croissante  $g$  avec  $g(0) = 0$ ,  $g(1) = 1$ , est cohérente si  $g$  est concave sur  $[0,1]$ . Nous renvoyons le lecteur à

théorique et de l'usage actuel, que  $\rho(X)$  ne dépend que de la distribution de  $X$  (c.-à-d. que  $\rho(X)$  est une fonctionnelle de la distribution de  $X$  et elle est appelée mesure de risque invariante en loi).

## 1.2 Contrôle ex post et robustesse des mesures de risque

**Contrôle ex post** : En fin de compte, si nous nous servons d'un modèle, c'est pour évaluer le degré de risque de  $S$  et obtenir un nombre  $\rho(S)$ . Dans de nombreux cas, il est possible de construire plusieurs modèles en concurrence qui se ressemblent tous du point de vue de la disponibilité des données et qui ne diffèrent qu'en ce qui concerne les hypothèses ad hoc qu'ils renferment.

Une façon naturelle de comparer les modèles en concurrence consiste à employer une mesure d'erreur qui utilise les prévisions ponctuelles et les observations réelles. Plus précisément, la performance d'un modèle peut être résumée à l'aide de la moyenne  $\bar{T}$  de la fonction  $T$  de score, calculée sur  $n$  points de prévision, c.-à-d. que

$$\bar{T} = \sum_{i=1}^n T(x_i, y_i), \quad (1.4)$$

où le  $i^{\text{e}}$  point de prévision correspond au couple  $(x_i, y_i)$  dans lequel  $x_i$  est la prévision ponctuelle et  $y_i$  l'observation ( $i = 1, 2, \dots, n$ ). Comme exemples typiques de fonction de score, citons l'erreur quadratique  $T(x, y) = (x - y)^2$  et l'erreur absolue  $T(x, y) = |x - y|$ .

[Gneiting \[2011\]](#) a montré que la fonction  $T$  de score à utiliser doit être adaptée à la mesure de risque considérée, sinon on risque d'en tirer de mauvaises conclusions. Cet auteur soutient qu'il faut évaluer la qualité du modèle (utilisé pour prévoir la fonctionnelle  $\rho(s)$ ) en utilisant une fonction de score qui produirait cette fonctionnelle comme prévision ponctuelle optimale. Si une fonction de score est donnée, la prévision optimale (à supposer que les observations soient distribuées de manière identique et indépendante), après application de la règle de Bayes, s'obtient comme suit :

$$x^* = \arg \min_x E(T(x, S)), \quad (1.5)$$

Par exemple, si la fonction de score est l'erreur quadratique, on sait que la prévision optimale correspond à la moyenne de  $S$ , tandis que si la fonction de score est l'erreur absolue, la solution est donnée par la médiane de  $S$ . Si cette correspondance entre mesure de risque (fonctionnelle) et fonction de score existe, la mesure de risque est qualifiée d'*elicitable*. Par exemple, la moyenne et la médiane sont *elicitable*. La Var est elle aussi *elicitable*, car l'utilisation d'une fonction de score linéaire par morceaux (généralisée) est compatible avec les estimations obtenues par la VaR. Toutefois, ce ne sont pas toutes les mesures de risque qui sont *elicitable* : l'écart-type ne l'est pas, et, notamment, la TVaR ne l'est pas non plus (cf. aussi [Ziegel \[2014\]](#) et [Embrechts et Hofert \[2014\]](#)).

Lorsque les mesures de risque ne sont pas *elicitable*, il peut y avoir des incohérences au moment de comparer les prévisions ponctuelles de différents modèles et(ou) de différents

---

[Wang \[2000\]](#), [Bäuerle et Müller \[2006\]](#) et [Föllmer et Schied \[2010\]](#) et aux ouvrages de référence qu'ils citent s'il souhaite prendre connaissance des études sur les mesures de risque et leurs propriétés.

spécialistes. Supposons que vous ayez un modèle dont on sait qu'il produit la meilleure estimation de la perte du portefeuille selon la VaR à 99,5 %, mais qu'il existe aussi un autre modèle dont on sait qu'il produit une estimation non optimale selon la VaR à 99,5 %. Si vous utilisez l'erreur quadratique comme fonction de score (qui n'est pas compatible avec la VaR à 99,5 %) pour évaluer les estimations à l'aide de la VaR à 99,5 %, il se peut que vous choisissiez en fin de compte des modèles non optimaux, simplement parce que vous avez utilisé la mauvaise mesure pour évaluer les estimations obtenues à l'aide de la VaR à 99,5 %.

Cette analyse appelle quelques commentaires : la TVaR n'est pas *elicitable* mais elle l'est indirectement, car elle peut être décomposée en une moyenne conditionnelle et un quantile, qui sont tous deux *elicitable* séparément. De plus, le contrôle ex post nécessite un riche échantillon de données de prévisions et d'observations qu'il n'est pas facile d'obtenir dans le contexte des évaluations de solvabilité, dont l'horizon est habituellement d'un an. Qui plus est, l'argument de la concordance qui sert à lier une mesure de risque à une fonction de score optimale repose sur l'hypothèse que toutes les observations sont distribuées de manière identique et indépendante, ce qui n'est pas toujours le cas en pratique.

**Robustesse** : Un autre sujet important est celui de la robustesse de la mesure de risque sous l'effet d'une erreur de spécification du modèle ou de petites variations dans les données. Du point de vue de l'organisme de réglementation, la mesure de risque utilisée devrait réellement être stable sous l'effet de diverses hypothèses de modélisation ou de petites variations dans les données. Dans le contexte de Solvabilité II, deux assureurs possédant le même portefeuille devraient obtenir la même VaR relativement à ce portefeuille. Toutefois, lorsque le bon modèle ne peut être identifié avec une (quasi) certitude, les assureurs peuvent utiliser deux modèles différents et obtenir des résultats de VaR considérablement différents. Par exemple, [Chernih et al. \[2010\]](#) ont montré qu'il était possible de construire un modèle de portefeuille de risques de crédit qui est compatible avec le modèle standard MKMV<sup>7</sup>, à l'exception près que le modèle MKMV admet une dépendance gaussienne entre les rendements d'actif, tandis que [Chernih et al. \[2010\]](#) emploient une différente copule (qui, néanmoins, produit les mêmes corrélations que celles du modèle MKMV). Les deux modèles sont donc parfaitement compatibles pour ce qui est des données disponibles sur l'exposition, la perte en cas de défaut, les probabilités de défaut et les corrélations, mais lorsque ces modèles sont employés pour estimer la VaR à 99,5 % d'un portefeuille typique de prêts, leurs résultats peuvent différer d'un facteur pouvant atteindre 15; pour de plus amples témoignages et exemples, cf. aussi [Heyde et Kou \[2004\]](#), [Kou et al. \[2013\]](#), [Bernard et al. \[2013b\]](#) et [Bernard et al. \[2015\]](#). Au vu de ces observations, [Bernard et al. \[2013b\]](#) déconseillent l'utilisation de la VaR à des niveaux de confiance élevés (p. ex., 99,5 %) comme base de calcul des exigences en capital. On notera aussi que si la mesure de risque externe n'est pas robuste, les institutions peuvent contourner la

---

<sup>7</sup> Le modèle MKMV est utilisé par un grand nombre d'institutions financières pour évaluer le degré de risque de portefeuilles de risques de crédit. De plus, le dispositif standard de Bâle III s'en sert pour déterminer le capital requis que les banques doivent détenir compte tenu du risque de crédit qu'elles encourent; cf. [Comité de Bâle \[2010b\]](#). Par ailleurs, le cadre de Solvabilité II utilise cette formule pour décider du montant de capital que les assureurs doivent détenir par mesure de prévention contre les conséquences défavorables de la faillite d'une ou de plusieurs contreparties.

réglementation en choisissant un modèle qui permet de réduire significativement les exigences en capital ou en manipulant les données d'entrée.

## 2. Agrégation et diversification

### 2.1 Avantages de la diversification et sous-additivité

Sur le site Web de l'organisme de réglementation canadien ([BSIF \[2014\]](#)), on y lit : « [...] nous définissons l'agrégation comme l'approche utilisée pour calculer individuellement et collectivement les éléments de risque. Un crédit pour la diversification est accordé lorsque la méthode d'agrégation des risques produit des résultats qui sont inférieurs à la somme de tous les éléments de risque. » Les avantages de la diversification peuvent découler du groupement de risques d'un certain type tel que le risque d'assurance, ou du groupement de plusieurs types de risque (p. ex., le risque d'assurance et le risque d'actif), entre entités ou entre régions. On y note combien il est difficile de déterminer les avantages de la diversification en période de crise. Les exigences en capital sont établies de façon que la société puisse faire face aux périodes de crise, et c'est précisément en ces moments-là que les avantages de la diversification s'estompent. La réduction du capital au titre de la diversification ne devrait être autorisée que lorsque les avantages de la diversification demeurent valables en périodes de crise. Certains avantages devraient cependant être reconnus. Cf. [BSIF \[2014\]](#) pour une analyse des avantages de la diversification entre le risque de volatilité et, respectivement, le risque de mortalité, le risque de morbidité, le risque de longévité et le risque de déchéance.

Considérons des portefeuilles dont les pertes respectives sont  $S_1$  et  $S_2$ , et soit  $\varrho$  une mesure de risque servant à fixer les exigences en capital; c.-à-d. que  $\varrho(S_1)$  représente le capital pour le premier portefeuille,  $\varrho(S_2)$  représente le capital pour le second portefeuille et  $\varrho(S_1 + S_2)$  le capital pour la position combinée (fusion des portefeuilles). Nous supposons que la nature des pertes  $S_1$  et  $S_2$  ne change pas du fait de la fusion des portefeuilles. Toutefois, dans la réalité, la fusion ou le fractionnement de portefeuilles peut entraîner, entre autres choses, un changement de direction, de stratégie commerciale ou de structure des coûts et ainsi modifier la distribution marginale des pertes considérées.

Nous donnons ci-dessous la définition standard d'un avantage de la diversification, désigné par  $DB(\varrho, S_1, S_2)$

$$DB(\varrho, S_1, S_2) = \varrho(S_1) + \varrho(S_2) - \varrho(S_1 + S_2). \quad (2.6)$$

$DB(\varrho, S_1, S_2)$  représente donc le gain (la perte) découlant de la fusion des deux portefeuilles. Il est clair que si  $\varrho$  est cohérent (et donc sous-additif), l'avantage de la diversification est non négatif, ce qui, intuitivement, a du sens, car le groupement de risques présente des avantages. Afin de confirmer cette intuition, notons que :

$$(S_1 + S_2 - \varrho(S_1) - \varrho(S_2))^+ \leq \sum_{j=1}^2 (S_j - \varrho(S_j))^+. \quad (2.7)$$

L'inégalité (2.7) indique que le risque d'insuffisance lié à la détention du portefeuille fusionné est toujours inférieur à la somme des risques d'insuffisance liés à la détention des portefeuilles individuels, lorsque le capital de solvabilité requis est additif. Elle signifie que, du point de vue

de l'organisme de réglementation, la fusion est avantageuse, car le risque d'insuffisance diminue lorsque les capitaux sont regroupés. La raison en est claire : dans le portefeuille fusionné, l'insuffisance de capital de l'une des entités peut être compensée par le gain de l'autre. En résumé, *la fusion réduit l'insuffisance de capital*. L'inégalité (2.7) indique donc que le capital de solvabilité correspondant à la position combinée peut être inférieur à la somme des capitaux de solvabilité correspondant aux deux portefeuilles individuels. Ces observations confortent la croyance générale selon laquelle le capital de solvabilité requis (mesure de risque) devrait être sous-additif. De fait, lorsqu'il y a fusion de deux portefeuilles individuels, l'organisme de réglementation prend en compte la sous-additivité aussi longtemps que se vérifie

$$(S_1 + S_2 - \varrho(S_1 + S_2))^+ \leq \sum_{j=1}^2 (S_j - \varrho(S_j))^+$$

À cet égard, notons que la condition de sous-additivité implique que

$$(S_1 + S_2 - \varrho(S_1 + S_2))^+ \geq (S_1 + S_2 - \varrho(S_1) - \varrho(S_2))^+, \quad (2.8)$$

et que, par conséquent, pour certaines valeurs observées  $(s_1, s_2)$ , nous pouvons avoir

$$(s_1 + s_2 - \varrho(S_1 + S_2))^+ > (s_1 - \varrho(S_1))^+ + (s_2 - \varrho(S_2))^+.$$

L'utilisation d'une mesure de risque sous-additive peut donc donner lieu à une plus grande insuffisance de capital que la somme des insuffisances des entités individuelles, c.-à-d. que

$$(s_1 + s_2 - \varrho(S_1 + S_2))^+ > (s_1 - \varrho(S_1))^+ + (s_2 - \varrho(S_2))^+$$

peut se vérifier ([Dhaene et al. \[2008\]](#)). En conséquence, bien que la sous-additivité soit une propriété acceptable du point de vue des organismes de réglementation, ils devraient limiter le degré de sous-additivité afin d'empêcher que  $(S_1 + S_2 - \varrho(S_1 + S_2))^+$  ne devienne trop risquée par rapport à  $(S_1 - \varrho(S_1))^+ + (S_2 - \varrho(S_2))^+$ .

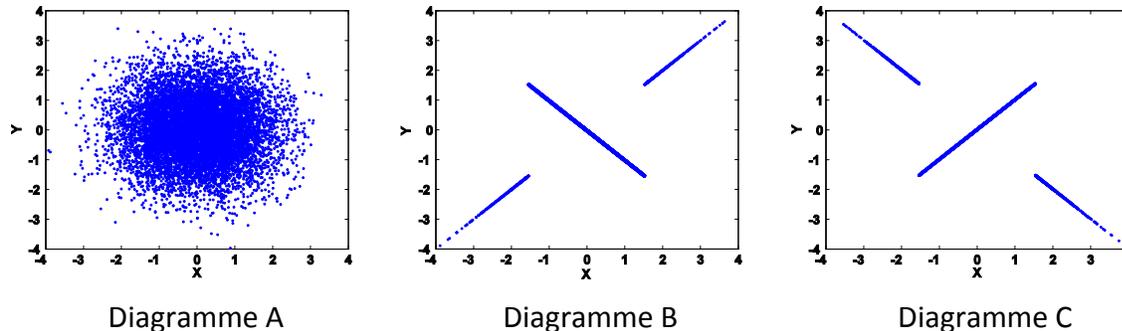
À cet égard, il importe aussi de faire remarquer que, pour les actionnaires, l'avantage d'une fusion n'est pas clairement établi. Nous l'expliquons comme suit. Pour un portefeuille  $j$  ( $j = 1, 2$ ), les fonds disponibles en fin d'année sont donnés par  $(\varrho(S_j) - S_j)^+$ . Si la perte  $S_j$  est inférieure au capital  $\varrho(S_j)$ , les fonds qui appartiennent aux actionnaires (à la fin de la période de référence) seront donnés par  $\varrho(S_j) - S_j$ , tandis que si la perte  $S_j$  est supérieure à  $\varrho(S_j)$ , l'unité commerciale rattachée à ce portefeuille est ruinée et les fonds disponibles sont égaux à zéro. Vu que

$$(\varrho(S_1) + \varrho(S_2) - S_1 - S_2)^+ \leq \sum_{j=1}^2 (\varrho(S_j) - S_j)^+, \quad (2.9)$$

nous remarquons qu'il est préférable, du point de vue des actionnaires, de détenir séparément les deux portefeuilles, essentiellement parce que, dans ce cas, une protection est intégrée, ce qui empêche la mauvaise performance d'un portefeuille d'avoir une incidence sur l'autre. En fait, les intérêts des actionnaires et des organismes de réglementation ne sont pas tout à fait les mêmes; à ce sujet, cf. aussi [Dhaene et al. \[2008\]](#) et [Dhaene et al. \[2009\]](#).

## 2.2 L'erreur de n'utiliser que les corrélations

Certains professionnels semblent croire que, lorsqu'il s'agit d'agréger deux risques, il suffit de savoir quel est leur coefficient de corrélation. Cette (fausse) intuition est sans doute attribuable à l'usage répandu et à l'importance<sup>8</sup> de la distribution normale multivariée, qui est entièrement caractérisée par la spécification des moyennes, des écarts-types et des paires de corrélations (Embrechts et al. [1999, 2002]). Or, il faut savoir que la distribution normale multivariée comporte déjà de manière inhérente le choix d'une dépendance (gaussienne) spécifique et que seules les corrélations sont requises pour pouvoir paramétrer cette dépendance gaussienne. De fait, il est facile de construire deux variables aléatoires normales ayant un coefficient de corrélation spécifique, mais dont la distribution conjointe (à deux variables) *n'est pas* normale. Pour illustrer ce fait, soit  $X$  et  $Y$  deux variables aléatoires normales standards et indépendantes. Notons, en particulier, qu'elles ont une dépendance gaussienne avec une corrélation nulle. Ensuite, nous définissons  $Z_c$  comme étant  $Z = -X$  si  $|X| < c$  et  $Z = X$  si  $|X| > c$  ( $c > 0$ ). Il est facile de voir que  $Z$  suit aussi une distribution normale standard : elle affiche une corrélation positive parfaite avec  $X$  aux extrémités de la distribution et une corrélation négative parfaite ailleurs. Ensuite, on peut choisir  $c^*$  de façon que la corrélation entre  $X$  et  $Z_{c^*}$  soit égale à zéro ( $c^* \approx 1,538$ ). Ainsi, lorsque  $p > \Phi(c^*)$  ( $\Phi(\cdot)$  indique la fonction de répartition de la variable aléatoire normale standard),  $\text{VaR}_p(X + Z_{c^*}) = 2\Phi^{-1}(p)$  alors que  $\text{VaR}_p(X + Y) = \sqrt{2}\Phi^{-1}(p)$ . La figure 2.1 en donne une illustration numérique.



**Figure 2.1 :** Illustration de trois situations dans lesquelles les variables aléatoires suivent une distribution normale standard et ont une corrélation nulle.

Un autre exemple illustrant les défauts des corrélations a trait à la mesure du risque d'un portefeuille de prêts de crédit. Pour expliquer cette idée, considérons les risques  $X_i$  ( $i = 1, 2, \dots, n$ ) qui indiquent des événements de défaut, en sorte que  $S = X_1 + X_2 + \dots + X_n$  représente le nombre de défauts du portefeuille. Plus précisément, soit  $p_i$  la probabilité que la  $i^{\text{e}}$  société fasse défaut et  $p_{ij}$  la probabilité conjointe que la société  $i$  et la société  $j$  fassent

<sup>8</sup> La distribution à plusieurs variables est au cœur de nombreuses théories et applications, telles que la régression linéaire, l'analyse en composantes principales, le modèle d'évaluation des actifs financiers, l'analyse de moyenne-variance de Markowitz, l'analyse discriminante, l'agrégation du capital (p. ex. Bâle III et Solvabilité II), modélisation du portefeuille de crédit (modèle KMV de Moody).

toutes deux défaut. Les corrélations des probabilités de défaut,  $\rho_{ij}^D$  ( $i, j = 1, 2, \dots, n$ ), sont ensuite données par

$$\rho_{ij}^D = \frac{P_{ij} - P_i P_j}{\sqrt{P_i(1 - P_i)}\sqrt{P_j(1 - P_j)}}. \quad (2.10)$$

En d'autres termes, les corrélations ne donnent une information complète que sur l'interaction de deux événements de défaut (événements groupés par paires), mais non réellement sur la manière dont trois prêts ou plus interagissent. À cet égard, nous notons l'absence en soi de statistiques suffisantes sur les défauts (les défauts simultanés sont des événements très rares en soi) et on peut s'attendre à ne pas pouvoir estimer de façon sûre des probabilités de défaut conjointes d'ordre supérieur. Autrement dit, l'évaluation du risque d'un portefeuille de risques de crédit est soumise à l'incertitude du modèle<sup>9</sup>. Par exemple, le fameux modèle MKMV associe les défauts des sociétés au comportement du rendement des actifs et suppose que ces rendements suivent une loi normale multivariée. Or, cette hypothèse n'en est qu'une parmi d'autres et rien n'indique qu'elle approche la réalité. [Bernard, Rüschenndorf et Vanduffel \[2013b\]](#) ont évalué l'incidence de l'incertitude du modèle sur les calculs de la VaR. Lorsqu'on utilise  $p = 99,5 \%$  comme base pour calculer la VaR et le capital requis (comme dans Bâle III et Solvabilité II), les résultats des modèles dans l'industrie se situent d'habitude à l'intérieur d'un large éventail de valeurs possibles de la VaR. Par contre, le risque de modélisation semble être plus limité lorsqu'on utilise des niveaux de probabilités plus modérés pour évaluer la VaR. Ces auteurs concluent qu'il pourrait être utile d'imposer des contraintes supplémentaires aux modèles lorsqu'ils servent à fixer le capital requis. Par exemple, il est possible d'utiliser les bornes de la VaR obtenues pour établir une valeur minimale à la VaR que produit le modèle interne, ou d'imposer un modèle particulier aux diverses institutions pour le calcul du capital requis, car cela faciliterait la comparaison des niveaux de capital et permettrait d'assurer une concurrence loyale.

### 2.3 L'impact des microcorrélations

Nous évoquons ici un autre point faible de la corrélation. [Kousky et Cooke \[2012\]](#) ont expliqué comment les risques catastrophiques étaient habituellement caractérisés par de la dépendance et une distribution à queue épaisse. Lorsque les distributions des pertes ont une queue épaisse, la probabilité d'un événement ne diminue que lentement, par rapport à sa sévérité, ce qui signifie que les grosses pertes ne sont pas exceptionnelles (pour une explication plus mathématique, nous renvoyons le lecteur à [Kousky et Cooke \[2012\]](#)). On a pu montrer qu'un grand nombre de catastrophes naturelles affichaient une distribution à queue épaisse. Tel qu'il est amplement expliqué en détail dans [Kousky et Cooke \[2012\]](#), l'étude des catastrophes introduit un autre type de dépendance, la dépendance de queue. La dépendance de queue se réfère à la probabilité qu'une variable dépasse un certain centile, sachant qu'une autre variable a déjà dépassé ce centile. En termes plus simples, cela signifie que les mauvaises choses ont tendance à survenir ensemble. Il est clair qu'une catastrophe peut frapper simultanément plusieurs branches d'un assureur (assurances habitation, automobile, maladie, biens des

<sup>9</sup> [Duffie et Singleton \[2012\]](#).

entreprises, etc.). Après avoir étudié plus de 700 événements de tempête en France, [Lescourret et Robert \[2006\]](#) ont constaté une dépendance de queue de la sorte pour les branches d'assurance. De plus, les risques catastrophiques tendent à afficher une corrélation spatiale en raison de la forte dépendance entre les sinistres qui découlent d'une catastrophe particulière. En pratique, cette corrélation diminue en fonction de la distance spatiale entre les polices. Lorsqu'elle tombe à zéro, elle permet aux assureurs de diversifier leurs risques en détenant des polices dans différentes régions. [Kousky et Cooke \[2012\]](#) ont constaté, à regret, qu'en ce qui concerne les avantages de la diversification, « près de zéro » n'équivaut pas à zéro. Même de petites corrélations positives moyennes entre les polices, que les auteurs qualifient de « microcorrélations », peuvent fausser l'agrégation des risques.

Le principal problème lié à l'existence de la microcorrélation est dû au fait que la loi des grands nombres ne s'applique pas lorsque les risques ne sont pas indépendants, même si ceux-ci affichent un coefficient de corrélation qui est très près de zéro. Ce phénomène est très bien expliqué dans les ouvrages de [Kousky et Cooke \[2009\]](#), de [Cooke et Kousky \[2010\]](#) et [Cooke et al. \[2011\]](#) ont appliqué des risques catastrophiques dans [Kousky et Cooke \[2012\]](#). L'idée de base est très simple et elle se fonde sur les situations dans lesquelles les polices affichent une petite corrélation positive moyenne (disons de 0,04, qui est la corrélation moyenne que [Cooke et Kousky \[2010\]](#) ont observé entre les réclamations d'assurance aux États-Unis, à l'échelle du comté). Cooke et Kousky ont montré combien une toute petite corrélation positive entre les polices pouvait vite devenir pernicieuse.

Soit  $X_1, \dots, X_n$  et  $Y_1, \dots, Y_n$  deux ensembles de variables aléatoires ayant la même variance moyenne  $\sigma^2$  et la même covariance moyenne  $C$  (dans et entre les ensembles). La corrélation de la somme des  $X$  et de la somme des  $Y$  s'obtient aisément comme suit :

$$\text{corr}\left(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i\right) = \frac{n^2 C}{n\sigma^2 + n(n-1)C} \quad (2.11)$$

Le principal problème tient au fait que cette formule tend vers 1 au fur et à mesure que  $n$  augmente, si  $C$  est non nulle (même très petite) et que  $\sigma^2$  est finie. Si toutes les variables sont indépendantes,  $C = 0$  et la corrélation dans (2.11) est égale à zéro. Pour mettre en évidence cette amplification de la corrélation, [Kousky et Cooke \[2009\]](#) se sont servi de données sur les réclamations d'assurance inondation. Ils ont choisi au hasard des paires de comtés des États-Unis et calculé leur corrélation. L'histogramme en vert de la figure 2.2 montre 500 de ces corrélations. La corrélation moyenne est de 0,04. Si quelques comtés ont affiché des corrélations fortes et positives, la plupart d'entre elles étaient très petites et avoisinaient zéro. Plutôt qu'étudier les corrélations entre deux comtés choisis au hasard, ils ont ensuite fait la somme de 100 comtés choisis au hasard, qu'ils ont mis en corrélation avec la somme de 100 autres comtés distincts choisis au hasard. Après avoir répété cette opération 500 fois, ils ont obtenu l'histogramme en bleu, où la moyenne de 500 de ces corrélations (de 100) est de 0,23. L'histogramme en rouge représente 500 corrélations (de 500) avec une valeur moyenne de 0,71. La hausse spectaculaire de la corrélation est le fait des microcorrélations entre les variables individuelles.

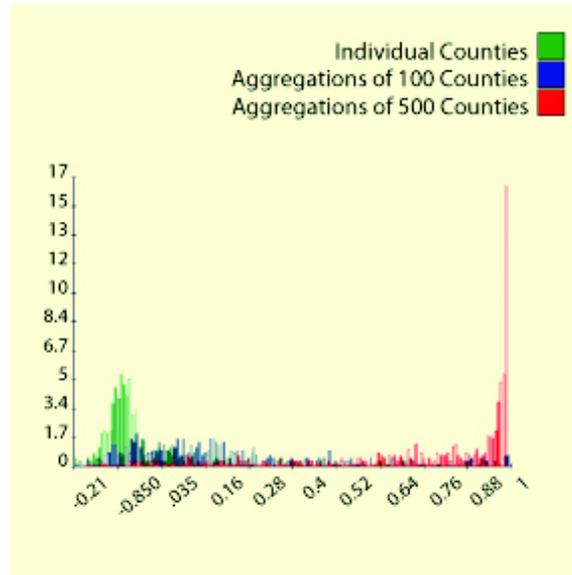


Figure 2.2 : Reproduction de la figure 11 de [Kousky et Cooke \[2009\]](#) à titre illustratif.

## 2.4 Ajustement d'une distribution multivariée

En pratique, il existe des techniques statistiques efficaces et précises permettant d'estimer les distributions marginales respectives de  $\mathbf{X} = (X_1, \dots, X_d)$ . En revanche, la structure de dépendance conjointe de  $\mathbf{X}$  est souvent beaucoup plus difficile à saisir : l'inférence statistique de données pluridimensionnelles s'accompagne de problèmes de calcul et de convergence, et le choix de distributions multivariées est assez limité en comparaison de la modélisation de distributions marginales. Pourtant, le choix d'une mauvaise hypothèse de dépendance peut entraîner d'importantes conséquences en matière de gestion des risques. Par exemple, le mauvais usage de la copule gaussienne multivariée peut avoir pour effet de sous-estimer de façon importante la probabilité de défauts simultanés dans un grand panier d'entreprises ([McNeil et al. \[2010\]](#)).

Le moyen le plus facile (et, par conséquent, le plus répandu) de modéliser une distribution multivariée consiste à utiliser une distribution gaussienne ou une distribution de Student multivariée. L'avantage de cette dernière est qu'elle affiche une certaine dépendance de queue. Toutefois, il existe des limites relativement à cette dépendance multivariée, car il n'y a qu'un seul paramètre de degrés de liberté qui détermine la dépendance de queue de toutes les paires de variables.

De façon plus générale, la distribution multivariée peut être décomposée en distributions marginales  $F_{X_i}$ ,  $i = 1, 2, \dots, d$  (qui représentent les risques individuels) et une fonction copule  $C$ , ainsi nommée, reflétant la dépendance. Plus précisément, le théorème de [Sklar \[1959\]](#) dit qu'il existe un vecteur  $(U_1, U_2, \dots, U_d)$  de variables aléatoires standards distribuées uniformément, tel que

$$\mathbf{X} = (F_{X_1}^{-1}(U_1), F_{X_2}^{-1}(U_2), \dots, F_{X_n}^{-1}(U_d)). \quad (2.12)$$

où « = » désigne l'égalité dans la distribution. L'égalité (2.12) montre donc que les propriétés distributionnelles du portefeuille  $\mathbf{X}$  sont complètement spécifiées par les distributions marginales  $F_{X_i}$  ( $i = 1, 2, \dots, d$ ) de ses composantes de risque et la distribution conjointe  $C$  des  $(U_1, U_2, \dots, U_d)$  décrivant l'interaction des risques du portefeuille.

Les copules ont fait l'objet d'une étude approfondie par [Joe \[1997\]](#) et [Nelsen \[2007\]](#). Vu qu'il existe de grandes familles de copules bidimensionnelles, il est relativement facile de modéliser la dépendance entre deux variables. Les copules bidimensionnelles dont l'usage est le plus répandu sont les copules archimédiennes, pour lesquelles il existe une importante littérature traitant de l'estimation et de la qualité de l'ajustement; cf. [Joe \[1997\]](#). Bedford et Cooke [\[2001, 2002\]](#) ont ensuite proposé de construire une copule multivariée en se servant de paires de copules comme composantes de base. Ils ont aussi fait des représentations graphiques consistant en une suite d'arbres hiérarchiques, qu'ils ont appelés copules en vignes régulières. Ce modèle multivarié, également désigné sous le nom construction de paires de copules, permet de décomposer un modèle multivarié complexe en des composantes bidimensionnelles plus simples. [Czado \[2010\]](#) en donne un aperçu. Cette approche est très flexible et permet à la dépendance entre n'importe quel sous-ensemble de deux variables de se différencier. Pour en savoir plus sur les moyens d'estimer les paramètres des copules en vignes, se reporter à [Kurowicka et Cooke \[2006\]](#). Comme remplacement aux constructions de copules par paires, [Hofert \[2012\]](#) propose l'utilisation d'un modèle hiérarchique; cf. [Okhrin et al. \[2013\]](#) pour les questions d'estimation. Les copules archimédiennes hiérarchiques ont été étudiées par [Hofert et Pham \[2013\]](#) et ont été utilisées par [Savu et Trede \[2010\]](#). On trouvera dans [Embrechts et Hofert \[2013\]](#) une vue d'ensemble complète de la question de la dépendance dans les grandes dimensions.

## 2.5 Résumé

Il est essentiel de prendre en compte la dépendance entre les composantes de risque lorsqu'il s'agit d'évaluer le risque agrégé d'un portefeuille. Nous avons montré que la sous-additivité d'une mesure de risque était justifiée du point de vue d'un organisme de réglementation. Autrement dit, il est justifié que les sociétés obtiennent des avantages de la diversification lorsqu'elles agrègent les risques. Toutefois, la prudence est de mise : les avantages de la diversification sont souvent évalués au moyen des corrélations, mais celles-ci constituent une piètre mesure de la dépendance. Il ne s'agit que d'un seul chiffre qui ne suffit pas à décrire l'interaction complexe des composantes de risque. Nous terminons la section 2 en examinant la façon de décrire la dépendance dans son intégralité.

## 3. Survol de la réglementation à l'heure actuelle

Le rapport du [Comité de Bâle \[2010a\]](#) décrit les méthodes de modélisation qu'utilisent les institutions financières et les organismes de réglementation dans divers pays pour agréger les risques. Il expose aussi les conditions dans lesquelles ces techniques d'agrégation fonctionnent comme prévu dans le modèle, et propose des possibilités d'amélioration. Le Comité émet des doutes quant à la fiabilité des résultats qui découlent de modèles internes d'agrégation des

risques qui tiennent compte des avantages de la diversification. Selon ce rapport : [traduction] « Les résultats des modèles doivent faire l'objet d'un examen attentif et être analysés avec prudence lorsqu'il s'agit de déterminer si les avantages de la diversification sont fiables et robustes. » Dans la section qui suit, nous ferons une courte synthèse des résultats de leurs recherches et de ceux d'autres organismes de réglementation.

### 3.1 Dispositifs de réglementation

**Dispositif de réglementation bancaire de Bâle III.** On calcule le total du capital minimal requis comme étant la somme du capital requis au titre du risque de crédit, du risque opérationnel et du risque de marché, sans tenir compte des avantages de la diversification entre les trois types de risque. L'idée selon laquelle l'absence de diversification correspond à la pire des situations du portefeuille n'est pas tout à fait exacte. En principe, cette propriété se vérifie lorsqu'une mesure de risque cohérente est utilisée, mais ce n'est pas forcément le cas pour d'autres mesures de risque telles que la VaR. En d'autres termes, il pourrait être possible d'agrèger les risques de façon que la VaR du risque agrégé soit supérieure à la somme des VaR.

En ce qui concerne le risque de marché, les banques ont le choix entre deux méthodes. Elles peuvent tirer avantage de la diversification si elles utilisent l'approche des modèles internes. Avec la méthode d'évaluation standard, le capital minimal requis pour le risque de marché est égal à la somme des charges de capital calculées pour chacun des types de risque (risque de taux d'intérêt, risque actions, risque de change, risque marchandises, risque prix des options).

**Au Canada, Test du capital minimal (TCM) et Montant minimal permanent requis pour le capital et l'excédent (MMPRCE).** Au Canada, le capital exigé des assureurs IARD se fonde sur le TCM. S'appuyant sur des facteurs, le TCM agrège les risques sous forme d'une somme qui autorise un crédit explicite pour la diversification entre le risque d'assurance et la somme du risque de crédit et du risque de marché, de façon que le total du capital requis pour ces risques soit inférieur à la somme du capital requis pour eux individuellement.

Par contre, le capital exigé des sociétés d'assurance-vie au Canada est calculé selon le MMPRCE établi par le Bureau du surintendant des institutions financières (BSIF), qui emploie des méthodes plus poussées à certains égards. « Le MMPRCE impose des exigences en capital pour les composantes de risque suivantes : risque de rendement insuffisant de l'actif, risque de mortalité, risque de morbidité, risque de déchéance, risque de désintermédiation, risqué lié aux garanties des fonds distincts. » [traduction] ([Comité de Bâle \[2010a\]](#)). Une partie des avantages de la diversification peut être prise en compte dans le calcul du risque de mortalité, du risque de morbidité et du risqué lié aux garanties des fonds distincts, mais le MMPRCE total s'obtient par la somme de chaque risque sans réduction possible pour la diversification. Ici encore il est (implicitement) supposé qu'il s'agit de la pire situation possible. On trouvera sur le site Web du BSIF ([www.osfi-bsif.gc.ca](http://www.osfi-bsif.gc.ca)) de plus amples informations au sujet du TCM et du MMPRCE.

**Solvabilité II.** Selon Solvabilité II, le capital de solvabilité requis se définit comme étant la VaR à 99,5 % sur un horizon d'un an. Lorsqu'ils agrègent les risques, les assureurs peuvent tirer avantage de la diversification : ils ont la possibilité d'utiliser un modèle interne (aucune méthode particulière prescrite) ou une formule standard. La formule standard agrège les

risques à l'aide d'une matrice de corrélations (méthode des variances-covariances) afin de tenir compte des dépendances.

**Dispositif suisse s'appliquant aux sociétés d'assurances.** Depuis 2008, tous les assureurs en Suisse doivent appliquer le Test de solvabilité suisse. À l'exemple de Solvabilité II, il existe un modèle standard et il est possible d'utiliser un modèle interne. Le modèle standard traite séparément les risques suivants : risque de marché, risque de crédit (défaut de la contrepartie), risque d'assurance de dommages, risque d'assurance-vie et risque d'assurance-maladie. À l'heure actuelle, les risques opérationnels ne font pas partie du Test. La diversification entre les catégories de risque est prise en compte dans tous les cas. Les sociétés d'assurance-vie utilisent la méthode d'agrégation variances-covariances, tandis que les assureurs IARD agrègent les risques avec plus de soin, cherchant la distribution du risque agrégé avant de calculer l'insuffisance attendue (ou la TVaR).

**Aux États-Unis, dispositif de solvabilité basé sur les risques d'assurance.** Nous terminons notre court survol des dispositifs de réglementation en usage dans le monde par celui, basé sur les risques (RBC), des États-Unis. La formule, standardisée, est valable dans tous les États, ce qui permet facilement de comparer les sociétés. Chaque type d'assureur (vie, IARD, maladie) applique une formule distincte. Les avantages de la diversification sont pris en compte en calculant une matrice de covariances entre les risques individuels afin de réduire le capital total et de le ramener à une valeur inférieure à celui correspondant à la somme des risques individuels.

Dans le calcul du capital basé sur les risques, la formule consiste en la racine carrée d'une somme de carrés, ce qui revient à appliquer une hypothèse très simple pour agréger les risques en supposant qu'ils sont parfaitement corrélés (corrélation de 1) ou indépendants (corrélation de 0) ([BSIF \[2014\]](#)).

### 3.2 Comparaison des dispositifs de réglementation internationaux et commentaires

En règle générale, les règlements prennent en compte la diversification en tenant compte des effets de la corrélation afin de réduire le capital total (du moins dans certaines sous-catégories). Globalement, nous constatons que les organismes de réglementation supposent tous implicitement que la somme des chiffres sur les risques correspond à la pire des situations. « Aucun avantage de la diversification » revient à « faire la somme des chiffres sur les risques (VaR) ».

La méthode la plus facile pour agréger les risques est celle des variances-covariances, qui est mentionnée explicitement dans Solvabilité II et dans le Test de solvabilité suisse et qui est aussi utilisée par l'organisme de réglementation australien ([BSIF \[2014\]](#)). Elle repose sur l'hypothèse que la matrice des corrélations suffit à décrire la dépendance et qu'il est possible d'agréger les risques en se fondant sur cette matrice. Sa force réside dans sa simplicité, mais elle n'est correcte que pour des distributions elliptiques multivariées telles que la distribution gaussienne multivariée. Par ailleurs, la corrélation étant une mesure linéaire de la dépendance, elle ne prend pas en compte adéquatement la dépendance de queue. L'utilisation d'une pareille méthode pour agréger les risques peut sans doute donner une idée « globale » de la distribution, mais elle ne permet pas d'évaluer le risque se situant dans la

queue de la distribution; on notera que les exigences en capital sont habituellement basées sur des mesures du risque se situant dans la queue, telles que la VaR à 99,5 %, qui reflète essentiellement un événement qui survient tous les 200 ans.

Plutôt que d'utiliser la méthode des variances-covariances pour agréger les risques individuels, on peut utiliser des copules, ce qui est assez flexible et permet de séparer l'évaluation des risques de la distribution marginale des risques individuels et de leur dépendance. En spécifiant une certaine copule pour modéliser la dépendance, il est ensuite possible de reconnaître la dépendance de queue entre les risques. Toutefois, le choix de la « bonne » copule est une tâche très difficile qui est sujette à un important risque de modélisation, comme nous le verrons plus loin. Les méthodes statistiques servant à ajuster un modèle multivarié font intervenir un grand nombre de paramètres et de familles de copules. De plus, l'interprétation des résultats du modèle exige une bonne connaissance de la théorie des copules pour pouvoir comprendre l'impact de chaque hypothèse faite au sujet de la dépendance. Pour les institutions, il s'agit là d'une chose préoccupante et difficile.

Une autre façon de prendre en compte les risques dans la queue et la dépendance de queue est de bien comprendre « d'où vient la dépendance », et de modéliser les déterminants réels de la dépendance existant entre les risques individuels du portefeuille et de comprendre leurs interactions. Dans son rapport, le [Comité de Bâle \[2010a\]](#) suggère d'agréger les effets de divers scénarios. Pour l'essentiel, cette façon de faire revient à déterminer la situation financière de la société sous l'effet de certains événements et à faire la somme des profits et pertes correspondant aux diverses positions. En d'autres termes, cela signifie qu'il faut incorporer des données connues au sujet de la dépendance dans des situations financières précises.

À l'annexe [B](#), nous proposons une méthode permettant d'évaluer le risque de modélisation, qui s'inscrit dans une certaine mesure dans cet esprit, car elle permet d'incorporer les données existantes sur la structure de dépendance entre les risques selon certaines conjonctures mondiales. L'analyse de scénarios présente un net avantage, en ce sens que le modèle multivarié est alors basé sur des déterminants du risque qui sont clairement identifiés (qui peuvent être simulés par exemple), et elle oblige la société à comprendre le modèle multivarié choisi; ce n'est plus un ensemble complexe de copules, mais la dépendance entre les facteurs est obtenue au moyen de facteurs raisonnables. Tel qu'il est indiqué dans le rapport du [Comité de Bâle \[2010a\]](#), les résultats de l'agrégation des effets des scénarios sont plus faciles à interpréter et ils ont des implications économiques et financières plus significatives, mais, ici encore, il faut avoir des connaissances approfondies pour pouvoir identifier les déterminants des risques et formuler des ensembles de scénarios cohérents avec des propriétés statistiques pertinentes, et savoir les utiliser pour obtenir une distribution complète des pertes sera quand même une tâche difficile. Un grand nombre des données d'entrée de ces sortes de modèles se fondent sur le jugement d'experts. Tout bien considéré, il n'existe pas de solution unique au problème de l'agrégation des risques. Chaque méthode comporte ses avantages et ses inconvénients et peut être utile dans certaines situations et inutile dans d'autres.

#### 4. Risque de modélisation de la dépendance

Comme nous l'avons dit à maintes reprises, l'un des principaux problèmes de l'agrégation des risques tient à la difficulté de modéliser la dépendance entre un grand nombre de risques, c.-à-d. que l'agrégation des risques est soumise au risque de modélisation. Plus particulièrement, nous avons montré, à la section [1](#), qu'il existait plus d'une façon de mesurer le risque, et à la section [2](#), que la corrélation ne suffisait pas à mesurer la dépendance et l'information complète sur la dépendance contenait beaucoup plus d'information. Toutefois, comme nous l'avons vu à la section [3](#), les organismes de réglementation de par le monde débattent des avantages de la diversification et proposent des lignes directrices pour les estimer. Mais cette question ne fait pas consensus. Il s'avère que la modélisation de la dépendance s'accompagne d'un important risque de modélisation.

En annexes, nous donnons des exemples spécifiques permettant de mieux comprendre le risque de modélisation lié à l'agrégation. L'annexe [A](#) présente une analyse de la façon de minimiser ou de maximiser une certaine mesure de risque  $\rho(\cdot)$  du risque agrégé lorsque les distributions des composantes de risque sont connues mais non leur interdépendance (en accord avec la méthode d'[Embrechts et al. \[2013\]](#)). Cette approche est utile pour évaluer le risque de modélisation de la dépendance, qui est l'un des plus importants facteurs dans l'évaluation du risque agrégé.

Or, les limites du risque de modélisation de la dépendance, obtenues par l'approche décrite à l'annexe [A](#) (cf. aussi [Embrechts et al. \[2013\]](#)), sont d'habitude trop larges pour avoir quelque utilité en pratique. Elles ne tiennent pas compte du tout de l'information sur la dépendance et ne considèrent que l'information sur les distributions marginales. Quelques études ont été faites sur le risque de modélisation avec de l'information partielle sur la structure de dépendance. Citons, entre autres, [Cheung et Vanduffel \[2013\]](#) qui ont étudié les limites d'ordre convexes avec une variance donnée; [Embrechts et Puccetti \[2006\]](#) qui ont analysé les limites de la distribution de  $S$  lorsque la copule de  $\mathbf{X}$  est bornée par une copule donnée; [Tankov \[2011\]](#) qui a étudié les limites de  $S$  lorsque  $n = 2$  et que la copule est soumise à des contraintes; [Bernard et al. \[2013b\]](#), lorsqu'une borne supérieure à la variance du risque agrégé est imposée, et [Bernard et al. \[2014a\]](#), lorsque des moments d'ordre supérieur sont donnés.

À l'annexe [B](#), nous présentons un cadre qui permet une quantification pratique du risque de modélisation, que [Bernard et Vanduffel \[2015a\]](#) ont mis au point. Fait à noter, à la différence de l'annexe [A](#), nous n'ignorons plus l'information disponible sur la dépendance. Nous supposons que les modélisateurs du risque ont créé un modèle multivarié « aussi bon que possible » pour un certain portefeuille. Toutefois, aucun modèle n'est parfait, et la mesure dans laquelle l'erreur de spécification du modèle proposé influe sur la mesure du risque devrait être évaluée. Notre cadre comporte un algorithme permettant aux actuaires de prendre en compte le risque de modélisation d'une manière très pratique.

Ces résultats rendent possible l'identification des mesures de risque pour lesquelles des informations additionnelles tirées d'un modèle multivarié bien ajusté réduisent le risque de modélisation de façon significative, faisant d'elles de bons choix à utiliser par les gestionnaires

du risque et les organismes de réglementation. Notre approche peut produire des limites qui sont considérablement plus étroites que celles présentées dans la littérature (non contraintes), en tenant compte de l'information disponible tirée d'un modèle multivarié ajusté et en permettant une évaluation plus réaliste du risque de modélisation. Toutefois, le risque de modélisation demeure une grande préoccupation et nous déconseillons la réglementation basée sur des VaR à des très hauts niveaux de confiance, vu que cette évaluation n'est pas en mesure de bénéficier des efforts déployés par les gestionnaires du risque pour ajuster un modèle multivarié. Par exemple, nous avons constaté, lors d'expériences numériques, que le portefeuille VaR à un très haut niveau de confiance (comme celui utilisé à l'heure actuelle dans le dispositif de réglementation de Bâle) était soumis à un fort risque de modélisation, dont l'intervalle des valeurs possibles resterait grand même si on connaissait presque parfaitement la distribution multivariée. En fait, on pourrait même ne pas être en mesure de réduire le risque de modélisation tel qu'[Embrechts et al. \[2013\]](#) l'ont calculé (cf. aussi à l'annexe [A](#)), lorsqu'aucune information sur la dépendance entre les risques n'est utilisée.

Nous faisons remarquer qu'il pourrait être intéressant de considérer une contrainte « globale » pour rétrécir davantage les limites. La variance est une statistique globale naturelle sur la distribution du risque agrégé et il serait relativement facile d'approfondir notre étude en appliquant des techniques semblables à celles utilisées dans [Bernard et al. \[2013b\]](#) pour prendre en compte une variance maximale possible du portefeuille agrégé.

Enfin, nous supposons que les distributions marginales sont fixes et connues. Afin de prendre en compte l'incertitude éventuelle des distributions marginales, nous pourrions amplifier leur queue. Par exemple, une distorsion (la transformée de Wang) pourrait être appliquée au moment de la nouvelle discrétisation (au lieu d'utiliser  $\hat{f}_i$ ).

**Superadditivité de la VaR :** Nous concluons cette section par la question des conséquences de l'agrégation. Plus particulièrement, nous traitons de la superadditivité de la VaR. La comonotonie est le pire des cas de dépendance, selon les décideurs qui ont une aversion pour le risque, mais cela ne donne pas la VaR maximale d'un portefeuille (on trouvera de plus amples détails à l'annexe [A](#)). La VaR correspondant au pire des cas ne se produit pas facilement lorsque les risques sont parfaitement corrélés. Étant donné que la VaR est additive pour les risques comonotones, il existe donc une dépendance telle que

$$VaR_p(X_1 + X_2 + \dots + X_n) \geq VaR_p(X_1) + VaR_p(X_2) + \dots + VaR_p(X_n) \quad (4.13)$$

L'absence d'avantages de la diversification est une situation que les professionnels en exercice acceptent difficilement. De plus, l'utilisation de la VaR peut conduire à un classement incohérent des risques, vu que la plus grande valeur possible de la mesure de risque ne correspond pas au scénario de la parfaite dépendance. Une question importante est celle où l'inégalité exprimée en [\(4.13\)](#) est stricte, c.-à-d. quand a-t-on la (stricte) superadditivité et dans quelle mesure la superadditivité est-elle importante? Il n'est pas difficile de montrer qu'on peut toujours trouver une dépendance telle que l'inégalité exprimée en [\(4.13\)](#) est stricte à moins que  $VaR_q(X_i)$  ne soit constante pour  $q \geq p$  (cf. aussi [Bernard, Rüschendorf et Vanduffel \[2013b\]](#)). Cette remarque nous permet de tirer les conclusions suivantes :

- Lorsque seules les distributions marginales sont connues et que le portefeuille contient des risques *non limités*, la VaR maximale possible (en trouvant la pire des dépendances possibles) peut être nettement supérieure à la VaR obtenue dans le cas monotone (dans lequel la VaR est additive). Par exemple, [Embrechts et al. \[2013\]](#) ont montré, dans leur figure 5, que pour un portefeuille de risques obéissant à la loi de Pareto(2), la borne supérieure de la VaR est environ deux fois plus grande que la VaR comonotone (lorsque les risques marginaux sont supposés être comonotones). Cf. aussi [Embrechts et al. \[2014\]](#). De manière plus générale, [Puccetti et Rüschendorf \[2012b\]](#) ont montré que, sous des conditions peu restrictives, la pire des VaR se comportait, de manière asymptotique, comme la pire des TVaR. Intuitivement, ce résultat s'explique de la manière suivante. La VaR (mesurée selon un certain niveau de probabilité  $p$ ) d'une somme comonotone n'est, bien entendu, qu'un point particulier de la fonction quantile de cette somme. Maintenant, si l'on modifie la dépendance comonotone dans la queue *supérieure* des supports marginaux (à partir du niveau  $p$ ), on est en mesure d'ajuster les quantiles supérieurs de la somme (à partir du niveau  $p$ ). Vu que la fonction quantile est non décroissante, on voit clairement que la plus grande des VaR peut s'obtenir en modifiant la dépendance de façon que la fonction quantile de la somme devienne constante sur  $(p, 1)$ . La valeur constante est donc la VaR maximale et elle est égale à la TVaR comonotone ([Bernard et al. \[2013b\]](#)).

Bien que, pour une  $p$  donnée, il y ait des structures de dépendance qui produisent une VaR supérieure à la VaR comotone, cela peut ne pas se produire dans la réalité.

- Vu que les sociétés d'assurances ont habituellement une responsabilité limitée, la VaR ne peut être (strictement) superadditive lorsque les niveaux de probabilité sont élevés (ce qui est généralement le cas dans le contexte des évaluations de solvabilité). En fait, dans ce cas, la VaR obtenue par l'utilisation d'un modèle particulier sera probablement sous-additive. Cette caractéristique revêt de l'importance, car la violation de la propriété de sous-additivité est une bonne raison de rejeter une mesure de risque, en particulier la VaR.
- La situation décrite précédemment met en évidence l'importance vitale d'avoir de l'information sur la dépendance si on veut pouvoir construire des modèles qui produisent des chiffres sur les risques qui sont fiables en ce sens que leurs limites supérieures et inférieures se situent toujours dans un intervalle raisonnable. Par exemple, on peut raisonnablement supposer que les risques sont positivement dépendants, ou que la variance du risque agrégé peut être estimée avec précision au moyen d'une analyse statistique des pertes observées, ou que de l'information sur la fonction copule est disponible. À cet égard, les résultats des études publiées sur les intervalles de la VaR en présence d'information additionnelle sur la dépendance sont plus limités et ponctuels. [Rüschendorf \[1991\]](#), [Embrechts et Puccetti \[2010a\]](#) et [Embrechts et al. \[2013\]](#) ont examiné la situation dans laquelle certaines des distributions bivariées sont connues, et [Denuit et al. \[1999\]](#) ont étudié les limites de la VaR en supposant que la distribution conjointe des risques est bornée par une certaine distribution. Toutefois, les limites proposées dans ces ouvrages sont souvent difficiles à

mettre en application, surtout dans le cas de portefeuilles non homogènes multidimensionnels, et ces limites ne permettent pas nécessairement de rétrécir de façon significative les limites non contraintes; pour une illustration dans le contexte de la modélisation d'un portefeuille de risques de crédit, cf. aussi [Chernih et al. \[2010\]](#). Toutefois, ces observations contrastent avec les constatations de [Bernard et al. \[2013b\]](#), qui considèrent la présence d'une contrainte de variance sur la somme du portefeuille comme étant une source d'information sur la dépendance, et ils montrent que cela permet de façon significative de rétrécir les limites non contraintes de la VaR.

Nous rappelons que la VaR est la mesure de risque la plus répandue dans les dispositifs de réglementation. Par exemple, en Europe, la réglementation actuelle des institutions financières (Bâle III) repose officiellement sur le concept d'actifs pondérés en fonction des risques, mais, pour l'essentiel, elle s'appuie sur la VaR. Il se peut donc qu'une approche basée sur les actifs pondérés en fonction des risques ne convienne pas s'il faut agréger les risques pour calculer la VaR d'un portefeuille. La majorité des chercheurs se sont toujours opposés à l'utilisation de la VaR, car celle-ci ne vérifie pas la condition de sous-additivité. Récemment, on a tendance à vouloir abandonner la VaR au profit de la TVaR; cf. [Embrechts et al. \[2014\]](#), [Comité de Bâle \[2012\]](#) et [Comité de Bâle \[2013\]](#).

## 5. Conclusions

Les récentes turbulences, telles que la crise des prêts hypothécaires à risque, ont incité les organismes de réglementation et les institutions financières à revoir attentivement leurs modèles de risques et à mieux comprendre dans quelle mesure les résultats des évaluations des risques, basées sur ces modèles, résistent aux changements dans les hypothèses sous-jacentes.

C'est pourquoi nous avons observé dernièrement la publication d'importants ouvrages traitant de l'agrégation des risques et des avantages de la diversification. Il faut s'attendre à ce que l'on propose de nouvelles approches concernant l'agrégation des risques et à ce que la question du risque de modélisation, inhérent à l'agrégation des risques, fasse l'objet d'un grand nombre d'études.

À la section [4](#), nous avons analysé brièvement les derniers développements touchant l'évaluation du risque de modélisation. À l'annexe [B](#), nous exposons une méthode pratique d'évaluation du risque de modélisation qui tient compte d'un ensemble type d'informations disponibles. Ces informations peuvent être tirées d'une modélisation statistique, telle qu'un modèle multivarié ajusté aux données en main (et on peut en avoir confiance chaque fois que les données sont en quantité suffisante), mais elles peuvent aussi être tirées de scénarios ou d'opinions d'experts. On suppose qu'on a certaines informations sur les scénarios extrêmes. Par exemple, on suppose que lorsqu'un gros réassureur fait faillite, les assureurs qui sont réassurés par lui enregistreront des pertes, et ce, simultanément (ce qui correspond à une situation comonotone dans la queue). Si de telles informations sont disponibles, elles peuvent être intégrées et il peut bien être possible de réduire les limites de la VaR à des niveaux élevés.

## Annexe A – Risque de modélisation de la dépendance au moment de l'agrégation des risques

À l'occasion du Symposium sur la GRE de 2014, les chercheurs ont reçu un prix de la PRMIA pour leur document intitulé [A New Approach to Assessing Model Risk in High Dimensions](#) (9 février 2014). Ce document propose une démarche pour évaluer le risque de modélisation, qui figure également dans notre document. En outre, des documents semblables ont été diffusés dans le [Journal of Banking and Finance](#) et dans *Dependence Modeling*. Les annexes [A](#) et [B](#) discutent des approches d'évaluation du risque de modélisation qui sont abordées dans ces documents.

L'un des principaux problèmes de l'agrégation des risques tient à la difficulté de modéliser la dépendance entre un grand nombre de risques, c.-à-d. que l'agrégation des risques est soumise au risque de modélisation. Nous analysons ici la façon de minimiser ou de maximiser une mesure  $\varrho(\cdot)$  du risque agrégé lorsque les distributions des composantes de risque sont connues mais non leur interdépendance (en accord avec la méthode d'[Embrechts et al. \[2013\]](#)). Dans la prochaine section, nous ferons le même exercice mais cette fois en supposant que des informations additionnelles sur la dépendance sont disponibles (suivant la méthode proposée récemment par [Bernard et Vanduffel \[2015a\]](#)).

Ci-après,  $\mathbf{X} = (X_1, X_2, \dots, X_d)$  représente le portefeuille considéré avec les distributions marginales données  $F_{X_i}$  ( $i = 1, 2, \dots, d$ ), et nous nous intéressons aux propriétés de  $\varrho(S)$ , où

$S = \sum_{i=1}^d X_i$ . Par commodité, nous supposons que toutes les moyennes sont finies.

L'égalité (2.12) nous rappelle que les propriétés distributionnelles du portefeuille  $\mathbf{X}$  sont complètement spécifiées si l'on sait aussi quelle est la copule qui décrit l'interaction entre les risques du portefeuille. Dans ce cas, la distribution multivariée de  $\mathbf{X}$  est connue et, à l'évidence, il existe une seule valeur possible de  $\varrho(S)$ . Toutefois, lorsque la structure de dépendance n'est pas spécifiée,  $\varrho(S)$  peut prendre un éventail de valeurs possibles selon la structure de dépendance choisie. Nous cherchons à trouver les valeurs minimale et maximale possibles de  $\varrho(S)$  qui reflètent le degré de risque de modélisation. Intuitivement, lorsque la dépendance est forte,  $S$  devient un risque « plus variable » et  $\varrho(S)$  devrait atteindre sa valeur maximale. Réciproquement, s'il y a beaucoup de compensation entre les risques,  $\varrho(S)$  devrait être petite. L'un des moyens bien connus de décrire la variabilité des risques est ce que l'on appelle l'ordre convexe. En termes mathématiques, l'ordre convexe<sup>10</sup>,  $\leq_{\text{cx}}$ , entre les variables aléatoires  $X, Y$  se définit comme suit

$$X \leq_{\text{cx}} Y \quad \text{si } E(f(X)) \leq E(f(Y))$$

pour toutes les fonctions convexes  $f(\cdot)$  pour lesquelles l'espérance existe. À noter que

$$E(X) \leq_{\text{cx}} X \tag{A.14}$$

<sup>10</sup> Pour de plus amples informations sur cet ordre dans le contexte de l'actuariat, cf. par exemple [Müller et Stoyan \[2002\]](#), [Denuit et al. \[2005\]](#), [Denuit et al. \[1999\]](#) et [Dhaene et al. \[2002\]](#).

et aussi que  $X \leq_{\text{cx}} Y$  implique que  $X$  et  $Y$  ont la même moyenne mais que  $Y$  a la plus grande variance. L'ordre convexe correspond bien aux préférences des investisseurs qui ont une aversion pour le risque et elle est très utile pour quantifier l'incertitude au sujet de  $\varrho(S)$ . Plus précisément, lorsque la mesure de risque  $\varrho(\cdot)$  est compatible avec l'ordre convexe, les limites de l'ordre convexe se transforment en limites de la mesure de risque<sup>11</sup>. C'est le cas de la variance ou de la TVaR, par exemple<sup>12</sup>. Pour ce qui est de la VaR, cette mesure de risque n'est pas en soi compatible avec l'ordre convexe, mais il existe néanmoins des liens étroits entre les limites de la VaR et celles de l'ordre convexe, comme nous l'expliquerons plus loin (cf. aussi [Bernard et al. \[2013b\]](#)). Dans tous les cas, il est donc important de déterminer les limites convexes supérieures et inférieures des sommes de risques.

### A.1 Limites convexes supérieures et inférieures

La limite convexe supérieure d'un nombre général  $d$  de risques individuels est atteinte lorsque les risques affichent une dépendance maximale (risques comonotones), ce qui est une structure de dépendance facile à décrire. Plus précisément, dans le cas comonotone on considère en fait

$$\mathbf{X} = (F_{X_1}^{-1}(U_1), F_{X_2}^{-1}(U_2), \dots, F_{X_n}^{-1}(U_n)), \quad (\text{A.15})$$

dans laquelle maintenant

$$U_1 = U_2 = \dots = U_n := U. \quad (\text{A.16})$$

Intuitivement, il est clair que les variables  $X_i = F_i^{-1}(U)$  sont parfaitement dépendantes, car elles augmentent de façon maximale l'une dans l'autre. Nous obtenons donc que pour n'importe quelle somme de portefeuille  $S := \sum_i X_i$  dans laquelle les composantes de risque  $X_i$  sont distribuées en fonction de  $F_i$ ,

$$E(S) \leq_{\text{cx}} S \leq_{\text{cx}} \sum_{i=1}^n F_i^{-1}(U) \quad (\text{A.17})$$

Ce résultat a été prouvé dans de nombreux ouvrages (en particulier la deuxième inégalité), premièrement dans [Meilijson et Nádas \[1979\]](#) et puis dans [Rüschendorf \[1982\]](#).

Si la limite convexe supérieure est facile à atteindre, la limite convexe inférieure énoncée précédemment,  $E(S)$ , n'est pas atteignable (*sharp*) en général. De fait, l'obtention de limites convexes inférieures qui sont *sharp* est un problème très difficile, en particulier lorsque le nombre de dimensions est élevé<sup>13</sup>. Néanmoins, nous montrons ci-après qu'il existe un

<sup>11</sup> L'ordre convexe est un ordre naturel dans la catégorie des *risques admissibles*. [Bernard, Jiang et Wang \[2014b\]](#) ont introduit le concept de *risque admissible* pour décrire tout le risque agrégé possible  $S$  avec des distributions marginales connues mais une structure de dépendance inconnue.

<sup>12</sup> Toutes les mesures concaves du risque de distorsion sont compatibles avec l'ordre convexe.

<sup>13</sup> Pour  $d = 2$ , [Denuit et al. \[1999\]](#) et [Tankov \[2011\]](#) ont obtenu une limite convexe inférieure pour  $X_1 = F_1^{-1}(U)$  et  $X_2 = F_2^{-1}(1-U)$ ; pour  $d \geq 3$ , cf. [Bernard et al. \[2014b\]](#), qui ont obtenu certains résultats. De fait, l'existence d'une limite inférieure dite *sharp* est liée de près à la notion de *complete mixability* ([Wang et Wang \[2011\]](#)) comme

algorithme qui permet (du moins dans le cas des portefeuilles dont la taille est de moyenne à grande, soit ceux qui nous intéressent) de trouver une dépendance entre les risques, de façon que la somme  $S$  se comporte approximativement comme la constante  $E(S)$ . En d'autres termes, l'algorithme permet d'obtenir des limites convexes inférieures approximatives de  $S$ . Ensuite, nous expliquons la façon de trouver des limites de risque maximale et minimale pour les portefeuilles, lorsque nous employons la variance et la VaR comme mesure de risque.

## A.2 Algorithme de réarrangement

L'algorithme de réarrangement de [Puccetti et Rüschendorf \[2012a\]](#), qu'[Embrechts et al. \[2013\]](#) ont développé davantage, se veut une méthode pratique de construire la dépendance entre les variables  $X_j$  ( $j = 1, 2, \dots, d$ ) de façon que la somme du portefeuille  $S = X_1 + \dots + X_d$  devienne aussi petite que possible dans l'ordre convexe. Nous rappelons que cet algorithme est très utile pour trouver les limites minimales de la variance et de la TVaR (les limites maximales sont faciles à trouver et découlent de la comonotonie dans ce cas), et il s'avère tout aussi utile pour trouver les limites de la VaR, bien que celle-ci ne satisfait pas à l'ordre convexe.

Sans perte de généralité (pratique), nous supposons que les variables  $X_j$  sont discrétisées et qu'elles prennent  $n$  valeurs qui sont mises aléatoirement<sup>14</sup> dans une matrice  $\mathbf{A}$  :

$$\mathbf{A} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1d} \\ x_{21} & x_{22} & \dots & x_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nd} \end{bmatrix}. \quad (\text{A.18})$$

La matrice  $\mathbf{A}$  peut être vue comme une représentation d'une structure multivariée possible pour  $\mathbf{X} = (X_1, X_2, \dots, X_d)$ . À noter que nous ne modifions pas les distributions marginales respectives des  $X_j$  ( $j = 1, 2, \dots, d$ ) en réarrangeant les résultats d'une colonne, mais uniquement la dépendance entre les  $X_j$ .

1. Pour  $i$  allant de 1 à  $d$ , rendre la  $i^{\text{e}}$  colonne antimonotone avec la somme des autres colonnes.
2. Recommencer à partir de la colonne 1 et la rendre antimonotone avec les sommes des colonnes 2 à  $d$ .

À chaque étape de cet algorithme, nous rendons la  $j^{\text{e}}$  colonne antimonotone avec la somme des autres, de façon que les colonnes, disons  $X_j$  avant le réarrangement et  $\tilde{X}_j$  après le réarrangement, vérifient évidemment

---

nous l'expliquons plus loin; cf. aussi [Dhaene et al. \[2002\]](#), [Wang et Wang \[2011\]](#), [Embrechts et al. \[2013\]](#) et [Wang et al. \[2013\]](#) pour obtenir des informations de base et des résultats mathématiques.

<sup>14</sup> Par exemple, nous pouvons mettre dans chaque colonne de la matrice  $\mathbf{A}$  les éléments en ordre croissant, auquel cas nous avons une structure comonotone comme situation de départ (ce qui produit la plus grande somme possible dans l'ordre convexe).

$$\text{var}\left(\sum_{i=1}^d X_i\right) \geq \text{var}\left(\tilde{X}_j + \sum_{i \neq j} X_i\right).$$

De fait,

$$\text{var}\left(\sum_{i=1}^d X_i\right) = \text{var}\left(X_j + \sum_{i \neq j} X_i\right)$$

et son minimum lorsque  $X_j$  est antimonotone avec  $\sum_{i \neq j} X_i$ . À chaque étape de l'algorithme, la variance diminue<sup>15</sup> et elle est bornée inférieurement (par 0) et converge donc vers une limite  $\ell \geq 0$  (convergence d'une suite monotone de nombres réels). Si la variance devient nulle, nous avons une situation de parfaite *mixability*, c.-à-d. que la dépendance est telle que la somme devient une constante et est donc aussi convexe que possible (cf. (A.14)). Sinon, l'algorithme converge vers un minimum local. Il n'y a ensuite aucune garantie que ce minimum soit réellement le minimum de la variance de la somme optimisée sur toute la structure de dépendance, car ce minimum peut dépendre du point de départ. Toutefois, en pratique, il s'avère que la convergence est très rapide et l'on parvient habituellement à approcher la situation de *complete mixability* au bout de quelques itérations (sauf lorsque la taille du portefeuille est très petite). En particulier, l'algorithme donne d'excellents résultats dans le cas d'un portefeuille homogène (dans lequel toutes les  $X_j$  ont la même distribution).

Remarque A.1 L'algorithme décrit précédemment s'arrête toujours dans une situation où chaque colonne est antimonotone avec les autres<sup>16</sup>.

### A.3 Exemple d'application de l'algorithme de réarrangement

Afin d'illustrer l'algorithme décrit précédemment, nous présentons le cas très simple d'une matrice contenant huit lignes et trois colonnes (c.-à-d. que nous considérons un portefeuille contenant trois risques qui prennent diverses valeurs selon huit scénarios), qui est similaire à la matrice générale présentée en A.18.

<sup>15</sup> On notera que la situation dans laquelle toutes les colonnes sont antimonotones avec la somme de toutes les autres est une condition nécessaire et évidente de l'obtention d'une structure de dépendance qui minimise la variance.

<sup>16</sup> À chaque étape de l'algorithme, si une colonne n'est pas antimonotone avec la somme des autres, elle est réarrangée de manière à ce qu'elle le devienne, ce qui implique que la variance diminue de façon stricte (car l'antimonotonie est la seule structure de dépendance qui atteint la variance minimale). La matrice étant de taille finie, il y a donc un nombre fini de réarrangements possibles de cette matrice et la variance peut donc diminuer strictement un nombre fini de fois. Si, à partir d'un point dans chaque colonne, la variance ne change pas, cela veut dire que chaque colonne est antimonotone avec la somme des autres et que l'algorithme s'est donc arrêté.

$$\begin{bmatrix} 3 & 4 & 1 \\ 2 & 1 & 1 \\ 0 & 3 & 2 \\ 1 & 2 & 1 \\ 0 & 4 & 2 \\ 1 & 0 & 1 \\ 3 & 1 & 2 \\ 4 & 2 & 3 \end{bmatrix}. \quad (\text{A.19})$$

Nous débutons ici à partir de la structure comonotone et nous appliquons l'algorithme de réarrangement séquentiellement tel qu'il est décrit précédemment et nous obtenons ceci (c.-à-d. en appliquant séquentiellement l'algorithme sur la première, la deuxième et la troisième colonne) :

$$\begin{bmatrix} 3 & 4 & 1 \\ 2 & 1 & 1 \\ 0 & 3 & 2 \\ 1 & 2 & 1 \\ 0 & 4 & 2 \\ 1 & 0 & 1 \\ 3 & 1 & 2 \\ 4 & 2 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 4 & 1 \\ 3 & 1 & 1 \\ 0 & 3 & 2 \\ 3 & 2 & 1 \\ 0 & 4 & 2 \\ 4 & 0 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 4 & 1 \\ 3 & 2 & 1 \\ 0 & 4 & 2 \\ 3 & 2 & 1 \\ 0 & 3 & 2 \\ 4 & 0 & 1 \\ 2 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 4 & 1 \\ 3 & 2 & 1 \\ 0 & 4 & 1 \\ 3 & 2 & 1 \\ 0 & 3 & 2 \\ 4 & 0 & 2 \\ 2 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix}. \quad (\text{A.20})$$

À noter que, dans la dernière matrice, chaque colonne est antimonotone avec la somme des deux autres.

#### A.4 La variance et le risque de modélisation de la dépendance

**Proposition A.2** (Limites de la variance de  $\sum_{i=1}^d X_i$ ). Soit  $(X_1, X_2, \dots, X_d)$  un portefeuille avec distribution marginale respective  $F_i$ . Posons  $S = \sum_{i=1}^d X_i$ . Nous avons :

$$\text{var}(E(S)) = 0 \leq \text{var}(S) \leq \text{var}\left(\sum_{i=1}^d F_i^{-1}(U)\right),$$

dans laquelle  $U$  est une variable aléatoire distribuée uniformément sur  $(0, 1)$ .

La proposition [A.2](#) découle directement du fait que la variance est compatible avec l'ordre convexe et la relation d'ordre convexe [A.17](#). La limite inférieure que nous proposons ici correspond donc au cas dans lequel la somme du portefeuille est constante, c.-à-d. qu'il s'agit de la situation de *complete mixability* évoquée dans [Wang et Wang \[2011\]](#). Dans ce cas, nous disons que la limite inférieure est *sharp*, car il existe une structure de dépendance entre les risques  $X_i$  qui fait en sorte que la somme soit constante et affiche exactement une variance

nulle. Comme nous l'avons expliqué précédemment, l'algorithme de réarrangement cherche à atteindre cette situation bien que ce ne soit pas toujours possible, auquel cas la limite inférieure indiquée à la proposition [A.2](#) n'est pas *sharp*. Dans tous les cas, l'algorithme peut être vu comme un moyen d'estimer approximativement la limite convexe inférieure *sharp*.

Illustrons ces limites au moyen de l'exemple des huit observations présenté précédemment. La variance maximale s'obtient lorsque les risques affichent une dépendance comonotone (cf. [A.19](#)) et nous avons

$$S := \begin{bmatrix} 4 & 4 & 3 \\ 3 & 4 & 2 \\ 3 & 3 & 2 \\ 2 & 2 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 11 \\ 9 \\ 8 \\ 6 \\ 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}.$$

Pour ce qui est de la variance minimale, après avoir appliqué l'algorithme de réarrangement en [A.20](#), nous obtenons comme résultat

$$S := \begin{bmatrix} 1 & 4 & 1 \\ 3 & 2 & 1 \\ 0 & 4 & 1 \\ 3 & 2 & 1 \\ 0 & 3 & 2 \\ 4 & 0 & 2 \\ 2 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 6 \\ 5 \\ 6 \\ 5 \\ 6 \\ 5 \\ 5 \end{bmatrix}.$$

Nous constatons que la limite inférieure dans la proposition [A.2](#) n'est pas atteinte dans ce cas particulier, c.-à-d. qu'il n'existe aucune structure de dépendance entre  $X_1$ ,  $X_2$  et  $X_3$  qui fait en sorte que la somme soit constante. Toutefois, le résultat de l'algorithme de réarrangement demeure néanmoins une très bonne approximation de la somme qui est la plus petite possible du point de vue de l'ordre convexe. En d'autres termes, l'algorithme permet d'approcher les limites inférieures (*sharp*) de la variance.

## A.5 La VaR et le risque de modélisation de la dépendance

Étant donné que la comonotonie est le pire des cas de dépendance selon les décideurs qui ont une aversion pour le risque, on pourrait penser, intuitivement, que cette dépendance produit la VaR maximale d'un portefeuille, comme c'est le cas de la variance. Or, nous verrons que cette

intuition est fautive en règle générale. Nous notons tout d'abord que, pour toute somme  $S = \sum_{i=1}^d (X_i)$  et  $0 < p < 1$ ,

$$\text{VaR}_p(S) \leq \text{TVaR}_p(S) \quad (\text{A.21})$$

$$\leq B = \text{TVaR}_p\left(\sum_{i=1}^d F_i^{-1}(U)\right) \quad (\text{A.22})$$

De même, on constate que

$$A := \text{LTVaR}_p\left(\sum_{i=1}^d F_i^{-1}(U)\right) \leq \text{VaR}_p(S) \quad (\text{A.23})$$

où nous avons défini la TVaR à gauche (LTVaR) au niveau  $p$  ( $0 < p < 1$ ) comme étant

$$\text{LTVaR}_p(X_i) = \frac{1}{p} \int_0^p \text{VaR}_u[X_i] du. \quad (\text{A.24})$$

À noter que pour TVaR et LTVaR,

$$(\text{L})\text{TVaR}_p\left(\sum_{i=1}^d F_i^{-1}(U)\right) = \sum_i (\text{L})\text{TVaR}_p(F_i^{-1}(U)) \quad (\text{A.25})$$

En résumé, nous obtenons le résultat suivant.

**Théorème A.3** (Limites de la VaR de  $\sum_{i=1}^d X_i$ ). Soit  $(X_1, X_2, \dots, X_d)$  un portefeuille avec distribution marginale respective  $F_i$ . Posons  $S = \sum_{i=1}^d X_i$  et  $p \in (0, 1)$ . Alors,

$$\sum_{i=1}^n \text{LTVaR}_q(F_i^{-1}(U)) \leq \text{VaR}_q\left(\sum_{i=1}^d X_i\right) \leq \sum_{i=1}^n \text{TVaR}_q(F_i^{-1}(U)). \quad (\text{A.26})$$

Ces limites sont données et prouvées dans [Bernard et al. \[2013b\]](#). La question est donc de savoir si ces limites sont *sharp*. Pour résoudre ce problème, notons que

$$\text{VaR}_p\left(\sum_{i=1}^d F_i^{-1}(U)\right) \leq \text{TVaR}_p\left(\sum_{i=1}^d F_i^{-1}(U)\right). \quad (\text{A.27})$$

Ainsi, afin d'atteindre la limite supérieure  $B$ , l'idée est donc de débiter avec la dépendance comotone et puis de la modifier de façon que l'inégalité (A.26) devienne une égalité. Vu que  $\text{TVaR}_p$  est la moyenne de toutes les  $\text{VaR}_q$  supérieures sur l'intervalle  $[p, 1]$ , il est clair que l'égalité s'obtient si la VaR de la somme comotone  $\sum_{i=1}^d F_i^{-1}(U)$  devient constante sur  $[p, 1]$  (en modifiant cette dépendance comotone). Soit  $G_i$  la distribution de  $F_i$  lorsqu'elle est restreinte<sup>17</sup> à la partie  $p$  supérieure de  $F_i$ . Pour atteindre la limite supérieure, il faut donc

<sup>17</sup> Formellement,  $G_i$  est la distribution de  $F_i^{-1}(V)$ , où  $V$  est distribuée uniformément sur  $[q, 1]$ .

trouver une dépendance entre les risques (maintenant avec distributions marginales  $G_i$ ) telle que la somme correspondante devient constante (c.-à-d. que les risques sont *completely mixing*). En général, la propriété de *mixing* ne se vérifie pas et les limites indiquées ne sont donc pas *sharp*. Cependant, il est maintenant clair qu'on peut obtenir (des approximations) des limites *sharp* de la VaR en trouvant une dépendance entre les risques (avec distributions marginales  $G_i$ ) de façon que la somme correspondante devienne aussi petite et convexe que possible (cf. aussi [Bernard et al. \[2013b\]](#)). Un raisonnement similaire montre que, pour pouvoir atteindre la limite inférieure indiquée aussi près que possible, il faudrait modifier la dépendance comonotone de façon que la fonction quantile de la somme du portefeuille comonotone devienne aussi plate que possible sur l'intervalle  $[0, p]$ .

Nous nous fondons sur cette idée pour proposer un algorithme pratique permettant d'estimer approximativement les limites *sharp*. Nous montrons donc comment approcher les limites *sharp* à l'aide de l'exemple discret utilisé précédemment, lorsque le niveau  $p$  utilisé pour évaluer la VaR est de  $5/8$ . À noter que nous débutons par la structure comonotone dans la matrice.

$$\begin{bmatrix} 4 & 4 & 3 \\ 3 & 4 & 2 \\ 3 & 3 & 2 \\ 2 & 2 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 4 & 3 \\ 3 & 4 & 2 \\ 3 & 3 & 2 \end{bmatrix}.$$

Nous appliquons l'algorithme de réarrangement dans les trois lignes correspondantes.

$$\begin{bmatrix} 4 & 4 & 3 \\ 3 & 4 & 2 \\ 3 & 3 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 3 & 3 \\ 3 & 4 & 2 \\ 3 & 4 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 3 & 2 \\ 3 & 4 & 2 \\ 3 & 4 & 3 \end{bmatrix}.$$

Les sommes sont donc respectivement égales à 9, 9 et 10 et la VaR maximale est de 9. Pour obtenir la VaR minimale, on se sert des faibles valeurs de chaque  $X_i$  puis on applique l'algorithme à ces valeurs.

$$\begin{bmatrix} 4 & 4 & 3 \\ 3 & 4 & 2 \\ 3 & 3 & 2 \\ 2 & 2 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 2 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

L'application de l'algorithme donne

$$\begin{bmatrix} 2 & 2 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 2 & 1 \end{bmatrix}.$$

Les valeurs des sommes sont donc égales à 4, 3, 3, 3 et 3. La VaR minimale est donc égale à 4.

## Annexe B – Risque de modélisation de la dépendance lors de l’agrégation des risques et quantification du risque de modélisation

Nous présentons ici un cadre créé par [Bernard et Vanduffel \[2015a\]](#) et qui permet une quantification pratique du risque de modélisation. Nous supposons que les modélisateurs du risque ont créé un modèle multivarié « aussi bon que possible » pour un certain portefeuille  $(X_1, X_2, \dots, X_d)$ . Toutefois, aucun modèle n’étant parfait, nous voulons évaluer dans quelle mesure l’erreur de spécification du modèle proposé influe sur la mesure du risque de  $S = \sum_i X_i$ . Fait à noter, à la différence de l’annexe [A](#), nous n’ignorons plus l’information disponible sur la dépendance. Notre cadre comporte un algorithme permettant aux actuaires de prendre en compte le risque de modélisation d’une manière très pratique, comme nous le verrons en détail.

Ces résultats rendent possible l’identification des mesures de risque pour lesquelles des informations additionnelles tirées d’un modèle multivarié bien ajusté réduisent le risque de modélisation de façon significative, faisant d’elles de bons choix à utiliser par les gestionnaires du risque et les organismes de réglementation. Par exemple, nous avons constaté, lors d’expériences numériques, que le portefeuille VaR à un très haut niveau de confiance (comme celui utilisé à l’heure actuelle dans le dispositif de réglementation de Bâle) était soumis à un fort risque de modélisation, dont l’intervalle des valeurs possibles resterait grand même si on connaissait presque parfaitement la distribution multivariée. En fait, on pourrait même ne pas être en mesure de réduire le risque de modélisation tel qu’[Embrechts et al. \[2013\]](#) l’ont calculé (cf. aussi à l’annexe [A](#)), lorsqu’aucune information sur la dépendance entre les risques n’est utilisée.

L’idée sur laquelle s’appuie notre approche est intuitive et correspond à des situations réelles. Supposons que nous avons  $N$  vecteurs d’observations à  $d$  dimensions,  $\{(x_{1i}, \dots, x_{di})\}_{i=1, \dots, N}$  et qu’un modèle multivarié a déjà été ajusté à cet ensemble de données. En d’autres termes, il y a une distribution conjointe de  $(X_1, X_2, \dots, X_d)$  qui est disponible (modèle de référence).

Toutefois, nous savons que le modèle est soumis à une erreur de spécification, surtout en raison du manque de données. Nous fractionnons donc  $\mathbb{R}^d$  en deux sous-ensembles :  $\mathcal{F}$  est désigné sous le nom de région « fixe » ou « fiable », et  $\mathcal{U}$  porte le nom de région « non fixe » ou « non fiable ».  $\mathcal{U}$  représente la région dans laquelle les données ne sont pas considérées comme étant assez fiables (riches) pour conclure que l’ajustement du modèle (dans cette région) est de qualité. À noter que

$$\mathbb{R}^d = \mathcal{F} \cup \mathcal{U}.$$

Si l’on a parfaitement confiance dans le modèle, toutes les observations se trouvent dans la partie « fiable » ( $\mathcal{U}=\emptyset$ ) et le risque de modélisation est nul. Par contre,  $\mathcal{F} = \emptyset$  si on n’a aucune confiance dans l’ajustement de la dépendance, ce qui correspond au cas étudié par [Embrechts et al. \[2013\]](#) (cf. aussi l’annexe [A](#)).

Un problème connexe a déjà été étudié pour les portefeuilles bidimensionnels ( $d = 2$ ) lorsque certaines informations sur la dépendance (copule) sont disponibles; cf. par exemple [Tankov](#)

[2011], [Bernard et al. \[2012\]](#) et [Bernard et al. \[2013a\]](#). [Tankov \[2011\]](#) a recours à des scénarios de dépendance extrêmes pour trouver des limites non basées sur un modèle pour les prix de certains dérivés bivariés, tandis que [Bernard et al. \[2014\]](#) et [Bernard et al. \[2014\]](#)<sup>18</sup> utilisent ces scénarios pour déterminer des stratégies d'investissement optimales pour les investisseurs soumis à des contraintes qui sont fonction de la conjoncture. Bien que ces deux applications puissent être utiles pour trouver les limites de copules dans le cas bivarié, la gestion des risques fait intervenir habituellement plus de deux risques. Malheureusement, la recherche de limites pour des copules dans le cas général à  $d$  dimensions et en présence de contraintes est non seulement plus difficile, mais aussi moins utile pour les applications en gestion des risques. La raison en est que lorsque  $d > 2$ , dans la plupart des cas, la pire copule (sous contraintes) d'un vecteur  $(X_1, X_2, \dots, X_d)$  ne donne pas lieu à la plus grande valeur possible de la mesure de risque considérée de  $S = \sum_{i=1}^d X_i$ , du fait que les distributions marginales ont aussi un impact; cf. par exemple [Bernard et al. \[2014b\]](#) pour des illustrations.

Qui plus est, il existe très peu d'ouvrages qui traitent explicitement des problèmes multidimensionnels et de l'existence d'informations (partielles) sur la structure de dépendance : [Rüschendorf \[1991\]](#), [Embrechts et Puccetti \[2010b\]](#) et [Embrechts et al. \[2013\]](#) ont examiné la situation dans laquelle certaines des distributions bivariées sont connues; [Denuit et al. \[1999\]](#) ont étudié les limites de la VaR en supposant que la distribution conjointe des risques est bornée par une distribution, et [Bernard et al. \[2013b\]](#) ont calculé des limites de la VaR lorsque la variance de la somme est connue. Ils ont été en mesure de montrer comment l'information sur la variance du risque agrégé permettait de réduire considérablement la distance entre la VaR maximale et la VaR minimale. De plus, ils ont proposé un algorithme (une extension de l'algorithme de réarrangement) que les actuaires peuvent utiliser de manière très pratique pour évaluer les limites de la VaR (et ainsi le risque de modélisation lié à l'évaluation de la VaR) lorsque seule la variance du risque agrégé est connue.

Le contexte de tous ces ouvrages est parfois difficile à rapprocher avec l'information disponible en pratique; ou encore il ne fait pas usage de toutes les informations disponibles sur la dépendance. De plus, bien que les limites proposées dans ces ouvrages soient *sharp* (atteignables), elles sont souvent difficiles à calculer numériquement, surtout lorsqu'il y a un grand nombre de dimensions et que les risques ne sont pas homogènes. On notera aussi que les limites obtenues ne permettent pas toujours de resserrer de manière significative les limites non contraintes, ce qui laisse penser qu'il faut avoir des informations additionnelles sur la dépendance pour pouvoir obtenir de meilleures limites; pour des résultats connexes, cf. aussi [Wang et Wang \[2011\]](#), [Embrechts et al. \[2013\]](#), [Wang et al. \[2013\]](#) et [Bernard et al. \[2014b\]](#).

Dans cette dernière section, nous étudions donc les limites des mesures du risque agrégé  $S$ , en utilisant l'information sur la distribution conjointe *multivariée* de ses composantes  $X_i$  (qui renferme l'information sur la dépendance) plutôt que l'information sur la copule. Nous proposons deux méthodes pour calculer les limites des mesures du risque. La première méthode pourrait être non paramétrique, elle s'appuie sur l'algorithme de réarrangement

<sup>18</sup> Cet ouvrage a obtenu de la SOA le prix Redington 2015.

d'[Embrechts et al. \[2013\]](#) et elle permet de trouver des limites *sharp*. La seconde méthode produit des limites simples, analytiques, mais qui ne sont pas forcément *sharp*. En pratique, nous montrons aussi à l'aide d'exemples que les limites analytiques ne sont pas loin d'être *sharp*.

Le fil de cette section se déroule comme suit. Tout d'abord, nous présentons la méthode pratique en utilisant l'algorithme de réarrangement (section [B.3](#)). Cette méthode peut être appliquée directement à l'aide des données en main (sans ajuster un modèle), en sorte que dans ce cas, le risque de modélisation peut être évalué de manière complètement non paramétrique. Cette méthode s'appuie sur l'algorithme de réarrangement créé récemment par [Puccetti et Rüschendorf \[2012a\]](#) et étudié de manière approfondie par [Embrechts et al. \[2013\]](#). Elle a recours à une version discrétisée du problème décrit précédemment et elle utilise une représentation matricielle pour estimer approximativement les structures de dépendance correspondant au pire des cas. Ensuite, à la section [B.4](#), nous présentons la forme analytique et nous l'illustrons au moyen de simulations de risques qui suivent une  $N(0, 1)$  et de risques qui obéissent à une loi de Pareto. Nous donnons des limites qui peuvent être calculées directement (en utilisant, par exemple, des simulations de Monte Carlo) mais qui ne sont pas forcément *sharp*. Nos résultats numériques indiquent que, dans un espace multidimensionnel, les limites calculées selon la méthode directe de la section [B.4](#) sont près des limites non paramétriques calculées à la section [B.3](#). Autrement dit, si la recherche de limites *sharp* constitue théoriquement un problème difficile, les illustrations numériques suggèrent que l'algorithme que nous proposons à la section [B.4](#) conduit à l'obtention de limites presque *sharp*. Les résultats numériques montrent aussi que les nouvelles limites donnent habituellement de meilleurs résultats que les limites non contraintes déjà disponibles dans la littérature et ils permettent donc une évaluation plus réaliste du risque de modélisation. Cela dit, le risque de modélisation demeure une question préoccupante, surtout lorsqu'on utilise une mesure de risque comme la VaR qui se focalise sur les événements de « type extrême ».

## B.1 Fondements théoriques et hypothèses

Soit  $(X_1, X_2, \dots, X_d)$  un vecteur aléatoire à l'étude. Posons  $\mathcal{F} \subset \mathbb{R}^d$  et  $\mathcal{U} = \mathbb{R}^d \setminus \mathcal{F}$ . Nous supposons que nous savons :

- i) la distribution marginale  $F_i$  de  $X_i$  sur  $\mathbb{R}$  pour  $i = 1, 2, \dots, d$ ;
- ii) la distribution de  $(X_1, X_2, \dots, X_d) \mid \{(X_1, X_2, \dots, X_d) \in \mathcal{F}\}$ ;
- iii) la probabilité  $p_f := P((X_1, X_2, \dots, X_d) \in \mathcal{F})$ , ainsi que  $p_u := P((X_1, X_2, \dots, X_d) \in \mathcal{U}) = 1 - p_f$ .

La distribution conjointe de  $(X_1, X_2, \dots, X_d)$  n'est donc pas complètement spécifiée (sauf si  $\mathcal{F} = \mathbb{R}^d$  et  $\mathcal{U} = \emptyset$ ). Par conséquent, les mesures de risque (p. ex., la VaR) de la somme agrégée  $\sum_{i=1}^d X_i$  ne peuvent être calculées avec précision. De fait, il existe de nombreux vecteurs  $(Y_1, Y_2, \dots, Y_d)$  qui concordent avec  $(X_1, X_2, \dots, X_d)$  pour ce qui est des propriétés i), ii) et iii), mais ils ont une différente mesure de risque de leur somme. Ci-après, nous cherchons à

trouver les valeurs extrêmes possibles de la mesure de risque considérée, car l'écart entre le minimum et le maximum peut servir à mesurer le risque de modélisation. Formellement, nous utilisons la définition suivante du risque de modélisation. Cette définition s'inscrit dans le même esprit que celui de [Barrieu et Scandolo \[2015\]](#).

**Définition B.1** (Risque de modélisation) Soit  $(X_1, X_2, \dots, X_d)$  un vecteur aléatoire vérifiant i), ii) et iii) et supposons qu'on utilise une mesure de risque  $\varrho(\cdot)$  (invariante en loi) pour évaluer le risque de  $\sum_{i=1}^d X_i$ . Définissons

$$\varrho_{\mathcal{F}}^{\pm} := \sup \left\{ \varrho \left( \sum_{i=1}^d Y_i \right) \right\}, \quad \varrho_{\mathcal{F}}^{\bar{-}} := \inf \left\{ \varrho \left( \sum_{i=1}^d Y_i \right) \right\}$$

où le supremum et l'infimum sont pris sur tous les autres (distributions conjointes de) vecteurs aléatoires  $(Y_1, Y_2, \dots, Y_d)$  qui respectent i), ii) et iii). Le risque de modélisation qui correspond à une sous-estimation du risque du fait que l'on calcule une estimation directe de  $\varrho(\sum X_i)$  dans un certain modèle de référence choisi (c.-à-d. lorsqu'une certaine distribution multivariée pour  $(X_1, X_2, \dots, X_d)$  a été spécifiée) se définit comme étant

$$\frac{\varrho_{\mathcal{F}}^{\pm} - \varrho(\sum_{i=1}^d X_i)}{\varrho_{\mathcal{F}}^{\pm}} \tag{B.28}$$

et, dans le même ordre d'idées, le risque de modélisation correspondant à une surestimation est donné par

$$\frac{\varrho(\sum_{i=1}^d X_i) - \varrho_{\mathcal{F}}^{\bar{-}}}{\varrho_{\mathcal{F}}^{\bar{-}}} \tag{B.29}$$

Le reste de cette section vise à obtenir les valeurs maximale et minimale possibles  $\varrho_{\mathcal{F}}^{\pm}$  et  $\varrho_{\mathcal{F}}^{\bar{-}}$  de  $\varrho(\sum_{i=1}^d X_i)$ . Les derniers ouvrages traitant du risque de modélisation se sont focalisés principalement sur le cas de complète incertitude quant à la dépendance entre les risques  $X_i$  ( $i = 1, 2, \dots, d$ ), c.-à-d. lorsque  $\mathcal{F} = \emptyset$  (cf. l'annexe A où nous avons examiné le travail d'[Embrechts et al. \[2013\]](#) en ce qui concerne la VaR; cf. aussi [Bernard et al. \[2014b\]](#) concernant une mesure de risque convexe). Dans cette section, nous considérons le cas dans lequel l'information sur la dépendance se traduit en distributions conjointes qui sont partiellement connues.

À cet égard, il sera utile de considérer la variable indicatrice  $\mathbb{I}$  correspondant à l'événement «  $(X_1, X_2, \dots, X_d) \in \mathcal{F}$  »

$$\mathbb{I} := \mathbb{1}_{(X_1, X_2, \dots, X_d) \in \mathcal{F}} \tag{B.30}$$

en sorte que nous puissions exprimer, de la façon suivante, les probabilités qu'un vecteur aléatoire prenne des valeurs dans  $\mathcal{U}$  :

$$p_f = P(\mathbb{I} = 1) \text{ and } p_u = P(\mathbb{I} = 0). \quad (\text{B.31})$$

Considérons aussi une variable aléatoire standard  $U$  distribuée uniformément qui est indépendante de l'événement «  $(X_1, X_2, \dots, X_d) \in \mathcal{F}$  » (et qui est donc aussi indépendante de  $\mathbb{I}$ ), ainsi qu'un vecteur aléatoire  $(Z_1, Z_2, \dots, Z_d)$  défini par

$$Z_i = F_{X_i | (X_1, X_2, \dots, X_d) \in \mathcal{U}}^{-1}(U), \quad i = 1, 2, \dots, d, \quad (\text{B.32})$$

où  $F_{X_i | (X_1, X_2, \dots, X_d) \in \mathcal{U}}^{-1}$  désigne l'inverse (gauche) de la fonction de répartition

$$F_{X_i | (X_1, X_2, \dots, X_d) \in \mathcal{U}}(x) := P(X_i \leq x | (X_1, X_2, \dots, X_d) \in \mathcal{U}).$$

À noter que  $F_{X_i | (X_1, X_2, \dots, X_d) \in \mathcal{U}}^{-1}(x)$  peut être calculé, car la distribution marginale de  $X_i$  est connue et la distribution conjointe de  $(X_1, X_2, \dots, X_d)$  est connue sur  $\mathcal{F}$  (cf. les propriétés i) et ii)). Par ailleurs, vu que tous les  $Z_i$  ( $i = 1, 2, \dots, d$ ) augmentent dans la variable  $U$  (commune),  $(Z_1, Z_2, \dots, Z_d)$  est un vecteur comonotone avec distribution conjointe connue. Définissons aussi

$$T := F_{\sum_i X_i | (X_1, X_2, \dots, X_d) \in \mathcal{F}}^{-1}(U). \quad (\text{B.33})$$

$T$  est donc une variable aléatoire avec distribution  $F_{\sum_i X_i | (X_1, X_2, \dots, X_d) \in \mathcal{F}}(x)$ .

Bien que la plupart de nos résultats se vérifient en règle générale ou puissent être aisément étendus à d'autres contextes, nous nous focaliserons sur les limites de la variance, de la VaR et de la TVaR (qui ont été définies à la section précédente en [\(A.21\)](#), [\(A.22\)](#) et [\(A.24\)](#)).

## B.2 Cadre pratique pour la méthode non paramétrique

Nous suivons le même contexte que celui déjà présenté à l'annexe [A](#). Nous avons  $N$  observations du vecteur à  $d$  dimensions  $(x_{i1}, x_{i2}, \dots, x_{id})$  pour  $i = 1, \dots, N$ . Nous désignons par  $M = (x_{ij})$  la matrice correspondante  $N \times d$ . Ces  $N$  observations peuvent être simplement  $N$  vecteurs de données observées ou  $N$  vecteurs de valeurs simulées obtenues d'une distribution multivariée ajustée de  $(X_1, X_2, \dots, X_d)$ . Dans les deux cas, chaque observation  $(x_{i1}, x_{i2}, \dots, x_{id})$  a lieu naturellement selon une probabilité  $\frac{1}{N}$  (possiblement avec répétitions). Nous supposons que la matrice  $M$  contient assez de données pour permettre une description précise des distributions marginales de  $X_k$  ( $k = 1, 2, \dots, d$ ), de sorte que la matrice  $M$  puisse être effectivement considérée comme une représentation du vecteur aléatoire  $(X_1, X_2, \dots, X_d)$  à l'étude. Définissons  $S_N$  par  $S_N(i) = \sum_{k=1}^d x_{ik}$  pour  $(i = 1, 2, \dots, N)$ . En d'autres termes,  $S_N$  peut être vue comme une variable aléatoire qui prend la valeur  $S_N(i)$  dans l'« état »  $i$  pour  $i = 1, 2, \dots, N$ . En général, il peut être difficile de trouver des limites *sharp* pour des mesures de

risque de  $S = \sum_i X_i$ . Ce qui suit a pour but de résoudre ce problème à l'aide du pendant « échantillonné »  $S_N$  de  $S$ , plutôt que  $S$  elle-même.

Comme dans le cadre théorique précédent, nous supposons que la distribution conjointe de  $(X_1, X_2, \dots, X_d)$  n'est pas complètement spécifiée. Dans le contexte de la représentation matricielle  $M$  du vecteur  $(X_1, X_2, \dots, X_d)$ , nous supposons que la matrice  $M$  est effectivement divisée en deux parties :  $\mathcal{F}_N$  désigne la sous-matrice des observations fiables  $(x_{i1}, x_{i2}, \dots, x_{id})$  et  $\mathcal{U}_N$  comporte le reste des observations. Ci-après, l'ensemble  $\mathcal{F}_N$  sera appelé partie « fixe » ou « fiable » et  $\mathcal{U}_N$  sera la partie « non fiable ». Si l'on a parfaitement confiance dans toutes les observations, la partie « non fiable » ne contient aucun élément ( $\mathcal{U}_N = \emptyset$ ) et  $S_N$  peut servir à évaluer le risque de  $S$ . Par contre, si l'on n'a pas confiance dans la dépendance observée,  $\mathcal{F}_N = \emptyset$ . Dans ce cas, les observations  $(x_{i1}, x_{i2}, \dots, x_{id})$  ne servent qu'à modéliser les distributions marginales  $F_k$  ( $k = 1, 2, \dots, d$ ) et ne permettent pas de tirer des conclusions concernant la dépendance. Il est donc important de faire remarquer que le réarrangement des valeurs  $x_{ik}$  ( $i = 1, 2, \dots, N$ ) dans la  $k^e$  colonne ne modifie pas la distribution marginale empirique de  $X_k$ , mais seulement la dépendance observée (interaction des éléments des différentes colonnes). Soit  $\ell_f$  le nombre d'éléments dans  $\mathcal{F}_N$ , et  $\ell_u$  le nombre d'éléments dans  $\mathcal{U}_N$ , de sorte que

$$N = \ell_f + \ell_u.$$

Sans perte de généralité, il est pratique de modifier la matrice  $M$  en changeant l'ordre des lignes de façon que la « région fiable » corresponde aux  $\ell_f$  premières lignes et la région non fiable aux dernières lignes. Ce faisant, nous avons seulement réaffecté les états  $i = 1, 2, \dots, N$  sans nuire à la capacité de  $M$  de décrire adéquatement les propriétés distributionnelles (invariantes en loi) de  $(X_1, X_2, \dots, X_d)$ . De même, conformément à la définition de la sous-matrice  $\mathcal{U}_N$ , nous pouvons réarranger les valeurs *dans les colonnes* de  $\mathcal{U}_N$  (et donc dans les parties correspondantes de  $M$ ), car cette opération génère une nouvelle matrice qui est considérée tout aussi fiable que la matrice initiale (notons, en fait, que nous ne connaissons pas la dépendance entre les  $X_i$ , conditionnellement au vecteur  $(X_1, X_2, \dots, X_d) \in \mathcal{U}$ ).

Sans perte de généralité, nous pouvons donc toujours supposer que la matrice  $\mathcal{U}_N$  représente une dépendance comonotone (dans chaque colonne, les valeurs sont triées en ordre décroissant, c.-à-d. que  $x_{m_1k} \geq x_{m_2k} \geq \dots \geq x_{m_{\ell_u}k}$  pour tous les  $k = 1, 2, \dots, d$ ). Enfin, pour  $\mathcal{F}_N$  (et donc dans les parties correspondantes de  $\mathcal{X}_N$ ) nous pouvons supposer que les  $\ell_f$  observations  $(x_{i_1,1}, x_{i_1,2}, \dots, x_{i_1,d})$  paraissent de manière que, pour les sommes des composantes, c.-à-d.  $s_j := x_{i_1,1} + x_{i_1,2} + \dots + x_{i_1,d}$  ( $j = 1, 2, \dots, \ell_f$ ), nous ayons  $s_1 \geq s_2 \geq \dots \geq s_{\ell_f}$ . Dorénavant, les points de données observés sont consignés dans la matrice  $M$  suivante :

$$M = \begin{bmatrix} x_{i_1 1} & x_{i_1 2} & \dots & x_{i_1 d} \\ x_{i_2 1} & x_{i_2 2} & \dots & x_{i_2 d} \\ \vdots & \vdots & \vdots & \vdots \\ x_{i_{\ell_f} 1} & x_{i_{\ell_f} 2} & \dots & x_{i_{\ell_f} d} \\ x_{m_1 1} & x_{m_1 2} & \dots & x_{m_1 d} \\ x_{m_2 1} & x_{m_2 2} & \dots & x_{m_2 d} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m_{\ell_u} 1} & x_{m_{\ell_u} 2} & \dots & x_{m_{\ell_u} d} \end{bmatrix}, \quad (\text{B.34})$$

où la région grise représente  $\mathcal{F}_N$  et la région blanche,  $\mathcal{U}_N$ . Les vecteurs correspondants  $S_N^f$  et  $S_N^u$  se composent, respectivement, des sommes des composantes de chaque observation des parties fiable et non fiable :

$$\begin{bmatrix} S_N^f \\ S_N^u \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{\ell_f} \\ \tilde{s}_1 := x_{m_1 1} + x_{m_1 2} + \dots + x_{m_1 d} \\ \tilde{s}_2 := x_{m_2 1} + x_{m_2 2} + \dots + x_{m_2 d} \\ \vdots \\ \tilde{s}_{\ell_u} := x_{m_{\ell_u} 1} + x_{m_{\ell_u} 2} + \dots + x_{m_{\ell_u} d} \end{bmatrix}. \quad (\text{B.35})$$

Bien que les  $s_1 \geq s_2 \geq \dots \geq s_{\ell_f}$  soient fiables, les sommes  $\tilde{s}_i$  changent lorsque le choix de la dépendance dans  $\mathcal{U}_N$  se modifie. De fait, l'ensemble  $\{i_1, \dots, i_{\ell_f}\}$  peut être vu comme la collection des états (scénarios) dans lequel les observations correspondantes sont fiables, tandis que l'ensemble  $\{m_1, \dots, m_{\ell_u}\}$  indique les états dans lesquels il existe des doutes quant à la structure de la dépendance.

À des fins pédagogiques, nous présentons maintenant un simple exemple de ce cadre. Il servira tout au long de cette section à illustrer chaque algorithme que nous proposons. Cet exemple fictif ne représente pas un ensemble réaliste d'observations, puisque dans de vraies applications, il y aurait un grand nombre d'observations (ici  $N = 8$ ) et possiblement un grand nombre de variables (ici  $d = 3$ ). Les huit observations sont indiquées ci-dessous, avec trois observations fiables ( $\ell_f = 3$ ), qui figurent dans la région grise de la matrice :

$$\begin{bmatrix} 3 & 4 & 1 \\ 1 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 2 & 1 \\ 2 & 4 & 2 \\ 3 & 0 & 1 \\ 1 & 1 & 2 \\ 4 & 2 & 3 \end{bmatrix} . \tag{B.36}$$

Sans perte de généralité, nous pouvons donc approfondir l'analyse et examiner la matrice  $M$  suivante ainsi que les vecteurs de sommes  $S_N^f$  et  $S_N^u$  suivants :

$$M = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 4 & 2 \\ 0 & 2 & 1 \\ 4 & 3 & 3 \\ 3 & 2 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} , \quad S_N^f = \begin{bmatrix} 8 \\ 8 \\ 3 \end{bmatrix} , \quad S_N^u = \begin{bmatrix} 10 \\ 7 \\ 4 \\ 3 \\ 1 \end{bmatrix} . \tag{B.37}$$

Enfin, en employant une notation abusive (compléter par 0 afin que  $S_N^f$  et  $S_N^u$  prennent huit valeurs), nous obtenons la représentation suivante de  $S_N$  :

$$S_N = \mathbb{I}S_N^f + (1 - \mathbb{I})S_N^u, \tag{B.38}$$

où  $\mathbb{I} = 1$  si  $(x_{i1}, x_{i2}, \dots, x_{id}) \in \mathcal{F}_N$  ( $i = 1, 2, \dots, N$ ). De fait,  $S_N^f$  peut aisément être considéré comme le pendant échantillonné de  $T$  que nous avons utilisé précédemment (cf. la définition [B.33](#)), tandis que  $S_N^u$  est une somme comonotone et correspond à la version échantillonnée de  $\sum_{i=1}^d Z_i$ . Dans cette dernière section, nous cherchons à trouver les structures de dépendance correspondant à la pire des situations et qui permettent une évaluation rigoureuse des risques de la somme  $S$  ( $S_N$ ) d'un portefeuille, ce qui équivaut à réarranger les résultats dans les colonnes de la partie non fiable  $\mathcal{U}_N$ , de façon à maximiser (respectivement minimiser) la mesure de risque considérée pour  $S_N$ .

### B.3 Limites d'une mesure de risque donnée

Nous traitons ici de la *sharpness*<sup>19</sup> des limites supérieures et inférieures des trois mesures de risque définies précédemment, à savoir la variance, la TVaR et la VaR.

Nous avons traité des limites analytiques et de la *sharpness* de ces limites à la section précédente, plus particulièrement la proposition [A.2](#), qui donne des limites non contraintes

<sup>19</sup> Pour rappel, une limite relative à une mesure de risque est dite *sharp* s'il existe une structure de dépendance entre les risques telle que cette limite est atteinte.

pour la variance, et la proposition [A.3](#), qui donne des limites non contraintes pour la VaR. Elles peuvent naturellement s'appliquer au cas avec contraintes que nous étudions ici.

### B.3.1 Limites théoriques de la variance et de la TVaR

La prochaine proposition permet de calculer facilement des limites supérieure et inférieure pour la variance de la somme  $\sum_{i=1}^d X_i$  d'un portefeuille.

**Proposition B.2** (Limites de la variance de  $\sum_{i=1}^d X_i$ ). *Soit  $(X_1, X_2, \dots, X_d)$  un vecteur aléatoire qui vérifie les propriétés i), ii) et iii), et soit  $\mathbb{I}$  et  $(Z_1, Z_2, \dots, Z_d)$  tels qu'ils sont définis en [\(B.30\)](#) et [\(B.32\)](#). Nous avons :*

$$\text{var} \left( \mathbb{I} \sum_{i=1}^d X_i + (1 - \mathbb{I}) \sum_{i=1}^d E(Z_i) \right) \leq \text{var} \left( \sum_{i=1}^d X_i \right) \leq \text{var} \left( \mathbb{I} \sum_{i=1}^d X_i + (1 - \mathbb{I}) \sum_{i=1}^d Z_i \right).$$

On trouvera la preuve de la proposition [B.2](#) dans [Bernard et Vanduffel \[2015a\]](#). Les limites supérieure et inférieure énoncées dans cette proposition sont intuitives et sont une extension naturelle de la proposition [A.2](#). Lorsqu'il s'agit de calculer la variance de la somme  $\sum_{i=1}^d X_i$  d'un portefeuille, il faut considérer séparément les événements  $(X_1, X_2, \dots, X_d) \in \mathcal{F}$  et

$(X_1, X_2, \dots, X_d) \in \mathcal{U}$ . La distribution de  $\sum_{i=1}^d X_i$  est connue relativement à l'événement  $\{(X_1, X_2, \dots, X_d) \in \mathcal{F}\}$ , mais elle est inconnue relativement à l'événement  $\{(X_1, X_2, \dots, X_d) \in \mathcal{U}\}$ . Sur  $\mathcal{U}$ , on remplace la somme  $\sum_i X_i$  par la constante  $\sum_i E(Z_i)$  (pour calculer la limite inférieure et ainsi minimiser la variance) et par la somme comonotone  $\sum_i Z_i$  (pour calculer la limite supérieure et ainsi maximiser la variance). On notera en particulier que lorsque  $\mathcal{U} = \emptyset$ , la limite supérieure est égale à la limite inférieure et le risque de modélisation est nul.

Passons maintenant aux limites de la TVaR, qui sont similaires à celles de la variance (parce que les deux mesures sont compatibles avec l'ordre convexe).

**Proposition B.3** (Limites de la TVaR de  $\sum_{i=1}^d X_i$ ). *Soit  $(X_1, X_2, \dots, X_d)$  un vecteur aléatoire qui vérifie les propriétés i), ii) et iii), et soit  $\mathbb{I}$  et  $(Z_1, Z_2, \dots, Z_d)$  tels qu'ils sont définis en [\(B.30\)](#) et [\(B.32\)](#). Nous avons :*

$$TVaR_p \left( \mathbb{I} \sum_{i=1}^d X_i + (1 - \mathbb{I}) \sum_{i=1}^d E(Z_i) \right) \leq TVaR_p \left( \sum_{i=1}^d X_i \right) \leq TVaR_p \left( \mathbb{I} \sum_{i=1}^d X_i + (1 - \mathbb{I}) \sum_{i=1}^d Z_i \right).$$

Il n'y a pas de risque de modélisation (les limites se réduisent à la même valeur) lorsque  $\mathcal{U} = \emptyset$ .

Certaines des limites énoncées aux propositions [B.2](#) et [B.3](#) peuvent être *sharp*. En particulier, les limites *supérieures* de la variance et de la TVaR indiquées aux propositions [B.2](#) et [B.3](#) sont *sharp*, sans autres conditions. On notera, en effet, que le vecteur multivarié

$$(\mathbb{I}X_1 + (1 - \mathbb{I})Z_1, \mathbb{I}X_2 + (1 - \mathbb{I})Z_2, \dots, \mathbb{I}X_d + (1 - \mathbb{I})Z_d) \tag{B.39}$$

vérifie les conditions i) et ii). Par contre, les limites inférieures énoncées ne sont pas forcément *sharp*, car  $\mathbb{I}X_i + (1 - \mathbb{I})E(Z_i)$  n'est habituellement pas distribué avec  $F_i$  ( $i = 1, 2, \dots, d$ ). Pour pouvoir s'approcher desdites limites inférieures, il faudrait chercher à modifier la dépendance du vecteur  $(Z_1, Z_2, \dots, Z_d)$  afin que  $Z_1 + Z_2 + \dots + Z_d$  devienne constant (et ainsi égal à  $E(Z_1) + E(Z_2) + \dots + E(Z_d)$ ). Forts de cette constatation, nous proposons ci-après un algorithme permettant d'estimer approximativement les limites *sharp* lorsque la mesure de risque utilisée est l'écart-type ou la TVaR.

### B.3.2 Limites pratiques de l'écart-type et de la TVaR

Selon la proposition [B.2](#), il est clair que, pour pouvoir maximiser la variance de  $S_N$ , il nous faut formuler un scénario comonotone sur  $\mathcal{U}_N$ . Cependant, nous avons déjà initialisé une structure comonotone (sans perte de généralité), et les valeurs correspondantes des sommes sont exactement les valeurs  $\tilde{s}_i$  ( $i = 1, 2, \dots, \ell_u$ ) déclarées pour  $S_N^u$  dans ([B.35](#)). La limite supérieure de la variance est ensuite calculée comme étant

$$\frac{1}{N} \left( \sum_{i=1}^{\ell_f} (s_i - \bar{s})^2 + \sum_{i=1}^{\ell_u} (\tilde{s}_i - \bar{s})^2 \right), \quad (\text{B.40})$$

où la somme moyenne  $\bar{s}$  est donnée par

$$\bar{s} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^d x_{ij} = \frac{1}{N} \left( \sum_{i=1}^{\ell_f} s_i + \sum_{i=1}^{\ell_u} \tilde{s}_i \right). \quad (\text{B.41})$$

Pour atteindre la limite de la variance minimale trouvée à la proposition [B.2](#), les valeurs de  $S_N^u$  doivent être aussi près que possible les unes des autres; dans l'idéal,  $S_N^u$  doit être constante. À cet égard, le concept de *complete mixability* semble être un outil théorique. Par *complete mixability*, on entend la structure de dépendance qui rend la somme  $S_N^u$  constante ([Wang et Wang \[2011\]](#)). Pour que ce soit le cas, en pratique, nous appliquons l'algorithme de réarrangement d'[Embrechts et al. \[2013\]](#) à la matrice  $U_N$  (partie non fiable) afin de la rendre aussi près que possible de l'état de *complete mixability*. Pour tout savoir sur cet algorithme, se reporter à la section [A.2](#). Nous désignons par  $\tilde{s}_i^m$  les valeurs correspondantes des sommes de  $S_N^u$  après application de l'algorithme de réarrangement. Nous calculons ensuite la variance minimale comme suit :

$$\frac{1}{N} \left( \sum_{i=1}^{\ell_f} (s_i - \bar{s})^2 + \sum_{i=1}^{\ell_u} (\tilde{s}_i^m - \bar{s})^2 \right), \quad (\text{B.42})$$

où  $\bar{s}$  est calculée comme en ([B.41](#)).

Nous illustrons les limites supérieure et inférieure ([B.40](#)) et ([B.42](#)) pour la variance obtenue précédemment avec la matrice  $M$  d'observations fournie en ([B.37](#)). Ensuite nous utilisons la structure comonotone pour la partie non fiable de la matrice  $M$  et nous calculons les vecteurs

des sommes  $S_N^f$  et  $S_N^u$  telles qu'elles sont définies en (B.37). La somme moyenne est de  $\bar{s} = 5,5$ . La variance maximale est égale à

$$\frac{1}{8} \left( \sum_{i=1}^3 (s_i - \bar{s})^2 + \sum_{i=1}^5 (\tilde{s}_i^c - \bar{s})^2 \right) \approx 8,75.$$

Pour ce qui est de la limite inférieure, nous appliquons l'algorithme de réarrangement à  $U_N$  et nous obtenons

$$M = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 4 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 3 \\ 0 & 3 & 2 \\ 1 & 2 & 2 \\ 3 & 1 & 1 \\ 4 & 0 & 1 \end{bmatrix}, \quad S_N^f = \begin{bmatrix} 8 \\ 8 \\ 3 \end{bmatrix}, \quad S_N^u = \begin{bmatrix} 5 \\ 5 \\ 5 \\ 5 \\ 5 \end{bmatrix}. \quad (\text{B.43})$$

Avec une somme moyenne  $\bar{s} = 5,5$ , la variance minimale peut être calculée comme

$$\frac{1}{8} \left( \sum_{i=1}^3 (s_i - \bar{s})^2 + \sum_{i=1}^5 (\tilde{s}_i^m - \bar{s})^2 \right) \approx 2,5.$$

Supposons que nous voulions fixer la TVaR au niveau de probabilité  $p$ , de façon que, pour simplicité d'exposition,

$$k := N(1 - p), \quad (\text{B.44})$$

où  $k$  est un entier. À l'exemple du cas de maximisation de la variance, il s'ensuit de la proposition B.3 que, pour obtenir la TVaR maximale, nous devons formuler un scénario comonotone sur  $\mathcal{U}_N$ . Il suffit donc de choisir les  $k$  valeurs les plus grandes de  $S_N^f$  et  $S_N^u$  calculées en (B.35). Désignons ces valeurs par  $s_1^*, s_2^*, \dots, s_k^*$  (classées par ordre décroissant), et nous pouvons ensuite facilement calculer la TVaR maximale au niveau de probabilité  $p$ . De plus, la TVaR minimale est obtenue de la même manière que la variance minimale. Tout d'abord, nous appliquons l'algorithme de réarrangement à la partie non fiable de la matrice  $U_N$  afin de rendre la variance de la (nouvelle) somme  $S_N^u$  aussi petite que possible. Puis nous choisissons les  $k$  valeurs les plus grandes de  $S_N^f$  et  $S_N^u$ , disons  $s_1^*, s_2^*, \dots, s_k^*$  (classées par ordre décroissant), et nous calculons la TVaR minimale.

Reprenons l'exemple précédent. Nous fixons  $p = 5/8$ , de sorte que  $k = 3$ . Les  $k = 3$  plus grandes valeurs sont 8, 8 et 10 et la TVaR maximale est donc de  $26/3 (\approx 8,67)$ . Après application de l'algorithme de réarrangement, nous obtenons (B.43) pour  $S_N^u$  et ainsi les trois plus grands résultats que nous observons pour  $S_N^u$  et  $S_N^f$  sont 8, 8 et 5. La TVaR minimale est donc de  $21/3 = 7$ .

### B.3.3 Limites théoriques de la VaR

L'utilisation de la VaR est très répandue dans le milieu des services financiers. La proposition suivante produit des limites pour la VaR.

**Proposition B.4** (Limites de la VaR de  $\sum_{i=1}^d X_i$ ). Soit  $(X_1, X_2, \dots, X_d)$  un vecteur aléatoire qui vérifie les propriétés i), ii) et iii), et soit  $\mathbb{I}, (Z_1, Z_2, \dots, Z_d)$  et  $U$  tels qu'ils sont définis en (B.30) et (B.32). Nous définissons les variables  $L_i$  et  $H_i$  comme étant

$$L_i = \text{LTVaR}_U(Z_i) \text{ et } H_i = \text{TVaR}_U(Z_i)$$

et soit

$$M_p := \text{VaR}_p \left( \mathbb{I} \sum_{i=1}^d X_i + (1 - \mathbb{I}) \sum_{i=1}^d H_i \right), \quad m_p := \text{VaR}_p \left( \mathbb{I} \sum_{i=1}^d X_i + (1 - \mathbb{I}) \sum_{i=1}^d L_i \right)$$

Les limites de la VaR du risque agrégé sont données par

$$m_p \boxplus \text{VaR}_p \left( \sum_{i=1}^d X_i \right) \boxplus M_p. \quad (\text{B.45})$$

On trouvera la preuve dans [Bernard et Vanduffel \[2015b\]](#). Au départ, les variables  $L_i$  et  $H_i$  peuvent paraître bizarres. Toutefois, on notera que les variables  $Z_i$ , qui jouent un rôle capital dans les propositions [B.2](#) et [B.3](#), peuvent aussi être exprimées par  $Z_i = \text{VaR}_U(Z_i)$ , et ici nous utilisons simplement  $\text{TVaR}_U(Z_i)$  et  $\text{LTVaR}_U(Z_i)$  à la place. La proposition [B.4](#) a donc une forme<sup>20</sup> similaire à celle, respectivement, des propositions [B.3](#) et [B.2](#), mais les limites proposées ne sont habituellement pas *sharp*<sup>21</sup>. Nous remarquons que, dans le cas de l'absence d'incertitude (c.-à-d.  $\mathcal{U} = \emptyset$ ), il n'y a pas de risque modélisation, car  $\mathbb{I} = 1$ . Lorsque l'incertitude est totale, soit lorsque  $\mathcal{U} = \mathbb{R}^d$ ,  $\mathbb{I} = 0$  et nous nous retrouvons dans le cas de figure de la proposition [A.3](#), où la limite inférieure de la VaR d'un portefeuille est non contrainte (cf. aussi le théorème 2.1 de [Bernard, Rüschemdorf et Vanduffel \[2013b\]](#)). On notera aussi que les limites de la VaR ne sont pas *sharp* en général.

Aux fins de calcul, il est pratique d'avoir recours à une autre formulation des limites énoncées de la VaR.

**Proposition B.5** (Autre formulation des limites de la VaR). Soit  $(X_1, X_2, \dots, X_d)$  un vecteur aléatoire qui vérifie les propriétés i), ii) et iii), et soit  $\mathbb{I}, (Z_1, Z_2, \dots, Z_d)$  et  $T$  tels qu'ils sont définis en (B.30), (B.32) et (B.33). On se rappellera que  $p_f = P(\mathbb{I} = 1)$ . Définissons

<sup>20</sup> À noter que la VaR n'est pas compatible avec l'ordre convexe, bien qu'il existe certains liens (cf. [Bernard et al. \[2013b\]](#) et [Bernard et al. \[2012\]](#)).

<sup>21</sup> On notera, en effet, que les variables  $H_i$  et  $L_i$  ne sont pas distribuées comme  $(X_i | \mathbb{I} = 0)$ .

$$\alpha_* := \inf \left\{ \alpha \in (\alpha_1, \alpha_2) \mid \text{VaR}_\alpha(T) \geq \text{TVaR}_{\frac{p-p_f\alpha}{1-p_f}} \left( \sum_{i=1}^d Z_i \right) \right\},$$

où  $\alpha_1 = \max \left\{ 0, \frac{p+p_f-1}{p_f} \right\}$  et  $\alpha_2 = \min \left\{ 1, \frac{p}{p_f} \right\}$ . Alors, pour  $p \in (0, 1)$ ,

$$M_p = \begin{cases} \text{TVaR}_{\frac{p-p_f\alpha_*}{1-p_f}} \left( \sum_{i=1}^d Z_i \right) & \text{if } \frac{p+p_f-1}{p_f} < \alpha_* < \frac{p}{p_f} \\ \text{VaR}_{\alpha_*}(T) & \text{if } \alpha_* = \frac{p}{p_f} \\ \max \left\{ \text{VaR}_{\alpha_*}(T), \text{TVaR}_{\frac{p-p_f\alpha_*}{1-p_f}} \left( \sum_{i=1}^d Z_i \right) \right\} & \text{if } \alpha_* = \frac{p+p_f-1}{p_f}. \end{cases} \quad (\text{B.46})$$

Les expressions pour la limite inférieure  $m_p$  s'obtiennent en remplaçant, dans les énoncés ci-dessus, « TVaR » par « LTVaR ».

On trouvera la preuve de la proposition [B.5](#) à l'annexe de l'ouvrage de [Bernard et Vanduffel \[2015a\]](#). Nous pouvons illustrer la proposition [B.5](#) dans le cadre discret, dans lequel l'espace de probabilités  $\Omega$  comporte  $N$  états. On suppose que l'événement  $\{(X_1, X_2, \dots, X_d) \in \mathcal{F}\}$  correspond à l'ensemble  $\{\omega_1, \dots, \omega_{\ell_f}\}$ , tandis que  $\{(X_1, X_2, \dots, X_d) \notin \mathcal{F}\}$  correspond à

l'ensemble  $\{\omega_{\ell_f+1}, \dots, \omega_{\ell_f+\ell_u}\}$  avec  $\ell_f + \ell_u = N$  ( $N$  est le nombre d'états). Alors  $p_f = \frac{\ell_f}{N}$  et

$p_u = \frac{\ell_u}{N}$ . On suppose que  $\sum_{i=1}^d Z_i$  prend  $\ell_u$  valeurs  $\tilde{s}_1 \geq \tilde{s}_2 \geq \dots \geq \tilde{s}_{\ell_u}$ , et que  $T$  prend donc  $\ell_f$  valeurs  $s_1 \geq s_2 \geq \dots \geq s_{\ell_f}$ . Plus précisément, on suppose que  $N = 8$ ,  $\ell_f = 3$ ,  $\ell_u = 5$ ,

$s_1 = 8 \geq s_2 = 8 \geq s_3 = 3$  et  $\tilde{s}_1 = 10 \geq \tilde{s}_2 = 7 \geq \tilde{s}_3 = 4 \geq \tilde{s}_4 = 3 \geq \tilde{s}_5 = 1$ ,  $p_f = 3/8$  et  $p_u = 5/8$ .

Pour  $p = 5/8$ , nous appliquons [\(B.46\)](#) dans la proposition [B.5](#) et constatons que  $\alpha_* = 0,75$  et que la  $\text{VaR}_p$  maximale est égale à  $\text{TVaR}_{0,55}(\sum Z_i) = 8$ . Ces valeurs seront confirmées au moyen de l'algorithme décrit ci-après.

### B.3.4 Limites pratiques de la VaR

Aux fins du calcul de la VaR maximale, nous présentons un algorithme pouvant s'appliquer directement à la matrice  $M$  des données observées et conduire ainsi à l'obtention de limites non paramétriques pour la VaR. On se rappellera que les  $\ell_f$  premières lignes de la matrice  $M$  correspondent à  $\mathcal{F}_N$ , tandis que  $\ell_u$  désigne le nombre de lignes de  $\mathcal{U}_N$  ( $N = \ell_f + \ell_u$ ). Dans l'algorithme, nous utilisons aussi  $S_N^f$  et  $S_N^u$ , que nous considérons comme des variables aléatoires. Pour calculer la VaR au niveau de probabilité  $p$ , nous définissons

$$k := N(1-p) \quad (\text{B.47})$$

où nous supposons que  $k$  est un entier.

L'algorithme se fonde sur la proposition [B.5](#) et sur la motivation suivante. On se rappellera de l'analyse de la proposition [B.4](#) que les limites supérieure et inférieure de la VaR ne sont pas *sharp*, en général. Néanmoins, nous sommes en mesure de proposer un algorithme permettant de se rapprocher des limites *sharp*. Nous expliquons ici cette idée de manière approfondie. Nous posons  $p \in (0, 1)$  et faisons remarquer que, presque assurément,

$$\mathbb{I} \sum_{i=1}^d X_i + (1 - \mathbb{I}) \sum_{i=1}^d Z_i \leq \mathbb{I} \sum_{i=1}^d X_i + (1 - \mathbb{I}) \sum_{i=1}^d H_i.$$

En particulier<sup>22</sup>, pour tous  $\alpha, \beta$  dans  $[0, 1]$  tels que  $p_f \alpha + (1 - p_f) \beta = p$ ,

$$\max \left\{ \text{VaR}_\alpha(T), \text{VaR}_\beta \left( \sum_{i=1}^d Z_i \right) \right\} \leq M_p = \max \left\{ \text{VaR}_{\alpha_*}(T), \text{TVaR}_{\beta_*} \left( \sum_{i=1}^d Z_i \right) \right\} \quad (\text{B.48})$$

où  $\alpha_*$  est défini comme dans la proposition [B.5](#) et  $\beta_* = \frac{p - p_f \alpha_*}{1 - p_f}$ . Le point critique est de

choisir  $\alpha$  et  $\beta$ , ainsi qu'une dépendance entre les composantes du vecteur (comonotone)  $(Z_1, Z_2, \dots, Z_d)$ , de telle sorte que l'inégalité [\(B.48\)](#) devienne une égalité. Pareille égalité est obtenue lorsque l'on prend  $\beta = \beta_*$  (donc  $\alpha_* = \alpha$ ) et une dépendance dans le vecteur  $(Z_1, Z_2, \dots, Z_d)$  telle que

$$\text{VaR}_{\beta_*} \left( \sum_{i=1}^d Z_i \right) = \text{TVaR}_{\beta_*} \left( \sum_{i=1}^d Z_i \right). \quad (\text{B.49})$$

La meilleure approximation de la limite *sharp* de  $\text{VaR}_p \left( \sum_{i=1}^d X_i \right)$  a donc probablement lieu lorsque la fonction quantile (VaR) de  $\sum_{i=1}^d Z_i$  peut être (presque) entièrement aplatie sur  $[\beta_*, 1]$ . Dans les cas où cette caractéristique ne peut être obtenue (ou presque), on ne peut exclure la possibilité de trouver de meilleures approximations (par exemple, si la fonction quantile de  $\sum_{i=1}^d Z_i$  peut être aplatie sur un autre intervalle  $[\beta, 1]$  dans lequel  $\beta$  est proche de  $\beta_*$ ). Par un raisonnement similaire, on voit que, pour pouvoir approcher aussi près que possible la limite inférieure indiquée, il faudrait que la fonction quantile de la somme du portefeuille soit aussi aplatie le plus possible sur l'intervalle  $[0, \beta_*]$ . Nous nous fondons sur cette idée pour proposer ci-dessous un algorithme pratique permettant d'estimer approximativement les limites *sharp*.

Ici,  $S_N^f$  joue le rôle de  $T$  et  $S_N^u$  joue le rôle de  $\sum_{i=1}^d Z_i$  (cf. aussi [\(B.38\)](#)). Sans perte de généralité, on suppose que  $S_N^f$  prend les valeurs  $s_1 \geq s_2 \geq \dots \geq s_{\ell_f}$ .

<sup>22</sup> Cf. [Bernard et Vanduffel \[2015a\]](#) pour avoir plus de détails et une intuition de la preuve de ce résultat.

Algorithme de calcul de la VaR maximale

1. On se rappellera que  $p_f = \frac{\ell_f}{N}$ . Calculez  $m_1 := \max\{0, \ell_f - k\}$  (de façon que  $\alpha_1 = \frac{m_1}{\ell_f} = \max\{0, \frac{p + p_f - 1}{p_f}\}$ ) et  $m_2 := \min\{\ell_f, N - k\}$  (alors  $\alpha_2 = \frac{m_2}{\ell_f} = \min\{1, \frac{p}{p_f}\}$ ).
2. Calculez  $\alpha_*$  où

$$\alpha_* := \inf \left\{ \alpha \in (\alpha_1, \alpha_2) \mid \text{VaR}_\alpha(S_N^f) \geq \text{TVaR}_{\frac{p-p_f\alpha}{1-p_f}}(S_N^u) \right\}.$$

3. Appliquez l'algorithme de réarrangement aux  $\lfloor (1 - \beta_*)\ell_u \rfloor$  premières lignes de la partie  $\mathcal{U}_N$  non fiable de la matrice  $M$ , où  $\beta_* = \frac{p - p_f\alpha_*}{1 - p_f}$  et où  $\lfloor \cdot \rfloor$  désigne la partie entière d'un nombre. À noter que  $\lfloor (1 - \beta_*)\ell_u \rfloor = k + m_* - \ell_f$ , où  $m_* := \lfloor \alpha_* \ell_f \rfloor$ , et que  $m_1 \leq m_* \leq m_2$ .

4. Par notation abusive, nous désignons par  $S_N^u$  les sommes réarrangées dans la partie non fiable. Il s'agit de la dépendance qui permet éventuellement d'atteindre la VaR

maximale en rendant  $\text{TVaR}_{\frac{\ell_u - b_*}{\ell_u}}(S_N^u)$  aussi près que possible de  $\text{VaR}_{\frac{\ell_u - b_*}{\ell_u}}(S_N^u)$ . Pour calculer cette VaR maximale possible, calculez toutes les sommes (lignes) pour  $\mathcal{U}_N$  et  $\mathcal{F}_N$  et ordonnez-les de la plus grande à la plus petite,  $\tilde{s}_1 \geq \tilde{s}_2 \geq \dots \geq \tilde{s}_k \geq \dots \geq \tilde{s}_N$ . La VaR est donc  $\tilde{s}_k$ .

L'algorithme qui précède se veut un moyen rapide de calculer des limites pour la VaR du risque agrégé qui sont éventuellement atteignables. Il nécessite un seul lancement de l'algorithme de réarrangement. Toutefois, vu que ce dernier génère rarement une somme parfaitement constante sur la région où il est appliqué, il se peut que l'on puisse obtenir une meilleure limite en appliquant l'étape 3 aux  $k + m - \ell_f$  premières lignes de  $\mathcal{U}_N$  pour un certain autre  $m$  ( $m_1 \leq m \leq m_2$ ).

Illustration de l'algorithme pour obtenir la VaR maximale dans l'exemple avec  $d = 3$ ,  $N = 8$ ,  $k = 3$  avec la même matrice  $M$  donnée en (B.37), de sorte que  $\ell_f = 3$  et  $\ell_u = 5$ .

Dans ce cas,  $\alpha_* = 0,75$  et la VaR maximale théorique est égale à  $\text{TVaR}_{\frac{p-p_f\alpha_*}{1-p_f}} = 8$ . Dans

l'algorithme,  $m_* = \lfloor \alpha_* \ell_f \rfloor = \lfloor 2,25 \rfloor = 2$  et la VaR maximale est obtenue pour  $m_* = 2$  (soit en appliquant l'algorithme de réarrangement aux  $k + m_* - \ell_f = 2$  premières lignes de la partie non

fiable de la matrice). En utilisant toutes les valeurs possibles de  $m$ , nous montrons ci-dessous qu'il s'agit effectivement de la valeur optimale.

Nous cherchons la valeur minimale et la valeur maximale de  $m$ ,

$$m_1 = \max(0,0) = 0, m_2 = \min(3,8-3) = 3,$$

de sorte que le nombre de lignes auquel on peut considérer d'appliquer l'algorithme de réarrangement se situe entre 0 et 3, car

$$k + m_1 - \ell_f = 0, k + m_2 - \ell_f = 3.$$

La première VaR que nous calculons en prenant trois lignes de  $S_N^u$  ( $m+k-\ell_f = 3$  avec  $m=3$ ) est égale à VaR=7 :

$$M = \begin{bmatrix} 3 & 4 & 1 \\ 0 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 3 & 3 \\ 3 & 2 & 2 \\ 4 & 1 & 2 \end{bmatrix}, \quad S_N^f = \begin{bmatrix} 8 \\ 8 \\ 3 \end{bmatrix}, \quad S_N^u = \begin{bmatrix} 7 \\ 7 \\ 7 \end{bmatrix}$$

La deuxième valeur est égale à VaR=8 ( $m+k-\ell_f = 2$  avec  $m=2$ ) et elle est déjà la plus grande valeur possible :

$$M = \begin{bmatrix} 3 & 4 & 1 \\ 0 & 2 & 1 \\ 2 & 4 & 2 \\ 3 & 3 & 3 \\ 4 & 2 & 2 \end{bmatrix}, \quad S_N^f = \begin{bmatrix} 8 \\ 8 \\ 3 \end{bmatrix}, \quad S_N^u = \begin{bmatrix} 9 \\ 10 \end{bmatrix}.$$

En fait, une itération de plus ( $m=1$ ) ne changera pas la valeur de la VaR, et deux itérations de plus ( $m=0$ ) produiront un nombre inférieur.

L'algorithme de calcul de la VaR minimale est semblable à celui servant au calcul de la VaR maximale, où TVaR est remplacé par LTVaR pour calculer  $\alpha_*$ . On trouvera plus de détails et un exemple dans [Bernard et Vanduffel \[2015a\]](#).

#### **B.4 Exemples de limites des mesures de risque des portefeuilles avec modélisation par Monte Carlo de l'incertitude au sujet de la dépendance**

Nous traitons maintenant de la façon de simuler les limites analytiques de la variance, de la TVaR et de la VaR (obtenues respectivement aux propositions [B.2](#), [B.3](#) et [B.5](#)), en donnant quelques exemples.

##### **B.4.1 Limites de la variance**

Les limites supérieure et inférieure de la variance énoncées à la proposition [B.2](#) peuvent être calculées par intégration numérique ou par simulation de Monte Carlo. Si le nombre  $d$  de dimensions est élevé, il est évident que le meilleur moyen de calculer les limites théoriques est

d'avoir recours aux techniques de Monte Carlo (à l'aide de simulations tirées du modèle multidimensionnel ajusté sur  $\mathcal{F}$ ). Nous illustrons la proposition [B.2](#) au moyen d'un exemple. À cet égard, il convient d'utiliser l'écart-type comme mesure de risque et non la variance (il est clair que, dans ce cas, les limites sont égales à la racine carrée de celles présentées à la proposition [B.2](#)). Cette façon de faire permet de comparer de manière juste les résultats de cet exemple avec ceux des exemples suivants qui utiliseront la TVaR ou la VaR comme mesure de risque.

**Exemple B.6** (Distribution normale multivariée comme modèle de référence). On suppose que  $(X_1, \dots, X_d)$  est un vecteur aléatoire avec distributions marginales normales standards. De plus, on suppose que la distribution conjointe de  $(X_1, \dots, X_d)$  est une distribution normale standard multivariée avec paramètre de corrélation<sup>23</sup>  $\rho$  sur le sous-ensemble  $\mathcal{F} := [q_\beta, q_{1-\beta}]^d \subset \mathbb{R}^d$  (pour un certain  $\beta < 50\%$ ), où  $q_\gamma$  désigne le quantile de la variable aléatoire normale standard au niveau de probabilité  $\gamma$ . Dans le tableau [B.1](#), nous supposons que  $d = 20$  et nous donnons les limites supérieure et inférieure de l'écart-type de la somme du portefeuille pour divers niveaux de confiance  $\beta$  et niveaux de corrélation  $\rho$ . La première colonne ( $\beta = 0\%$ ) indique les résultats des cas dans lesquels il n'y a aucune incertitude quant à la distribution multivariée; elle sert donc de référence pour évaluer le risque de modélisation (cf. définition [B.1](#)). La dernière colonne ( $\beta = 50\%$ ) indique les limites pour les cas dans lesquels il y a une incertitude totale au sujet de la dépendance; elle correspond donc à la situation habituellement étudiée dans la littérature.

$\mathcal{F} = [q_\beta, q_{1-\beta}]^d$	$\mathcal{U} = \emptyset$				
	$\beta = 0\%$	$\beta = 0.05\%$	$\beta = 0.5\%$	$\beta = 5\%$	$\mathcal{U} = \mathbb{R}^d$ $\beta = 50\%$
$\rho = 0$	4.47	(4.4, 5.65)	(3.89, 10.6)	(1.23, 19.3)	(0, 20)
$\rho = 0.1$	7.62	(7.41, 8.26)	(6.23, 11.7)	(1.69, 19.2)	(0, 20)
$\rho = 0.5$	14.5	(13.8, 14.6)	(11.1, 15.4)	(3.74, 18.6)	(0, 20)

**Tableau B.1** : Dans la première colonne, nous indiquons l'écart-type de  $\sum_{i=1}^{20} X_i$  dans l'hypothèse de normalité multivariée (aucune incertitude quant à la dépendance, c.-à-d. que  $\mathcal{U} = \emptyset$ ). Les limites inférieure et supérieure de l'écart-type de  $\sum_{i=1}^{20} X_i$  sont indiquées par paires  $(q_{\mathcal{F}}^-, q_{\mathcal{F}}^+)$  selon divers niveaux de confiance  $\beta$ . Nous utilisons 3 000 000 simulations. Tous les chiffres consignés dans le tableau sont significatifs.

Au vu du tableau [B.1](#), l'impact du risque de modélisation sur l'écart-type peut être considérable même lorsque la distribution conjointe de  $(X_1, \dots, X_d)$  est presque parfaitement connue, soit lorsque  $\beta$  avoisine zéro ( $p_u$  est près de 0). Posons par exemple  $\beta = 0,05\%$  et  $\rho = 0$ . Dans ce cas,  $p_u = 1 - 0,999^{20} \approx 0,02$ , et nous constatons que l'utilisation de l'hypothèse de normalité multivariée (comme référence) pouvait conduire à sous-estimer l'écart-type de  $(5,65-4,47)/4,47 = 26,4\%$  et à le surestimer de  $(4,47-4,4)/4,4 = 1,6\%$ . Il semble donc que

<sup>23</sup> Une distribution normale standard multivariée est dite avoir un coefficient de corrélation  $\rho$  si la corrélation entre toutes les paires  $(X_i, X_j)$ ,  $i \neq j$ , est égale à  $\rho$ .

l'hypothèse de normalité multivariée ne soit pas particulièrement résistante face à l'erreur de spécification. Ici, en fait, il est évident qu'elle donne lieu à une situation dans laquelle il est plus probable de sous-estimer le risque que de le surestimer. De plus, l'exemple montre que l'ajout d'information additionnelle au sujet de la dépendance (soit lorsque  $\beta < 50\%$ ) peut modifier les limites non contraintes (cas dans lequel  $\beta = 50\%$ ) et il vient confirmer que la dépendance est importante lorsqu'il s'agit d'évaluer le risque d'un portefeuille. Par exemple, lorsque  $\beta = 0,5\%$  et  $\rho = 0$ , on a  $p_u = 1 - 0,99^{20} \approx 0,18$  et la limite supérieure non contrainte de l'écart-type perd près de la moitié de sa valeur (passant de 20 à 10,6).

$p_u$	$\mathcal{U} = \emptyset$	$\beta = 0,05\%$	$\beta = 0,5\%$	$\beta = 5\%$	$\mathcal{U} = \mathbb{R}^d$
	$\beta = 0\%$				$\beta = 50\%$
$\rho = 0$	0	0.02	0.18	0.88	1
$\rho = 0.1$	0	0.02	0.18	0.87	1
$\rho = 0.5$	0	0.016	0.12	0.66	1

**Tableau B.2 :** Probabilité  $p_u$  que  $(X_1, \dots, X_d)$  prenne des valeurs situées à l'extérieur du cube  $d$   $[q_\beta, q_{1-\beta}]^d$  selon un niveau de confiance  $\beta$  et un niveau de corrélation  $\rho$ . Nous utilisons 3 000 000 simulations.

Dans le tableau [B.2](#), nous indiquons, selon les niveaux de corrélation  $\rho$  et les niveaux de confiance  $\beta$  utilisés dans le tableau [B.1](#), la probabilité  $p_u$  que  $(X_1, \dots, X_d)$  prenne des valeurs situées à l'extérieur du cube  $d$   $\mathcal{F} = [q_\beta, q_{1-\beta}]^d$ . Cette façon de faire nous permet de mieux interpréter les résultats du tableau [B.1](#) et de mieux comprendre les effets du choix d'une autre structure pour la région fiable  $\mathcal{F}$ .

Dans l'exemple précédent, la région fiable est basée uniquement sur l'utilisation des densités marginales  $N(0,1)$ . Plus généralement, nous supposons que les densités marginales ont été ajustées aux densités empiriques  $\hat{f}_i$  pour  $i = 1, \dots, d$ , respectivement.  $\mathcal{F}$  est donc défini par

$$\mathcal{F} := \left\{ (x_{1k}, \dots, x_{dk}) \in \mathbb{R}^d / \forall j \in \{1, 2, \dots, d\}, \hat{f}_j(x_{jk}) \geq \varepsilon \right\} \quad (\text{B.50})$$

Dans le cas où les événements rares correspondent aux plus grands ou aux plus petits résultats des risques, cette approche est compatible avec l'utilisation d'un cube  $d$  comme région fiable. Un autre critère naturel permettant de déterminer la partie fiable de la distribution multivariée est celui consistant à commencer par une certaine densité multivariée ajustée  $\hat{f}$  (provenant, par exemple, d'un modèle gaussien multivarié, d'un modèle multivarié de Student ou d'un modèle de construction de paires de copules ([Aas et al. \[2009\]](#), [Czado \[2010\]](#))). La région fiable est ensuite basée sur les niveaux de contour de la densité. Nous renvoyons le lecteur à [Bernard et Vanduffel \[2015a\]](#) s'il veut obtenir plus de détails à ce sujet. Ces observations se comprennent aussi intuitivement, puisque l'écart-type est sensible aux résultats élevés et ces scénarios se produisent fréquemment lorsqu'on considère la limite supérieure (les événements extrêmes sont alors supposés être entièrement corrélés).

### B.4.2 Limites de la TVaR

Nous nous servons maintenant du même exemple pour illustrer et calculer les limites de la TVaR au moyen de simulations de Monte Carlo.

**Exemple B.7** (Distribution normale multivariée comme modèle de référence). Le tableau [B.3](#) indique les limites de la TVaR selon divers niveaux de probabilité  $p$  et selon divers niveaux de confiance  $\beta$  et de corrélation  $\rho$ . Les résultats concordent avec ceux de l'exemple précédent. Le risque de modélisation est toujours présent lorsque  $\beta$  est faible, mais, en même temps, la disponibilité de l'information sur la dépendance ( $\beta < 50\%$ ) permet de renforcer considérablement les limites non contraintes ( $\beta = 50\%$ ). Il est intéressant de noter que le degré de risque de modélisation dépend aussi de l'interaction du niveau de probabilité  $p$  utilisé pour évaluer la TVaR et du degré d'incertitude quant à la dépendance, tel qu'il est mesuré par  $\beta$ . Lorsque  $p$  est grand (p. ex.,  $p = 99,5\%$ ), une faible proportion d'incertitude du modèle (p. ex.,  $\beta = 0.05\%$ ) semble avoir un effet considérable sur le risque de sous-estimation découlant du modèle. Nous pouvons expliquer ce fait comme suit. La TVaR mesure essentiellement la moyenne de toutes les VaR supérieures et son niveau est donc principalement déterminé par les scénarios dans lesquels un ou plusieurs résultats des risques considérés sont élevés. Toutefois, ces scénarios ne sont pas considérés comme fiables pour représenter la dépendance (dans la queue) avec un impact négatif sur le niveau de la TVaR. De fait, pour un niveau donné de  $p$ , le risque de sous-estimation découlant du modèle augmente brusquement lorsque  $\beta$  augmente et il s'approche de son maximum déjà pour des valeurs faibles ou moyennes de  $\beta$ . Cet effet est encore mis en évidence lorsque  $p$  augmente. Autrement dit, la TVaR est très vulnérable à l'erreur de spécification du modèle, surtout lorsqu'elle est évaluée selon des niveaux de probabilité élevés.

$F = [q_\beta, q_{1-\beta}]^d$	$\mathcal{U} = \emptyset$ $\beta = 0\%$	$\beta = 0.05\%$	$\beta = 0.5\%$	$\beta = 5\%$	$\mathcal{U} = \mathbb{R}^d$ $\beta = 50\%$	
$p = 95\%$	$\rho = 0$	9.21	(9.12 , 11.6)	(8.49 , 27.5)	(3.36 , 41.3)	(-0.002 , 41.3)
	$\rho = 0.1$	15.7	(15.4 , 17.3)	(13.5 , 28.4)	(4.72 , 41.3)	(0.004 , 41.3)
	$\rho = 0.5$	29.9	(28.1 , 30.5)	(22.9 , 34.0)	(10.0 , 41.3)	(-0.002 , 41.3)
$p = 99.5\%$	$\rho = 0$	12.9	(12.8 , 30.4)	(12.1 , 57.9)	(7.52 , 57.9)	(-0.004 , 57.9)
	$\rho = 0.1$	22	(21.5 , 33.3)	(19.0 , 57.8)	(10.0 , 57.9)	(-0.002 , 57.9)
	$\rho = 0.5$	42	(37.4 , 47.6)	(29.6 , 57.9)	(15.2 , 57.9)	(0.019 , 57.9)

**Tableau B.3 :** TVaR<sub>95%</sub> et TVaR<sub>99,5%</sub> de  $\sum_{i=1}^{20} X_i$  sont indiquées en l'absence d'incertitude (modèle normal standard multivarié avec  $\mathcal{U} = \emptyset$ ). Les limites sont ensuite données selon divers niveaux de confiance  $\beta$ , de corrélation  $\rho$  et de probabilité  $p$ . Le calcul des limites se base sur 3 000 000 simulations. Tous les chiffres indiqués sont significatifs.

À l'exemple du cas de l'écart-type, on peut aussi utiliser une région fiable qui est basée sur les contours de la distribution normale multivariée, afin d'évaluer les limites supérieure et inférieure. Vu que les résultats sont similaires, nous n'en donnons pas les détails.

### B.4.3 Limites de la VaR

Nous analysons ici les limites de la VaR lorsque le modèle de référence est une distribution normale multivariée.

**Exemple B.8** (Distribution normale multivariée comme modèle de référence). Les limites de la VaR indiquées dans le tableau B.4 ont été obtenues en quelques minutes, au moyen de 3 000 000 simulations de Monte Carlo. Nous faisons les observations suivantes. Tout d'abord, le risque de modélisation est nettement présent même lorsque la dépendance est principalement connue (soit lorsque  $\beta$  est petit). De plus, le degré précis d'erreur de spécification dépend fortement du niveau de probabilité  $p$  utilisé pour évaluer la VaR. Considérons le modèle de référence avec  $\rho = 0$  (les risques sont indépendants et suivent une distribution normale standard) et  $\beta = 0$  % (aucune incertitude). Nous obtenons

$$VaR_{95\%} \left( \sum_{i=1}^{20} X_i \right) = \sqrt{20} \Phi^{-1}(95\%) = 7,35 \text{ et, de même, } VaR_{99,5\%} \left( \sum_{i=1}^{20} X_i \right) = 11,5,$$

$$VaR_{99,95\%} \left( \sum_{i=1}^{20} X_i \right) = 14,7. \text{ Toutefois, si } \beta = 0,05\%, p_u \approx 0,02, \text{ et le modèle de référence}$$

pourrait sous-estimer la VaR à 95 % de  $(8,08-7,36)/8,08 = 8,9$  % ou la surestimer de  $(7,36-7,27)/7,27 = 1,24$  %. Cependant, lorsqu'on utilise la VaR à 99,5 %, le degré de sous-estimation peut monter à  $(30,4-11,5)/30,4 = 62,2$  %, tandis que le degré de surestimation est égal à seulement  $(11,5-11,4)/11,4 = 0,9$  %. Le risque de sous-estimation augmente donc brusquement en fonction du niveau de probabilité qui est utilisé pour évaluer la VaR.

$F = [q_\beta, q_{1-\beta}]^d$	$\mathcal{U} = \emptyset$ $\beta = 0\%$	$\beta = 0.05\%$	$\beta = 0.5\%$	$\beta = 5\%$	$\mathcal{U} = \mathbb{R}^d$ $\beta = 0.5$
$p=95\%$ $\rho = 0$	7.36	( 7.27 , 8.08 )	( 6.65 , 27.5 )	( 0.79 , 41.3 )	( -2.17 , 41.3 )
$p=95\%$ $\rho = 0.1$	12.5	( 12.2 , 13.3 )	( 10.7 , 27.7 )	( 1.51 , 41.2 )	( -2.17 , 41.3 )
$p=95\%$ $\rho = 0.5$	23.8	( 22.9 , 24.2 )	( 18.9 , 30.9 )	( 6.97 , 41.2 )	( -2.17 , 41.3 )
$p=99.5\%$ $\rho = 0$	11.5	( 11.4 , 30.4 )	( 10.8 , 57.8 )	( 6.13 , 57.8 )	( -0.29 , 57.8 )
$p=99.5\%$ $\rho = 0.1$	19.6	( 19.1 , 31.4 )	( 16.9 , 57.8 )	( 8.23 , 57.8 )	( -0.29 , 57.8 )
$p=99.5\%$ $\rho = 0.5$	37.4	( 34.3 , 45.1 )	( 27.4 , 57.8 )	( 13.5 , 57.8 )	( -0.29 , 57.8 )
$p=99.95\%$ $\rho = 0$	14.7	( 14.6 , 71.0 )	( 13.8 , 71.1 )	( 9.31 , 71.1 )	( -0.036 , 71.1 )
$p=99.95\%$ $\rho = 0.1$	25.1	( 24.2 , 71.1 )	( 21.5 , 71.1 )	( 12.1 , 71.1 )	( -0.035 , 71.1 )
$p=99.95\%$ $\rho = 0.5$	47.7	( 41.3 , 71.1 )	( 32.3 , 71.1 )	( 17.2 , 71.1 )	( -0.036 , 71.1 )

**Tableau B.4 :**  $VaR_{95\%}$ ,  $VaR_{99,5\%}$  et  $VaR_{99,95\%}$  de  $\sum_{i=1}^{20} X_i$  sont indiquées en l'absence

d'incertitude (modèle normal standard multivarié avec  $\mathcal{U} = \emptyset$ ). Les limites sont ensuite données selon divers niveaux de confiance  $\beta$ , de corrélation  $\rho$  et de probabilité  $p$ . Nous utilisons 3 000 000 simulations et tous les chiffres indiqués sont significatifs.

Enfin, on note que lorsque des niveaux de probabilité très élevés sont utilisés pour calculer la VaR ( $p = 99,95$  %; se reporter aux trois dernières lignes du tableau B.4), les limites supérieures contraintes sont très près de la limite supérieure non contrainte, même lorsqu'il n'y a presque aucune incertitude au sujet de la dépendance ( $\beta = 0,05$  %). Les limites calculées par Embrechts et al. [2013] sont donc quasiment les meilleures limites possibles, même s'il semble que le modèle multivarié est connu selon un très haut niveau de confiance. Cela implique que tout

effort visant à ajuster avec précision un modèle multivarié ne permettra pas de réduire le risque de modélisation associé à la mesure de risque (et au capital requis).

On notera que lorsqu'aucune information sur la dépendance n'est disponible ( $\beta = 50\%$ ), les limites supérieure et inférieure indiquées à la proposition [B.4](#) se réduisent respectivement à  $\sum_{i=1}^d TVaR_p(X_i)$  et  $\sum_{i=1}^d LTVaR_p(X_i)$  et coïncident avec la limite inférieure  $A$  et la limite supérieure  $B$  données par [Bernard et al. \[2013b\]](#). À l'aide de leurs formules pour  $A$  et  $B$ , nous constatons que les limites de la  $VaR_p$  des sommes de 20 risques indépendants  $\mathcal{N}(0, 1)$  sont données par

$$A = -20 \frac{\phi(\Phi^{-1}(p))}{p}, \quad B = 20 \frac{\phi(\Phi^{-1}(p))}{1-p}$$

et nous remarquons qu'elles sont compatibles avec les limites indiquées dans le tableau [B.4](#). Par exemple, lorsque  $p = 95\%$ , nous obtenons  $(A, B) = (-2,17; 41,25)$ , ce qui concorde avec les chiffres du tableau [B.4](#).

**Exemple B.9** (Risques suivant une distribution de Pareto). Nous présentons un autre exemple avec les risques suivant une distribution de Pareto et nous constatons que les mêmes résultats peuvent être obtenus. Ils sont même amplifiés. Nous supposons que les risques individuels suivent tous une loi de Pareto de paramètre  $\theta = 3$ . Nous supposons qu'il y a  $d = 20$  risques suivant une Pareto de paramètre  $\theta$  et que leur dépendance est modélisée par une copule gaussienne de paramètre  $\rho$  (corrélation par paires). La fonction de répartition de chaque risque, pour  $x > 0$ , est définie par

$$F(x) = 1 - (1+x)^{-\theta}.$$

La  $VaR$  au niveau  $p \in (0,1)$  est donnée par

$$F_X^{-1}(p) = (1-p)^{-1/\theta} - 1.$$

Nous supposons que la zone fiable se base sur le fait que chaque distribution marginale se trouve entre le quantile de niveau  $1-\beta$  et  $\beta$

$$\mathcal{F} = \bigcap_{k=1}^d \{q_\beta \leq X_k \leq q_{1-\beta}\}$$

où

$$q_\beta = (1-\beta)^{-1/\theta} - 1$$

Après  $n = 3\,000\,000$  simulations de  $d = 20$  variables de Pareto, nous obtenons les résultats suivants pour les limites de la  $VaR$  :

$F = [q_\beta, q_{1-\beta}]^d$	$\mathcal{U} = \emptyset$				$\mathcal{U} = \mathbb{R}^d$
	$\beta = 0\%$	$\beta = 0.05\%$	$\beta = 0.5\%$	$\beta = 5\%$	$\beta = 0.5$
$\alpha=95\% \quad \rho = 0$	16.6	( 16 , 18.4 )	( 13.8 , 37.4 )	( 8.62 , 61.4 )	( 7.29 , 61.4 )
$\alpha=95\% \quad \rho = 0.1$	19.7	( 18.3 , 20.6 )	( 15.9 , 37.8 )	( 8.82 , 61.4 )	( 7.29 , 61.4 )
$\alpha=95\% \quad \rho = 0.5$	28	( 26.5 , 33.5 )	( 20.6 , 43.2 )	( 10.3 , 61.4 )	( 7.29 , 61.4 )
$\alpha=99.5\% \quad \rho = 0$	25.8	( 21.5 , 60.7 )	( 17.5 , 156 )	( 10.7 , 156 )	( 9.27 , 156 )
$\alpha=99.5\% \quad \rho = 0.1$	32.5	( 27.9 , 63 )	( 21.8 , 156 )	( 11.6 , 156 )	( 9.27 , 156 )
$\alpha=99.5\% \quad \rho = 0.5$	61.1	( 49 , 94.7 )	( 31.6 , 155 )	( 14 , 155 )	( 9.26 , 155 )
$\alpha=99.95\% \quad \rho = 0$	43.5	( 26.5 , 359 )	( 20.5 , 360 )	( 12.4 , 360 )	( 9.83 , 359 )
$\alpha=99.95\% \quad \rho = 0.1$	51.9	( 36.3 , 357 )	( 26.8 , 359 )	( 13.9 , 358 )	( 9.82 , 357 )
$\alpha=99.95\% \quad \rho = 0.5$	116	( 69.6 , 361 )	( 40.1 , 361 )	( 16.8 , 359 )	( 9.83 , 359 )

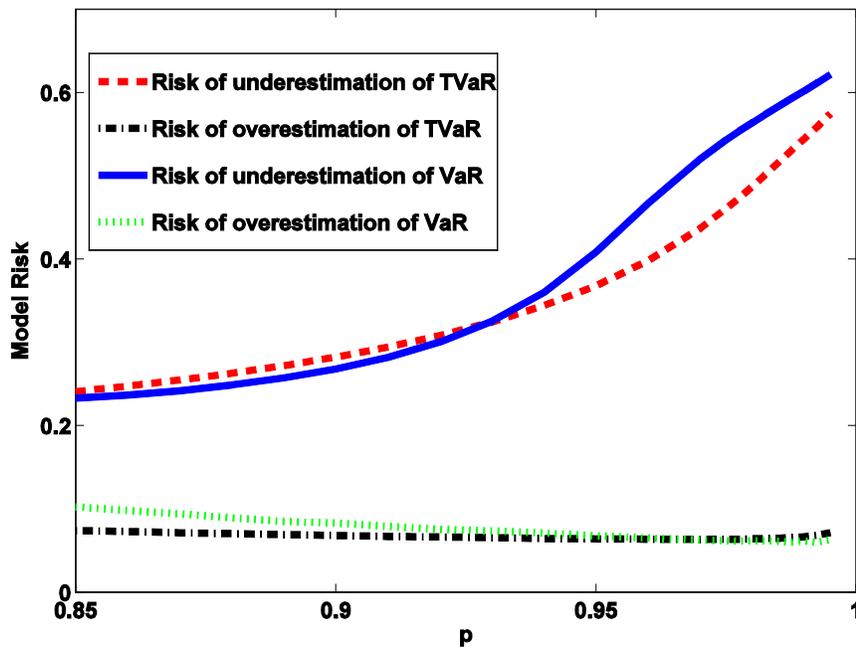
#### B.4.4 Autres considérations relatives au risque de modélisation

Considérons encore une fois un vecteur aléatoire  $(X_1, \dots, X_d)$  dont les distributions marginales suivent une loi normale standard et sont toutes corrélées selon un coefficient de 10 % (référence). Nous nous focalisons maintenant sur le risque de sous-estimation et de surestimation découlant du modèle; c.-à-d. que nous considérons les quantités

$$\frac{\varrho_{\mathcal{F}}^+ - \varrho(\sum_i X_i)}{\varrho_{\mathcal{F}}^+} \quad \text{and} \quad \frac{\varrho_{\mathcal{F}}^- - \varrho(\sum_i X_i)}{\varrho_{\mathcal{F}}^-}, \quad (\text{B.51})$$

qui ont été introduites à la section [B.4](#) (définition [B.1](#) et expressions [\(B.28\)](#) et [\(B.29\)](#)). La mesure de risque  $\varrho(\cdot)$  est la VaR et la TVaR, et, pour la région fiable  $\mathcal{F}$ , nous considérons les contours elliptiques tels que  $P((X_1, \dots, X_d) \in \mathcal{F}) = p_{\mathcal{F}}$ .

À la figure [B.3](#), nous représentons le risque de sous-estimation et de surestimation de la VaR et de la TVaR, respectivement, selon divers niveaux de probabilité  $p$  et en utilisant les mesures de risque ([B.51](#)) pour le risque de modélisation. D'après la figure [B.3](#), nous constatons qu'une légère erreur de spécification du modèle entraîne possiblement déjà une importante sous-estimation de la VaR et de la TVaR. Par contre, le risque de surestimation semble être moins grand. Nous pouvons expliquer ces observations comme suit. Dans le modèle de référence, il est supposé que les risques  $X_i$  ( $i = 1, 2, \dots, d$ ) suivent une loi normale multivariée, avec un coefficient de corrélation de seulement 10 %. Toutefois, en présence d'incertitude, les risques sont supposés être entièrement dépendants dans la région non fiable  $\mathcal{U}$  lorsqu'il s'agit de calculer la limite supérieure de la TVaR, et se comporter comme une constante lorsqu'il s'agit de calculer la limite inférieure (la somme du portefeuille est donc elle aussi constante dans ce cas). La dernière situation étant plus près de celle qui est présente dans le modèle de référence, le risque de surestimation de la TVaR est relativement faible. Inversement, le risque de sous-estimation est plutôt considérable. On observe les mêmes tendances pour les limites de la VaR, car celles-ci sont basées sur la TVaR d'une somme comonotone et diffèrent donc de la VaR sous le modèle de référence (les risques sont faiblement corrélés).



Risk of underestimation of TVaR = Risque de sous-estimation de la TVaR

Risk of overestimation of TVaR = Risque de surestimation de la TVaR

Risk of underestimation of VaR = Risque de sous-estimation de la VaR

Risk of overestimation of VaR = Risque de surestimation de la VaR

Model Risk = Risque de modélisation

**Figure B.3 :** Nous supposons que  $(X_1, X_2, \dots, X_{20})$  est une distribution standard multivariée avec corrélation par paires  $\rho = 0,1$ . Posons  $p_f = 90\%$ . Nous montrons le risque de surestimation ou de sous-estimation de  $\text{VaR}_p$  et de  $\text{TVaR}_p$ , comme une fonction de  $p$ .

Nous constatons aussi que lorsque le niveau de probabilité  $p$  est élevé, le risque de sous-estimer la VaR en raison du modèle semble être plus grand que le risque de sous-estimer la TVaR pour la même raison. Nous expliquons comme suit ce fait remarquable. Lorsque  $p$  est suffisamment grand (en comparaison de  $1 - p_u$ ), la pire des VaR et la pire des TVaR de la somme du portefeuille sont toutes deux basées sur les scénarios non fiables, car les plus grands résultats pour la somme du portefeuille correspondent habituellement aux scénarios non fiables. Ainsi, dans ce cas, la pire des VaR et la pire des TVaR ont tendance à être près l'une de l'autre<sup>24</sup>. Toutefois, l'écart entre la TVaR modélisée et la VaR modélisée demeure, bien entendu, strictement positif. Ensemble, les deux effets impliquent que le risque de sous-estimer la VaR en raison du modèle est plus important que le risque de sous-estimer la TVaR lorsque  $p$  est très grand. Cet exemple laisse donc penser que la VaR est plus sensible au risque

<sup>24</sup> [Puccetti et Rüschendorf \[2012a\]](#) ont montré que, sous des conditions peu restrictives et pour un ensemble donné de scénarios, la pire des VaR se comportait, de manière asymptotique, comme la pire des TVaR.

de modélisation que ne l'est la TVaR. Il montre aussi qu'un modèle peut bien s'ajuster aux données dans leur ensemble mais néanmoins ne pas être adéquat pour estimer la VaR à des niveaux de probabilité élevés.

## Annexe C – Définitions des notations mathématiques

Symbole	Explication	Exemples
$X_i$	Risque individuel ( $x_{i,j,d}$ dans la matrice)	$S = \sum_{i=1}^d X_i$
$S$	Risque agrégé ( $s_i$ dans la matrice)	
$\mathbf{X}$	Vecteur de perte du portefeuille	$\mathbf{X} = (X_1, \dots, X_d)$
$\{\dots\}$	Ensemble	$\{(X_{1i}, \dots, X_{di})\}_{i=1, \dots, N}$ , Un ensemble de $N$ vecteurs d'observations à $d$ dimensions
$\emptyset$	Ensemble vide	$\mathcal{F} = \emptyset$
$\varrho$	Mesure du risque	VaR, TVaR, <i>etc.</i>
$K$	Exigence de capital	
$()^+$ or $()_+$	« Insuffisance »	$(S - \varrho(S))_+ := \max(0, S - \varrho[S])$ , $(S - \varrho(S))_+$ , la perte totale que l'assureur n'est pas en mesure de rembourser, <i>est définie comme</i> $S - \varrho[S]$ , si elle est plus grande que 0.
$:=$	« défini comme »	
$\in$	Appartient à	$b \in \mathbb{R}$ , $b$ appartient aux nombres réels
$\bar{T}$	Moyenne de la fonction de score	Fonction de score $T$ : erreur quadratique, erreur absolue, <i>etc.</i>
$\operatorname{argmin}_x$	Argument du minimum	$x^* = \operatorname{argmin}_x E(T(x, S))$ , point $x$ pour lequel $E(T(x, S))$ atteint sa valeur minimale
$C$	Fonction copule	
$U$	Distribution uniforme	$\mathbf{X}^d = (F_{X_1}^{-1}(U_1), F_{X_2}^{-1}(U_2), \dots, F_{X_n}^{-1}(U_d))$
$\stackrel{d}{=}$	Partagent les mêmes fonctions de distribution	
$F_Y, F_Y^{-1}$	Fonction de répartition, fonction de répartition inverse de $Y$	
$\leq_{cX}$	Ordre convexe	$X \leq_{cX} Y$ , « $X$ et $Y$ ont la même moyenne mais que $Y$ a la plus grande variance »

$\tilde{X}_j$	$X_j$ après réarrangement	
U	Union de ... ou ...	
$\mathcal{F}$	Région « fixe » ou « fiable »	$\mathbb{R}^d = \mathcal{F} \cup \mathcal{U}$
$\mathcal{U}$	Région « non fixe » ou « non fiable »	
$\subset$	Sous-ensemble	$\mathcal{F} \subset \mathbb{R}^d$
$\setminus$	« excluant »	$\mathcal{U} = \mathbb{R}^d \setminus \mathcal{F}$
	Tel que	$(X_1, X_2, \dots, X_d) \mid \{(X_1, X_2, \dots, X_d) \in \mathcal{F}\}$
$q^+$	Borne supérieure (supremum) de $q$	$q_{\mathcal{F}}^+ := \sup\{q(\sum_{i=1}^d Y_i)\}$
$q^-$	Borne inférieure (infimum) de $q$	$q_{\mathcal{F}}^- := \inf\{q(\sum_{i=1}^d Y_i)\}$
$\mathbb{I}$	Variable indicatrice	$\mathbb{I} = 1$ , if $(x_{i_1}, x_{i_2}, \dots, x_{i_d}) \in \mathcal{F}_N$ ( $i = 1, 2, \dots, N$ ).
[.]	Partie entière	
$Z_i, T_i$	Variables aléatoires dans la partie non fixe, la partie fixe	$Z_i = F_{X_i   (X_1, X_2, \dots, X_d) \in \mathcal{U}}^{-1}(U)$ , $T_i = F_{X_i   (X_1, X_2, \dots, X_d) \in \mathcal{F}}^{-1}(U)$
$N$	$N$ vecteurs d'observations (scénarios)	
$\ell_f, \ell_u$	Nombre d'éléments dans $\mathcal{F}, \mathcal{U}$	$N = \ell_f + \ell_u$
$p_f, p_u$	Probabilité d'être dans $\mathcal{F}, \mathcal{U}$	$p_f = \ell_f/N$ , $p_u = \ell_u/N$
$p$	Niveau de probabilité	$Var_p$ , $k := N(1 - p)$
$\bar{s}$	Moyenne de $s$	
$\Omega$	Espace de probabilités	
$q_\gamma$	Quantile de la variable aléatoire normale standard à $\gamma$	$\mathcal{F} = [q_\beta, q_{1-\beta}]^d \subset \mathbb{R}^d$
$\beta$	Niveau de confiance	
$\rho$	Coefficient de corrélation	

$\forall$	« pour tous »	$\mathcal{F} := \{(x_{1i}, \dots, x_{di}) \in \mathbb{R}^d \mid \forall j \in \{1, 2, \dots, d\}, \hat{f}_j(x_{ji}) \geq \varepsilon\}$
$\hat{f}$	Densités empiriques	
$\phi, \Phi$ $\mathcal{N}(0,1)$	Fonction de densité, fonction de répartition de normale standard $\mathcal{N}(0,1)$	
$\cap$	Intersection	$\mathcal{F} = \bigcap_{k=1}^d \{q_\beta \leq X_k \leq q_{1-\beta}\}$

## Bibliographie

- Kjersti Aas, Claudia Czado, Arnoldo Frigessi et Henrik Bakken. « Pair-copula constructions of multiple dependence », *Insurance: Mathematics and Economics*, vol. 44, n° 2, p. 182–198, 2009.
- Philippe Artzner, Freddy Delbaen, Jean-Marc Eber et David Heath. « Coherent measures of risk », *Mathematical Finance*, vol. 9, n° 3, p. 203–228, 1999.
- Pauline Barrieu et Giacomo Scandolo. « Assessing Financial Model Risk », *European Journal of Operational Research*, vol. 242, n° 2, p. 546–556, 2015.
- Nicole Bäuerle et Alfred Müller. « Stochastic Orders and Risk Measures: Consistency and Bounds », *Insurance: Mathematics and Economics*, vol. 38, n° 1, p. 132–148, 2006.
- Tim Bedford et Roger M. Cooke. « Probability Density Decomposition for Conditionally Dependent Random Variables Modeled by Vines », *Annals of Mathematics and Artificial Intelligence*, vol. 32, n° 1-4, p. 245–268, 2001.
- Tim Bedford et Roger M. Cooke. « Vines—a new graphical model for dependent random variables », *The Annals of Statistics*, vol. 30, n° 4, p. 1031–1068, 2002.
- Carole Bernard et Steven Vanduffel. « A New Approach to Assessing Model Risk in High Dimensions », *Journal of Banking and Finance*, vol. 58, p. 166–178, 2015a.
- Carole Bernard et Steven Vanduffel. « Quantile of a Mixture », *Dependence Modeling*, à paraître, 2015b.
- Carole Bernard, Xiao Jiang et Steven Vanduffel. « A note on ‘Improved Fréchet Bounds and Model-Free Pricing of Multi-Asset Options’ by Tankov (2011) », *Journal of Applied Probability*, vol. 49, n° 3, p. 866–875, 2012.
- Carole Bernard, Yuntao Liu, Niall MacGillivray et Jinyuan Zhang. « Bounds on Capital Requirements for Bivariate Risk with Given Marginals and Partial Information on the Dependence », *Dependence Modeling*, vol. 1, p. 37–53, 2013a.
- Carole Bernard, Ludger Rüschendorf et Steven Vanduffel. « Value-at-Risk Bounds with Variance Constraints », *Journal of Risk and Insurance*, à paraître, 2013b.
- Carole Bernard, Phelim Boyle et Steven Vanduffel. « Explicit Representation of Cost-Efficient Strategies », *Finance*, vol. 35, n° 2, p. 5–55, 2014.
- Carole Bernard, Jit-Seng Chen et Steven Vanduffel. « Optimal Portfolios under Worst-Case Scenarios », *Quantitative Finance*, vol. 14, n° 4, p. 657–671, 2014.
- Carole Bernard, Michel Denuit et Steven Vanduffel. « Value-at-Risk Bounds with High-Order Moments », document de travail, 2014a.
- Carole Bernard, Xiao Jiang et Ruodu Wang. « Risk Aggregation with Dependence Uncertainty », *Insurance: Mathematics and Economics*, vol. 54, n° 1, p. 93–108, 2014b.
- Carole Bernard, Ludger Rüschendorf, Steven Vanduffel et Jing Yao. « How Robust is the Value-at-Risk of Credit Risk Portfolios? », *European Journal of Finance*, à paraître, 2015.

- BSIF. « Étude d'impact quantitative No. 4 : Généralités – agrégation et diversification – Information supplémentaire », document en ligne, 2014.
- Andrew Chernih, Luc Henrard et Steven Vanduffel. « Reconciling Credit Correlations », *The Journal of Risk Model Validation*, vol. 4, n° 2, p. 47–64, 2010.
- Ka-Chun Cheung et Steven Vanduffel. « Bounds for Sums of Random Variables when the Marginal Distributions and the Variance of the Sum are Given », *Scandinavian Actuarial Journal*, vol. 13, n° 2, p. 103–118, 2013.
- Comité de Bâle. « Developments in Modelling Risk Aggregation », Instance conjointe, Banque des règlements internationaux, 2010a.
- Comité de Bâle. « Bâle III: dispositif réglementaire mondial visant à renforcer la résilience des établissements et systèmes bancaires », texte préliminaire, 2010b.
- Comité de Bâle. « Fundamental review of the trading book », document de consultation, mai 2012.
- Comité de Bâle. « Fundamental review of the trading book: A revised market risk framework », document de consultation, octobre 2013.
- Roger M. Cooke et Carolyn Kousky. « The Limits of Securitisation: Micro correlations, Fat Tails, and Tail Dependence », *Rethinking Risk Measurement and Reporting Volume I: Uncertainty, Bayesian Analysis, and Expert Judgement*, série de publications sur les risques, Londres, p. 273–294, 2010.
- Roger M. Cooke, Carolyn Kousky et Harry Joe. « Micro Correlations and Tail Dependence », *Dependence Modeling: Vine Copula Handbook*, chap. 5, p. 89–112, 2011.
- Claudia Czado. « Pair-copula constructions of multivariate copulas », *Copula Theory and Its Applications*, Springer, p. 93–109, 2010.
- Michel Denuit, Christian Genest et Étienne Marceau. « Stochastic bounds on sums of dependent risks », *Insurance: Mathematics and Economics*, vol. 25, n° 1, p. 85–104, 1999.
- Michel Denuit, Jan Dhaene, Marc Goovaerts et Rob Kaas. « Actuarial Theory for Dependent Risks: Measures, Orders and Models », Wiley Online Library, 2005.
- Jan Dhaene, Michel Denuit, Marc Goovaerts, Rob Kaas et David Vyncke. « The Concept of Comonotonicity in Actuarial Science and Finance: Theory », *Insurance: Mathematics and Economics*, vol. 31, n° 1, p. 3–33, 2002.
- Jan Dhaene, Steven Vanduffel, Marc Goovaerts, Rob Kaas, Qihe Tang et David Vyncke. « Risk Measures and Comonotonicity: A Review », *Stochastic Models*, vol. 22, p. 573–606, 2006.
- Jan Dhaene, Roger J.A. Laeven, Steven Vanduffel, Gregory Darkiewicz et Marc J. Goovaerts. « Can a coherent risk measure be too subadditive? », *Journal of Risk and Insurance*, vol. 75, n° 2, p. 365–386, 2008.
- Jan Dhaene, Michel Denuit et Steven Vanduffel. « Correlation Order, Merging and Diversification », *Insurance: Mathematics and Economics*, vol. 45, n° 3, p. 325–332, 2009.

- Jan Dhaene, Andreas Tsanakas, Emiliano A. Valdez et Steven Vanduffel. « Optimal Capital Allocation Principles », *Journal of Risk and Insurance*, vol. 79, n° 1, p. 1–28, 2012.
- Darrell Duffie et Kenneth J. Singleton. « *Credit Risk: Pricing, Measurement, and Management* », Princeton University Press, 2012.
- Paul Embrechts et Marius Hofert. « Risk Measures and Dependence Modeling », *Handbook of Insurance*, Springer, p. 135–165, 2013.
- Paul Embrechts et Marius Hofert. « Statistics and Quantitative Risk Management for Banking and Insurance », *Annual Review of Statistics and Its Application*, vol. 1, p. 493–514, 2014.
- Paul Embrechts et Giovanni Puccetti. « Bounds for functions of dependent risks », *Finance and Stochastics*, vol. 10, n° 3, p. 341–352, 2006.
- Paul Embrechts et Giovanni Puccetti. « Risk Aggregation », *Copula Theory and Its Applications*, Springer, p. 111–126, 2010a.
- Paul Embrechts et Giovanni Puccetti. « Bounds for the sum of dependent risks having overlapping marginals », *Journal of Multivariate Analysis*, vol. 101, n° 1, p. 177–190, 2010b.
- Paul Embrechts, Alexander McNeil et Daniel Straumann. « Correlation: Pitfalls and Alternatives », *Risk Magazine*, Londres, vol. 12, p. 69–71, 1999.
- Paul Embrechts, Alexander McNeil et Daniel Straumann. « Correlation and Dependence in Risk Management: Properties and Pitfalls », *Risk management: Value-at-Risk and beyond*, M. Dempster et H. Moffatt (éditeurs), p. 176–223, 2002.
- Paul Embrechts, Giovanni Puccetti et Ludger Rüschendorf. « Model uncertainty and VaR aggregation », *Journal of Banking & Finance*, vol. 37, n° 8, p. 2750–2764, 2013.
- Paul Embrechts, Giovanni Puccetti, Ludger Rüschendorf, Ruodu Wang et Antonela Beleraj. « An Academic Response to Basel 3.5 », texte préliminaire, 2014.
- H. Föllmer et A. Schied. « Coherent and convex risk measures », *Encyclopedia of Quantitative Finance*, R. Cont (éditeur), New York, John Wiley & Sons, p. 355–363, 2010.
- Tilmann Gneiting. « Making and Evaluating Point Forecasts », *Journal of the American Statistical Association*, vol. 106, n° 494, p. 746–762, 2011.
- Chris C. Heyde et Steven G. Kou. « On the controversy over tailweight of distributions », *Operations Research Letters*, vol. 32, n° 5, p. 399–408, 2004.
- Marius Hofert. « A stochastic representation and sampling algorithm for nested Archimedean copulas », *Journal of Statistical Computation and Simulation*, vol. 82, n° 9, p. 1239–1255, 2012.
- Marius Hofert et David Pham. « Densities of nested Archimedean copulas », *Journal of Multivariate Analysis*, vol. 118, p. 37–52, 2013.
- Harry Joe. « *Multivariate Models and Multivariate Dependence Concepts* », vol. 73, CRC Press, 1997.
- Philippe Jorion. « *Value-at-Risk: The New Benchmark for Managing Financial Risk* », McGraw-Hill, 3<sup>e</sup> éd., 2006.

- Steven Kou, Xianhua Peng et Chris C. Heyde. « External Risk Measures and Basel Accords », *Mathematics of Operations Research*, vol. 38, n° 3, p. 393–417, 2013.
- Carolyn Kousky et Roger Cooke. « Explaining the Failure to Insure Catastrophic Risks », *The Geneva Papers on Risk and Insurance-Issues and Practice*, vol. 37, n° 2, p. 206–227, 2012.
- Carolyn Kousky et Roger M. Cooke. « The Unholy Trinity: Fat Tails, Tail Dependence, and Micro-Correlations », *Resources for the Future*, document de travail, p. 9–36, 2009.
- Dorota Kurowicka et Roger M. Cooke. « Uncertainty Analysis with High Dimensional Dependence Modelling », John Wiley & Sons, 2006.
- Laurence Lescourret et Christian Y. Robert. « Extreme dependence of multivariate catastrophic Losses », *Scandinavian Actuarial Journal*, vol. 2006, n° 4, p. 203–225, 2006.
- Alexander J. McNeil, Rüdiger Frey et Paul Embrechts. « Quantitative Risk Management: Concepts, Techniques, and Tools », Princeton University Press, 2010.
- Isaac Meilijson et Arthur Nádas. « Convex Majorization with an Application to the Length of Critical Paths », *Journal of Applied Probability*, p. 671–677, 1979.
- Alfred Müller et Dietrich Stoyan. « Comparison Methods for Stochastic Models and Risks », Chichester, John Wiley & Sons, 2002.
- Roger B. Nelsen. « An Introduction to Copulas », Springer, 2007.
- Ostap Okhrin, Yarema Okhrin et Wolfgang Schmid. « On the structure and estimation of hierarchical Archimedean copulas », *Journal of Econometrics*, vol. 173, n° 2, p. 189–204, 2013.
- Harry H. Panjer. « Measurement of Risk, Solvency Requirements and Allocation of Capital Within Financial Conglomerates », Institute of Insurance and Pension Research, Université de Waterloo, 2001.
- Giovanni Puccetti et Ludger Rüschendorf. « Computation of sharp bounds on the distribution of a function of dependent risks », *Journal of Computational and Applied Mathematics*, vol. 236, n° 7, p. 1833–1840, 2012a.
- Giovanni Puccetti et Ludger Rüschendorf. « Asymptotic equivalence of conservative VaR and ES-based capital charges », document de travail, 2012b.
- Ludger Rüschendorf. « Random Variables with Maximum Sums », *Advances in Applied Probability*, p. 623–632, 1982.
- Ludger Rüschendorf. « Bounds for Distributions with Multivariate Marginals », recueil de notes de cours, collection de monographies, p. 285–310, 1991.
- Cornelia Savu et Mark Trede. « Hierarchies of Archimedean copulas », *Quantitative Finance*, vol. 10, n° 3, p. 295–304, 2010.
- Abe Sklar. « Fonctions de répartition à n dimensions et leurs marges », vol. 8, 1959.
- Peter Tankov. « Improved Fréchet bounds and model-free pricing of multi-asset options », *Journal of Applied Probability*, vol. 48, n° 2, p. 389–403, 2011.

Andreas Tsanakas. « To split or not to split: Capital allocation with convex risk measures », *Insurance: Mathematics and Economics*, vol. 44, n° 2, p. 268–277, 2009.

Andreas Tsanakas et Evangelia Desli. « Measurement and Pricing of Risk in Insurance Markets », *Risk Analysis*, vol. 25, n° 6, p. 1653–1668, 2005.

Bin Wang et Ruodu Wang. « The complete mixability and convex minimization problems with monotone marginal densities », *Journal of Multivariate Analysis*, vol. 102, n° 10, p. 1344–1360, 2011.

Ruodu Wang, Liang Peng et Jingping Yang. « Bounds for the Sum of Dependent Risks and Worst Value-at-Risk with Monotone Marginal Densities », *Finance and Stochastics*, vol. 17, n° 2, p. 395–417, 2013.

Shaun Wang. « A class of distortion operations for pricing financial and insurance risks », *Journal of Risk and Insurance*, vol. 67, n° 10, p. 15–36, 2000.

Johanna F. Ziegel. « Coherence and elicibility », *Mathematical Finance*, 2014.